2009

ΤΕΙ ΚΡΗΤΗΣ – ΗΡΑΚΛΕΙΟ

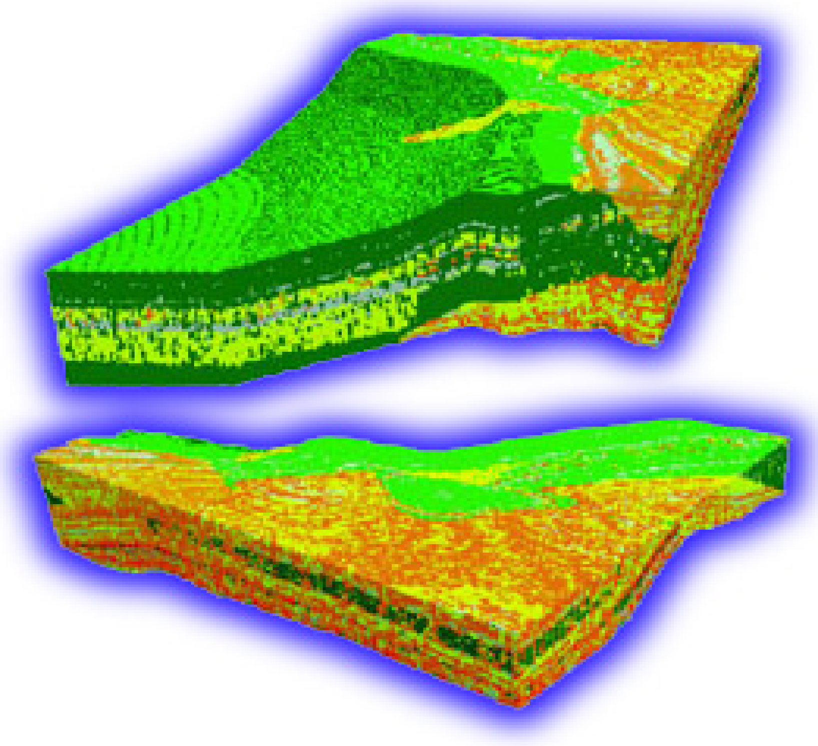
ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΠΟΛΥΜΕΣΩΝ

ΝΙΚΟΣ ΓΙΑΝΝΟΠΟΥΛΟΣ
ΑΜ 312

Πέμπτη, 10 Δεκεμβρίου 2009

ΤΕΧΝΟΛΟΓΙΚΟ
ΕΚΠΑΙΔΕΥΤΙΚΟ
ΙΔΡΥΜΑ ΚΡΗΤΗΣ

ΜΙΑ ΕΙΣΑΓΩΓΗ ΣΤΗ ΓΕΩΣΤΑΤΙΣΤΙΚΗ ΒΑΡΙΟΓΡΑΜΜΑΤΑ ΚΑΙ ΜΕΘΟΔΟΙ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ



Περίληψη

Η εργασία είναι μια προσπάθεια να εισάγει τον αναγνώστη στις βασικές έννοιες τις Γεωστατιστικής με μια σύντομη περιγραφή στην εκτίμηση αποθεμάτων και τις κύριες μεθόδους που χρησιμοποιούνται για αυτήν, τα βασικά στοιχεία Στατιστικής που έχουν σχέση με την Γεωστατιστική, καθώς και μια εισαγωγή στη Βαριογραφία και τα Βαριογράμματα και περιγραφή μερικών εφαρμογών, όπως εκφράζονται στην σχετική αρθρογραφία. Ο στόχος είναι η ενημέρωση και κατάρτιση στη Γεωστατιστική και τις μοντέρνες μεθόδους που χρησιμοποιούνται.

Στο ερώτημα που τίθεται ως προκείμενη του θέματος της παρούσης - ποιο είναι το αντικείμενο της Γεωγραφίας - η απάντηση φυσικά είναι ο χώρος και οι κανόνες που διέπουν το σύνολο των δραστηριοτήτων –κυρίως των ανθρωπίνων- σε σχέση με αυτόν.

Μια πιο συγκεκριμένη απάντηση θα μπορούσε να είναι ότι ουσιαστικά η Γεωγραφία ζητά να εξηγήσει τις διαδικασίες και τα πρότυπα διάταξης στη χωρική οργάνωση των ανθρώπινων δραστηριοτήτων.

Ένα βασικό ερώτημα που ξεχωρίζει τη γεωγραφία από τις άλλες επιστήμες είναι το πώς αντικείμενα, φαινόμενα, υπηρεσίες κλπ κατανέμονται στο χώρο. Το ερώτημα αυτό αποτελεί βάση για άλλα ερωτήματα όπως «πως προήλθε μια χωροθέτηση που παρατηρείται», ερωτήσεις που υπονοούν πως όσα περικλείονται σε μια περιοχή έχουν συγκεκριμένη διάταξη.

Η διάταξη με την οποία τα αντικείμενα, φαινόμενα, κλπ κατανέμονται στην επιφάνεια της γης ονομάζεται χωρική ανάλυση και αναλύεται στις τρεις παρακάτω βασικές χωρικές έννοιες:

Σχήμα: Ένα δισδιάστατο (ή τρισδιάστατο) χαρακτηριστικό μιας χωρικής τακτοποίησης που ορίζεται από μια κλειστή καμπύλη (ή επιφάνεια) η οποία οριοθετεί μια συλλογή αντικειμένων και παρέχει τη μέτρηση της επιφάνειας (ή υπερεπιφάνειας) κατανομής τους.

Χωρική Διασπορά: Είναι ένα μονοδιάστατο χαρακτηριστικό μιας χωρικής τακτοποίησης που μετρά την απόσταση μεταξύ ενός συνόλου αντικειμένων σε σχέση με ένα συγκεκριμένο σχήμα μιας δοσμένης επιφάνειας(περιοχής).

Χωρικό πρότυπο: Είναι ένα μηδενικής διάστασης χαρακτηριστικό μιας χωρικής τακτοποίησης που περιγράφει η θέση στο χώρο ενός συνόλου αντικειμένων σε σχέση με τα άλλα. Συνεπώς είναι χαρακτηριστικό μιας χωρικής διασποράς σημείων χωροθετημένων σε καθορισμένο σχήμα δεδομένης πυκνότητας.[a]

Κάθε χωρική μονάδα παρατήρησης (σημείο, γραμμή, επιφάνεια, φατνίο) έχει κάποιες ιδιότητες-χαρακτηριστικά τα οποία αποτυπώνονται και αναφορικά με το επίπεδο καταγραφής τους διακρίνονται σε τέσσερις κλίμακες μέτρησης. Ιεραρχημένα, με βάση τις δυνατότητές τους ως εργαλεία ανάλυσης χώρου και κυρίως την ακρίβειά τους είναι (σε επίπεδα):

Το ονομαστικό που περισσότερο κατηγοριοποιεί τα αντικείμενα πάρα μετρά τις ιδιότητές τους. Έχει αναγνωριστική και όχι μαθηματική δύναμη ταξινομώντας ουσιαστικά την καταγραφή μιας μεταβλητής σε ονομαστικές κατηγορίες

Το διατεταγμένο που συγκρίνει κατηγορίες πχ: μεγαλύτερου ή μικρότερου. Ταξινομεί λοιπόν αντικείμενα αναφορικά με την τιμή τους με μια πραγματικά μετρήσιμη αριθμητική κλίμακα. Η σύγκριση μπορεί να είναι διαφόρων ειδών.

Το ιεραρχικό που εκφράζει ποσοτικά μια ιεραρχική σχέση. Δηλαδή γνωρίζει ότι το αντικείμενο Α είναι πχ: μεγαλύτερο του Β και επίσης ξέρει και πόσο μεγαλύτερο.

Το αναλογικό έχει το πλεονέκτημα μιας φυσικής αφετηρίας δλδ ξεκινά από ένα σταθερό μηδέν. Αυτό βοηθά στη μέτρηση διαφορών και απόλυτων τιμών καθώς και στην εύρεση αναλογικής σχέσης ανάμεσα στις τιμές δύο μεταβλητών. Το βάρος η μάζα το μήκος μπορούν να μετρηθούν με αναλογική κλίμακα.

Κάθε επίπεδο έχει ποικίλους περιορισμούς γι’ αυτό κι ο ερευνητής πρέπει να αναζητά υψηλότερα επίπεδα μέτρησης που θα τον οδηγήσουν σε καλύτερη επιλογή αναλυτικών τεχνικών.

Συνεπώς οι χωρικές παρατηρήσεις διαφοροποιούνται ανάλογα με την κλίμακα μέτρησης των χαρακτηριστικών τους. Ένα παράδειγμα είναι ότι τα σημεία μπορούν να διαφοροποιηθούν ως προς το είδος της ασθένειας (ονομαστικό), τη διάταξη: χωριό, κωμόπολη, πόλη (διατάξεως) ή τέλος με μέτρηση του ύψους των ανθρώπων σε κάθε σημείο (αναλογικό).[b]

Βασικά Στοιχεία Στατιστικής- Γεωστατιστική και Στατιστική

Η Γεωστατιστική είναι κλάδος της Εφαρμοσμένης Στατιστικής και ασχολείται με φαινόμενα που μεταβάλλονται στο χώρο – έτσι βασίζεται στην θεωρία πιθανοτήτων και τη Στατιστική.

Βασικά Στοιχεία Στατιστικής (σε σχέση με την Γεωστατιστική)

Ο χώρος είναι η συνολική μάζα ή ο όγκος του φαινομένου που μας ενδιαφέρει ως πηγή δεδομένων.

Ο πληθυσμός είναι το σύνολο όλων των πιθανών στοιχείων – οντοτήτων, εφόσον ορίσουμε αυτές, που μπορούμε να πάρουμε από ένα χώρο.

Ένα γεγονός είναι μια συλλογή από δειγματοληπτικά στοιχεία που ορίζουν ένα αποτέλεσμα στο οποίο θα πρέπει να δώσουμε μια πιθανότητα.

Σε κάθε γεγονός Α μπορούμε να δώσουμε ένα αριθμό Pr(A) που λέγεται ‘πιθανότητα του γεγονότος Α’ ($0 \leq Pr(A) \leq 1$, $Pr(\Omega) = 1$, $Pr(\varnothing) = 0$ και έτσι η πιθανότητα του συμπληρωματικού ενός γεγονότος είναι $Pr(A') = 1 - Pr(A)$). Για του συμβολισμούς και την άλγεβρα δείτε το Πίνακα 1 (περιληπτικά):

Πίνακας 1. **Συμβολισμός Συνόλων**

Συμβολισμός	Ερμηνεία
$a \in A$, $a \notin A$	Το a είναι (δεν είναι) στοιχείο του A
S, Ω, U	Ο χώρος, ο δειγματοληπτικός χώρος
$A = \Omega$	Το A είναι βέβαιο
$\varnothing$	Κενό σύνολο (χωρίς στοιχεία)
$A = \varnothing$	Το A είναι αδύνατο

$A \subset B$, $A \subseteq B$	Το A είναι υποσύνολο του B
A = B	Το A ισούται του B
$A \cup B$	Η ένωση των A και B
$A \cap B$	Η τομή του A και B
A’	Το συμπληρωματικό του A, το μη A
$A \cap B = \varnothing$	Το A και B δεν έχουν κοινά στοιχεία

Κανόνας Άθροισης και Αμοιβαία Αποκλειστικότητα. Η πιθανότητα να συμβεί το ένα ή το άλλο από δύο αμοιβαία αποκλειστικά γεγονότα A και B είναι το άθροισμα της πιθανότητας να συμβεί το A και της πιθανότητας να συμβεί το B. Γενικά ισχύει:

$$\Pr(A \cup B \cup C...) = \Pr(A) + \Pr(B) + \Pr(C)...$$

Εάν τα γεγονότα δεν είναι αμοιβαία αποκλειστικά τότε:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

Κανόνας Πολλαπλασιασμού και Ανεξαρτησία. Δυο γεγονότα A και B είναι ανεξάρτητα εάν η πιθανότητα να συμβούν και τα δύο είναι το προϊόν των αντιστοίχων πιθανοτήτων τους. Δηλαδή:

$$\Pr(A \cap B) = \Pr(A) \times \Pr(B) \Rightarrow A, B \text{ ανεξάρτητα}$$

Πιθανότητα υπό Συνθήκη. Εάν A και B είναι δυο γεγονότα, η πιθανότητα να συμβεί το A εάν έχει ήδη συμβεί το B και συμβολίζεται:

$$\Pr(A \mid B)^{\text{c}}$$

Τυχαίες Μεταβλητές και Κατανομές

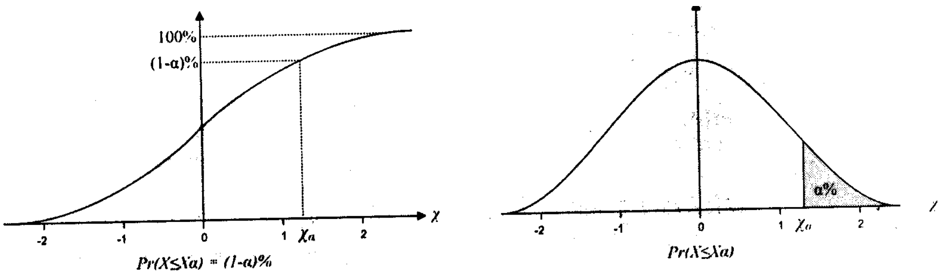
Τυχαία μεταβλητή (ΤΜ) που συμβολίζεται με το γράμμα *X,* είναι μια μεταβλητή που παίρνει αριθμητικές τιμές, που συμβολίζονται από μικρά γράμματα: $\{x_1, x_2, x_3, \ldots x_n\}$, σύμφωνα με το αποτέλεσμα ενός πειράματος στο οποίο έχουμε αποδώσει σχετικές πιθανότητες.

Συνάρτηση Αθροιστικής Κατανομής – Κάθε τυχαία μεταβλητή *X* ορίζεται από την συνάρτηση αθροιστικής κατανομής:

$$F(x) = \Pr[X \le x] \quad -\infty < x < \infty$$

Αν f(x) είναι η παράγωγος της συνάρτησης αθροιστικής κατανομής (η πυκνότητα της πιθανότητας στο σημείο *x*) τότε:

$$F(x) = \int_{-\infty}^{\infty} f(x)dx$$



Σχήμα 5. Συναρτήσεις αθροιστικής κατανομής και πυκνότητας πιθανότητας

Ιστόγραμμα – για να εκτιμηθεί η f(x) κατασκευάζουμε ένα ιστόγραμμα.

Σχετικά στοιχεία (Ροπές, Διακύμανση, Διασπορά κλπ)

Ροπές και αναμενόμενη τιμή

Η αναμενόμενη τιμή E[x] μιας τυχαίας μεταβλητής *X* βρίσκεται από την μέση τιμή της συνάρτησης για όλες τις πιθανές τιμές της μεταβλητής – ιδανική τιμή. *Παράδειγμα 11: Η πυκνότητα μιας τυχαίας μεταβλητής* $\mathbf{X}$ *είναι*

$$f(x) = \begin{cases} x/2, & 0 < x < 2 \\ 0, & \text{αλλιώς} \end{cases}$$

Υπολογίστε τη μέση τιμή της $\mathbf{X}$.

Έχουμε

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^2 x\frac{1}{2}xdx = \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 = \frac{4}{3}.$$

Η ροπή μιας δύναμης στη φυσική είναι το μέγεθος αυτής της δύναμης επί την απόσταση μεταξύ του σημείου εφαρμογής της και τον άξονα περιστροφής. Στην στατιστική ορίζουμε στατιστικές ροπές γύρω από οποιονδήποτε άξονα – δηλαδή γύρω από οποιαδήποτε επιθυμητή τιμή του *X.* Η *k* ροπή του *X* συμβολίζεται συνήθως με μ'$_k$ και ορίζεται ως:

$$\mu'_k = E[X^k] = \int_{-\infty}^{\infty} x^k f(x)dx$$

Έτσι η πρώτη ροπή είναι:

$\mu'_1 = m = E[x]$ που είναι ο μέσος

Έτσι ο μέσος μπορεί να θεωρηθεί ως το κέντρο βάρους της κατανομής.

> Γενικά, ο χωρικός μέσος δεν έχει έννοια ως αριθμητική τιμή, έχει νόημα μόνο ως σημείο στο χάρτη. Η θέση του χωρικού μέσου είναι «συνθετική». Δύο διαφορετικές κατανομές μπορούν να «δώσουν ίδιο χωρικό μέσο.

Μπορούμε να επίσης να υπολογίσουμε τις ροπές γύρω από οποιοδήποτε κέντρο (για παράδειγμα τον μέσο):

$$E[X - \mu'_1] = \int_{-\infty}^{\infty} (x - \mu'_1)^k f(x)dx$$

Η **Διακύμανση** είναι πολύ σημαντική στη στατιστική (μέτρο διασποράς γύρω από κάποια κεντρική τιμή) και στην γεωστατιστική. Είναι η δεύτερη ροπή γύρω από τον μέσο:

$$\begin{aligned}
&Var(X) = \sigma^2 = \mathrm{E}[X - m]^2 \\
&= \int_{-\infty}^{\infty} (x-m)^2 f(x)dx \\
&= \int_{-\infty}^{\infty} x^2 f(x)dx - m^2 \\
&= E[\mathrm{X}^2] - \mathrm{E}[X]^2
\end{aligned}$$

Στην πράξη η εκτίμηση της διακύμανσης Var(X) ή σ^2 πραγματοποιείται από το στατιστικό s^2:

$$s^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - m)^2$$

$$s^2 = \frac{1}{N}\left[\sum_{i=1}^{N}(x_i)^2\right] - m^2$$

Όπου m (ο αριθμητικός μέσος) εκτιμάται από τον μέσο των δειγμάτων $\overline{x}$.

Διακύμανση είναι μια στατιστική παράμετρος που μετρά την διασπορά. Εάν οι περισσότερες τιμές βρίσκονται κοντά στον μέσο, η διακύμανση θα είναι μικρή. Εάν οι τιμές απλώνονται σε ένα μεγάλο εύρος, η διακύμανση θα είναι μεγάλη.

Για να μετρήσουμε την διασπορά στις ίδιες μονάδες με τις ίδιες τις τιμές χρησιμοποιούμε την τετραγωνική ρίζα της διακύμανσης, **την τυπική απόκλιση:**

$$\sigma = \sqrt{\mathrm{var}[X]}$$

Για να συγκρίνουμε κατανομές μια παράμετρος που χρησιμοποιείται είναι **ο συντελεστής μεταβλητότητας** (CV – Coefficient of Variation):

$$CV = \frac{\sigma}{m}$$

Όταν συγκρίνουμε δύο κατανομές αυτή με τον υψηλότερο CV είναι πιο διασπαρμένη από αυτήν με τον χαμηλότερο CV. Τιμές του CV αρκετά μεγαλύτερες από 1,0 είναι τυπικές για δείγματα περιεκτικότητας από πολλά ορυχεία ουρανίου, χρυσού ή κασσίτερου. Κοιτάσματα βασικών μετάλλων έχουν CV από 1,0. Κοιτάσματα χρυσού έχουν CV μεγαλύτερο από 2,0. Όσο μεγαλύτερος είναι ο CV τόσο πιο δύσκολη είναι η γεωστατιστική ανάλυση και η εκτίμηση συνολικά.

Συνδιακύμανση και Συσχετισμός

Μια σημαντική ιδιότητα της διακύμανσης είναι η **συνδιακύμανση:**

Μεταξύ των μεταβλητών *X* και *Y* η συνδιακύμανση *Cov[XY]* καταλήγει σε περίπτωση ανεξαρτησίας μεταξύ των μεταβλητών στην απλοποιημένη μορφή:

Cov(XY)=E[XY]-E[X]•E[Y]

Κατανομές

Μερικές κατανομές έχουν ιδιαίτερη χρήση στην γεωστατιστική πρακτική και την εκτίμηση αποθεμάτων:

Gaussian Κατανομή (λέγεται και κανονική κατανομή)

Η Gaussian κατανομή έχει την συνάρτηση πιθανοτήτων:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$$

όπου *m* είναι ο μέσος και σ^2 η διακύμανση.

Λογαριθμική Κατανομή (για κατανομές που τείνουν να είναι θετικά *λοξές*.

$$f(y) = \frac{\log e}{\gamma\sigma\sqrt{2\pi}} \exp -\frac{1}{2}\left(\frac{\log\gamma - \log y}{\sigma}\right)^2$$

όπου γ είναι ο διάμεσος της κατανομής: [d]

Στατιστικός Έλεγχος

1) Διατυπώνεται η μηδενική υπόθεση Ho, που συνήθως είναι η αντίθετη από τη διερευνόμενη υπόθεση H1.
2) Επιλέγεται το κριτήριο (δειγματοληπτική κατανομή) που θα χρησιμοποιηθεί, ανάλογα με τα χαρακτηριστικά των κατανομών που συγκρίνονται και τις δυνατότητες που διαθέτει κάθε κριτήριο.
3) Ορίζεται το επίπεδο εμπιστοσύνης (100-α)% που θεωρείται αποδεκτό. Συνήθως χρησιμοποιείται α = 1, 5 ή 10.
4) Υπολογίζεται η τιμή που αντιστοιχεί στο κριτήριο που έχει επιλεγεί. Η παράμετρος αυτή υπολογίζεται από τα πραγματικά δεδομένα και πιο συγκεκριμένα:
5) Υπολογίζεται η πιθανότητα για να εμφανιστεί η υπολογισμένη τιμή του κριτηρίου, αν η Ho είναι σωστή. Οι τιμές αυτές υπάρχουν συνήθως σε πίνακες.
6) Συγκρίνεται η πιθανότητα να είναι σωστή η Ho με το επίπεδο εμπιστοσύνης που έχει επιλεγεί. Αν η πιθανότητα είναι μικρότερη από το επίπεδο εμπιστοσύνης, τότε δεν γίνεται

δεκτή η Ho, αλλά η H1. Αν είναι μεγαλύτερη από το επίπεδο εμπιστοσύνης, τότε συμπεραίνεται ότι δεν είναι δυνατό να απορριφθεί η Ho χωρίς όμως αυτό να σημαίνει αναγκαστικά ότι γίνεται αποδεκτή.

Κριτήριο στατιστικού ελέγχου Για το κριτήριο $\mathbf{x^2}$ η τιμή που υπολογίζεται είναι η μέτρηση των διαφορών ανάμεσα στις κατανομές. Δηλαδή:

$$x^2 = \sum_{i=1}^{n} \frac{\left(\Pi_r - A_r\right)^2}{A_r}$$

όπου: Πr = παρατηρούμενος αριθμός φατνίων με r σημεία.

Ar = αναμενόμενος αριθμός φατνίων με r σημεία.

Για την περίπτωση του κριτηρίου **t**, η τιμή υπολογίζεται:

$$t = \frac{\Pi - E\left(\Pi\right)}{\sigma_{\pi}\sqrt{n}}$$

όπου: **Π** = Παρατηρούμενη τιμή

Ε(Π) = Αναμενόμενη τιμή του δείκτη

$\boldsymbol{\sigma_{\pi}}$ = το τυπικό σφάλμα του δείκτη π

n = ο αριθμός των παρατηρήσεων (σημείων)

Ενώ για την περίπτωση της κανονικής κατανομής **Z**, η τιμή δίνεται: [e]

$$Z = \frac{\Pi - E\left(\Pi\right)}{\sigma_{\pi}}$$

Χωρομεταβλητές και Βαριογράμματα

Μοντέλα Πιθανοτήτων και Δειγματοληπτικές Μέθοδοι

Είναι γνωστό ότι οι γεωγραφικές μεταβλητές (περιεκτικότητα, πάχος κλπ) τείνουν να εμφανίζουν σημαντική διασπορά μικρής κλίμακας. **Αυτές οι μεταβολές είναι τόσο πολύπλοκες που δεν μπορούν να αποδοθούν σε όρους συνηθισμένων μαθηματικών συναρτήσεων. Αυτό μας αναγκάζει να υιοθετήσουμε μοντέλα πιθανοτήτων για την ανάλυση των χωρικών φαινομένων.**

Χωρομεταβλητές και Τυχαίες Συναρτήσεις

Ο όρος *Χωρομεταβλητή* (ΧΜ) συνδυάζει δύο στοιχεία:

- Ένα τυχαίο στοιχείο: αυτό περιλαμβάνει τοπικές και μικρές κλίμακας ανωμαλίες
- Ένα δομημένο ή *χωρικό* στοιχείο: αυτό περιλαμβάνει μεγάλης κλίμακας τάσεις των φαινομένων

Τυχαίες Συναρτήσεις – μια τυχαία συνάρτηση χαρακτηρίζεται από την κοινή κατανομή μια ομάδας τυχαίων μεταβλητών:

$Z(xi),Z(x2),\dots Z(xk)$ για όλα τα σημεία $x1,x2,\dots xk$

Στασιμότητα, Αυστηρή Στασιμότητα, Εσωτερική Υπόθεση

Οι επιπρόσθετες παραδοχές και υποθέσεις που χρειάζονται να γίνουν σε μια πραγματική περίπτωση ευτυχώς μειώνουν τον αριθμό των παραμέτρων από τις οποίες εξαρτάται η Τυχαία Συνάρτηση.

Η στασιμότητα έχει να κάνει με την υπόθεση ότι σε κάποια εφαρμογή μια μεταβλητή μπορεί να θεωρηθεί στάσιμη κατά την μετατόπιση.

Στην αυστηρή της έννοια η στασιμότητα απαιτεί ότι όλες οι ροπές της κατανομής να είναι αμετάβλητες κατά την μετατόπιση. Αυτό είναι δύσκολο στην πράξη.

Συχνά αρκεί να ισχύει αυτό που λέγεται **Αδύναμη ή 2ου Βαθμού Στασιμότητα** όπου απαιτούμε μόνο τις δύο πρώτες ροπές της κατανομές – τον μέσο και την συνδιακύμανση - να είναι αμετάβλητες κατά την μετατόπιση. Κάνουμε τις εξής παραδοχές:

Ότι η αναμενόμενη τιμή (ο μέσος) της ΤΑ $Z(x)$ είναι σταθερή για όλα τα σημεία x:

$E[Z(x)]=m(x)=m$ για κάθε x.

Ότι η συνάρτηση συνδιακύμανσης C(h) μεταξύ οποιονδήποτε δύο σημείων x και (x+h), όπου το (x+h) χωρίζεται από το x από μια διανυσματική απόσταση h, είναι ανεξάρτητη από την θέση των δύο αυτών σημείων. Δηλαδή:

$E[Z(x)-Z(x+h)]-m2=C(h)$

Ότι δηλαδή η συνδιακύμανση μεταξύ δύο σημείων εξαρτάται μόνο από την απόσταση και διεύθυνση μεταξύ τους, όχι από την συγκεκριμένη θέση των ίδιων των σημείων.

Η **Εσωτερική Υπόθεση** χρησιμοποιείται στην πράξη γιατί συχνά οι παραδοχές ακόμη και της αδύναμης στασιμότητας δεν ικανοποιούνται. Για παράδειγμα όταν υπάρχει κάποια σημαντική τάση στο μέσο δεν μπορούμε να θεωρήσουμε ότι η μέση τιμή είναι σταθερή.

Έτσι σε θεωρητική και πρακτική βάση είναι βολικό να μπορούμε να αποδυναμώσουμε ακόμα περισσότερο την υπόθεση στασιμότητας. Στην εσωτερική στασιμότητα υποθέτουμε ότι οι αυξήσεις της συνάρτησης είναι αδύναμα στάσιμες. Αυτό σημαίνει ότι ο μέσος της διακύμανσης των αυξήσεων, δηλαδή:

$$Z(x+h)-Z(x)$$

είναι ανεξάρτητος από την συγκεκριμένη θέση του *x*. Η εσωτερική υπόθεση μπορεί να δοθεί περιληπτικά ως εξής:

$$E[Z(x+h)-Z(x)]=0$$

$$Var[Z(x+h)-Z(x)]=2\gamma(h)$$

όπου γ(h) λέγεται *ημιβαριόγραμμα* (ή όπως συνήθως λέγεται *βαριόγραμμα* για συντομία.[f]

Τυχαία – Συστηματική - Στρωματοποιημένη Δειγματοληψία

- **Τυχαία Δειγματοληψία**

Στη χρήση ενός μοντέλου πιθανοτήτων επιτρέπεται να συμπεριλάβουμε την αβεβαιότητα εξετάζοντας τα δεδομένα που έχουμε σαν αποτέλεσμα μιας τυχαίας διαδικασίας.

Δύο βασικές στατιστικές συνθήκες:

- Τα μέλη ενός πληθυσμού έχουν ίσες πιθανότητες να εκλεγούν για το δείγμα
- Η επιλογή ενός μέλους δεν επηρεάζει την επιλογή ενός άλλου

Η επιλογή λοιπόν γίνεται ανεξάρτητα και με ίσες πιθανότητες.

Αποτελείται από τις εξής κατηγορίες δειγματοληψίας:

Δειγματοληψία από κατάλογο

Όταν η χωρική διάσταση ενός πληθυσμού είναι υπαρκτή αλλά όχι καθοριστική για την επίλυσή του προβλήματος χρησιμοποιούμε αυτού του είδους τη δειγματοληψία

Σημειακή δειγματοληψία

Τα μέλη του πληθυσμού ορίζονται στο χάρτη ως σημεία. Κάθε μέλος του πληθυσμού αντιπροσωπεύεται από τη συντεταγμένη του, τετμημένη X και τεταγμένη Y.

Γραμμική δειγματοληψία

Στη γραμμική δειγματοληψία θεωρείται ότι η περιοχή μελέτης είναι μια συνεχής επιφάνεια και η αναζητούμενη πληροφορία επιλέγεται καταγράφοντας την κατά μήκος μιας γραμμής

Επιφανειακή δειγματοληψία

Το δειγματοληπτικό υπόβαθρο (χάρτης) αποτελείται από φατνία ενός καννάβου από τα οποία επιλέγεται τυχαία ένας μικρότερος αριθμός για να αποτελέσει το δείγμα. Δημιουργείται ένα σύστημα συντεταγμένων που ορίζουν το δειγματοληπτικό υπόβαθρο.

- **Συστηματική Δειγματοληψία**

Με τη συστηματική δειγματοληψία η επιλογή των αριθμών που αποτελούν τα μέλη ενός δείγματος δεν γίνεται τυχαία αλλά υπάρχει η φροντίδα ώστε το διάστημα μεταξύ των παρατηρήσεων να είναι σταθερό και καθορισμένο αποφεύγοντας έτσι την ελλιπή κάλυψη του δειγματοληπτικού υπόβαθρου, μειονέκτημα που διακρίνει την τυχαία δειγματοληψία. Αποτελείται από τις εξής κατηγορίες δειγματοληψίας:

Δειγματοληψία από κατάλογο

Για την επιλογή των παρατηρήσεων χρησιμοποιείται ένα σταθερό βήμα β που ισούται με το αντίστροφο του δείγματος n προς τον πληθυσμό N. Αν έχω πληθυσμιακό υπόβαθρο 5000 αριθμών και δείγμα 50 αριθμών τότε το β=100. Ο πρώτος αριθμός επιλέγεται τυχαία μεταξύ του 1 και του β ώστε να έχουν όλοι πιθανότητα επιλογής. Εν συνεχεία επιλέγεται κάθε φορά ο β επόμενος αριθμός.

Σημειακή δειγματοληψία

Το υπόβαθρο είναι ένας χάρτης και τα μέλη του πληθυσμού είναι σημεία. Το πρώτο σημείο επιλέγεται τυχαία ενώ τα υπόλοιπα πρέπει να είναι σε ίση απόσταση μεταξύ τους κατά τους άξονες X και Y.

Γραμμική δειγματοληψία

Επιλέγεται η πρώτη γραμμή με τυχαίο τρόπο και στη συνέχεια το δειγματοληπτικό υπόβαθρο καλύπτεται από παράλληλες προς την αρχική γραμμές σε συγκεκριμένη απόσταση μετρούμενη στον άξονα της αρχής της πρώτης γραμμής.

Κανναβική δειγματοληψία

Εδώ η δειγματοληπτική μονάδα είναι μια τετραγωνική επιφάνεια. Η πρώτη δειγματική επιφάνεια επιλέγεται τυχαία και τα επόμενο με τέτοιο τρόπο ώστε η κάτω αριστερή γωνία κάθε φατνίου να βρίσκεται σε καθορισμένη απόσταση, κατά τους άξονες X και Y, από το προηγούμενο.

Συνοψίζοντας έχουμε τα εξής:

Δύο βασικά μειονεκτήματα της συστηματικής δειγματοληψίας είναι:

- Η επιλογή του πρώτου σημείου καθορίζει και τα υπόλοιπα συνεπώς δεν έχουν όλα τα σημεία ίση πιθανότητα εκλογής.

- Αν υπάρχουν περιοδικότητες στο δειγματοληπτικό υπόβαθρο η συλλογή σημείων σε κανονικά διαστήματα θα έχει ως αποτέλεσμα όλα τα σημεία του δείγματος να δίνουν κάθε φορά το ίδιο χαρακτηριστικό της περιοδικότητας.

Πλεονεκτήματα της συστηματικής δειγματοληψίας είναι:

- Είναι πολύ γρηγορότερο και απλούστερο να παίρνουμε συστηματικά δείγματα αφού αποφεύγουμε τη συνεχή χρήση του πίνακα τυχαίων αριθμών.
- Δίνει περισσότερο ομοιόμορφη κάλυψη του δειγματοληπτικού υποβάθρου.

Ουσιαστικά θεωρείται πως η συστηματική δειγματοληψία είναι περισσότερο αντικειμενική και τα οφέλη που προσφέρει σε βάθος χρόνου υπερκαλύπτουν τα σφάλματα της μικρής τελικά μεροληπτικότητας που παρουσιάζουν.

- **Συστηματικές τυχαίες δειγματοληψίες**

Αποτελεί συνδυασμό τυχαίας και συστηματικής δειγματοληψίας με πολλούς τρόπους:

Πολλαπλή επιλογή Αφετηρίας

Αντιμετωπίζει σφάλματα περιοδικότητας επιλέγοντας πάνω από μία τυχαία αφετηρία. Πχ για β=10 αν επιλεγούν οι αριθμοί 5 28 67 κλπ το δείγμα θα αποτελείται από τους αριθμούς 5 15 25 και εν συνεχεία 28 38 48 κλπ..

Τυχαίο διάστημα

Όταν επιλέγεται ένα δείγμα αποκλείεται η επιλογή άλλου δείγματος που βρίσκεται σε συγκεκριμένη απόσταση. Πχ σε ένα κύκλο η επιλογή ενός σημείου εντός του κύκλου αποκλείει την επιλογή άλλων σημείων που βρίσκονται εντός κύκλου επίσης με προκαθορισμένη ακτίνα α. Όσο προχωρά η επιλογή και οι αποστάσεις πλησιάζουν την ακτίνα α η δειγματοληψία πλησιάζει τη συστηματική.

Τυχαία Μεταβαλλόμενο βήμα

Σε αυτή τη δειγματοληψία το β δεν είναι σταθερό αλλά κάθε φορά μεταβάλλεται κατά ένα μικρό τυχαίο αριθμό λ.(τιμές β + λ, β – λ).

Το λ αυξομειώνεται τυχαία πχ αν επιλεγεί μόνος αριθμός από τον πίνακα τυχαίων αριθμών τότε το λ θα προστίθεται και το αντίστροφο.

- **Στρωματοποιημένη Δειγματοληψία**

Χρησιμοποιείται σε περιπτώσεις όπου ο πληθυσμός διαιρείται σε μικρότερες ομάδες που δειγματοληπτούνται ξεχωριστά, πολύ σημαντική διαδικασία όταν ο πληθυσμός αποτελείται από υποομάδες που χαρακτηρίζονται από διαφορετικό μέγεθος ή χαρακτηριστικά. Τα πλεονεκτήματα στρωματοποιημένης δειγματοληψίας:

- Εξασφαλίζει ότι κάθε τμήμα του υπό εξέταση πληθυσμού συμπεριλαμβάνεται στο δείγμα, ώστε να είναι αντιπροσωπευτικό.
- Αυξάνει σημαντικά την ποιότητα των εκτιμήσεων για τον πληθυσμό, που εν τέλει είναι και ο τελικός σκοπός κάθε δειγματοληψίας.

Αποτελείται από τις εξής κατηγορίες δειγματοληψίας:

Στρωματοποίηση σε κατάλογο

Είναι ιδιαίτερα εύκολη, για παράδειγμα αν πρέπει να ληφθεί δείγμα για τον τρόπο που αντιλαμβάνονται την απόσταση οι άνδρες και οι γυναίκες μιας κοινότητας απλά λαμβάνονται ίσα δείγματα των δύο

Φυσική Στρωματοποίηση

Η χωρική αυτή δειγματοληψία αναφέρεται στην ανεξάρτητη λήψη δειγμάτων σε κάθε στρώμα της περιοχής μελέτης.

Κανναβική Στρωματοποίηση

Η περιοχή μελέτης επικαλύπτεται από ένα κάνναβο με φατνίο ίσου μεγέθους και για κάθε υποπεριοχή υπολογίζεται το μέγεθος του δείγματος το οποίο καθορίζει τον αριθμό των σημείων σε κάθε φατνίο.

- **Στρωματοποιημένη Συστηματική μη Γραμμική Δειγματοληψία**

1. Σύνθετη Μέθοδος. Χαρακτηριστικά της:
2. Πλεονεκτήματα Τυχαίας
3. Πλεονεκτήματα Στρωματοποιημένης
4. Διαδικασία - Βήματα
5. Διαίρεση Υποβάθρου σε Καννάβους Ίσου Μεγέθους
6. Τυχαία Επιλογή Σημείου στο Πρώτο Φατνίο
7. Σταθερή Τετμημένη Πρώτης Σειράς - Τυχαία Τεταγμένη

8. Σταθερή Τεταγμένη Πρώτης Στήλης - Τυχαία Τετμημένη

9. Ανάλογη Επιλογή Δεύτερης Σειράς και Στήλης

- **Ιεραρχική Δειγματοληψία (Ανάγκη για Αυξανόμενη Λεπτομέρεια)**

Ιεραρχική Διαίρεση του Δειγματοληπτικού Υποβάθρου

Πλεονεκτήματα:

  - Μείωση Εξόδων Επιτόπιας Έρευνας
  - Διερεύνηση Διαφοροποιήσεων σε Διάφορες Χωρικές Κλίμακες

Μειονέκτημα:

  - Στατιστική Αναξιοπιστία[g]

Εκτίμηση Παρεμβολής

Η εκτίμηση αποθεμάτων από δειγματοληπτικά δεδομένα είναι η κύρια δραστηριότητα όσον αφορά την απόφαση που θα παρθεί στην περαιτέρω διαδικασία.

Δειγματοληψία

Η πρώτη εκτίμηση (αναγνώριση) είναι το αρχικό στάδιο ενός πιθανού έργου.

Οι πληροφορίες είναι ουσιαστικά ποιοτικές και μη σαφείς και διακριτές

Υπάρχουν λίγα ή μερικές φορές καθόλου δείγματα

Τα δείγματα που υπάρχουν τείνουν να βρίσκονται σε ευνοϊκές θέσεις, και γι' αυτό είναι ομαδοποιημένα.

Πρώτη συστηματική δειγματοληψία

Μετά τον καθορισμό της περιοχής μελέτης γίνεται μια συστηματική δειγματοληψία της ζώνης ενδιαφέροντος. Περιλαμβάνει:

Τον καλύτερο δυνατό ορισμό των ορίων του υπό μελέτη φαινομένου και, άρα, της δειγματοληψίας (γεωμετρία)

Εκτίμηση της συνολικής μέσης περιεκτικότητας

Εκτίμηση του συνολικού in-situ περιεκτικότητας[1].

Σπουδαίο ρόλο εδώ παίζουν h ακρίβεια των παρατηρήσεων (αριθμός και είδος δειγμάτων και η «κανονικότητα» της παρουσίας του φαινομένου).

Λεπτομερής δειγματοληψία

Λαμβάνονται πιο πυκνά δείγματα και έτσι έχουμε πιο πολλά δείγματα και καλύτερο ορισμό του ιστογράμματος της περιεκτικότητας. μπορούμε να ορίσουμε καλύτερα την χωρική κατανομή της

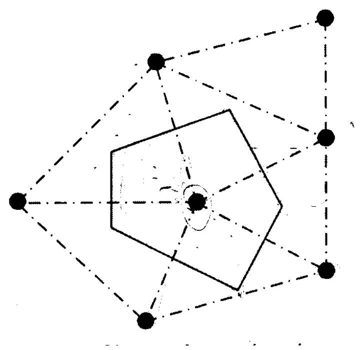
[1] Αν και από την αναδίφηση στη βιβλιογραφία προκυπτει ότι ο μόνος καθολικός κανόνας είναι η παντοδυναμία της τοπικότητας

περιεκτικότητας. Συνήθως έχουμε βελτιωμένα γεωγραφικά μοντέλα και κατά συνέπεια καλύτερη περιγραφική ανάλυση του φαινομένου.

Ανασκόπηση Τεχνικών Εκτίμησης Αποθεμάτων

Μέθοδοι πολυγώνων

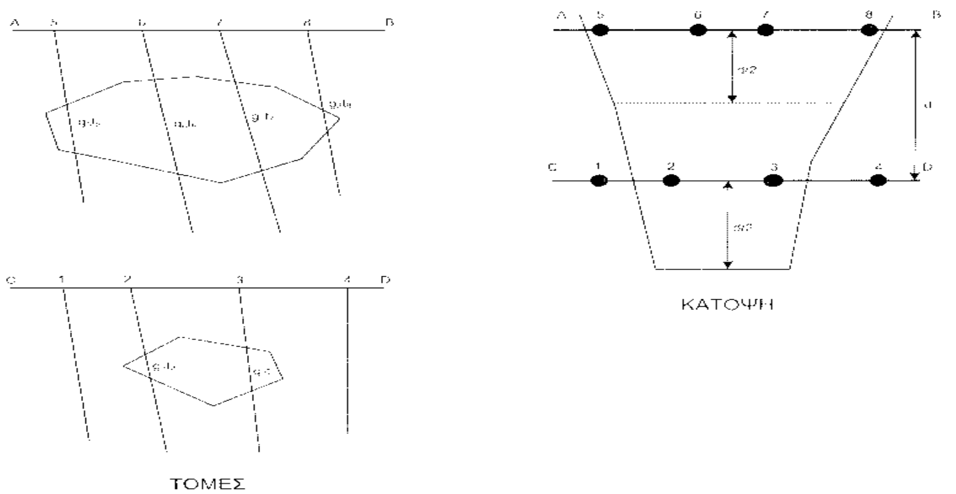
Στις πολυγωνικές μέθοδες που είναι και οι παλαιότερες κάθε δείγμα τοποθετείται στο κέντρο ενός πολυγώνου που ορίζεται από τις μεσοκαθέτους τμημάτων που ορίζονται από ζεύγη δειγμάτων (σχήμα 1) [2]



Σχήμα 1. Πολυγωνική εκτίμηση

Μέθοδοι τομών[3]

Ο εκτιμητής αυτός χρησιμοποιείται μόνο για συνολικούς πόρους. Η μέθοδος φαίνεται στο σχήμα 2.



Σχήμα 2. Μέθοδος τομών

---

[2] Εισαγωγή στην Γεωστατιστική, Καπαγερίδης, Ιωάννης, Εκδόσεις Ιων, Αθήνα, 2006 (σελ 25)

[3] Το ίδιο (σελ 26)

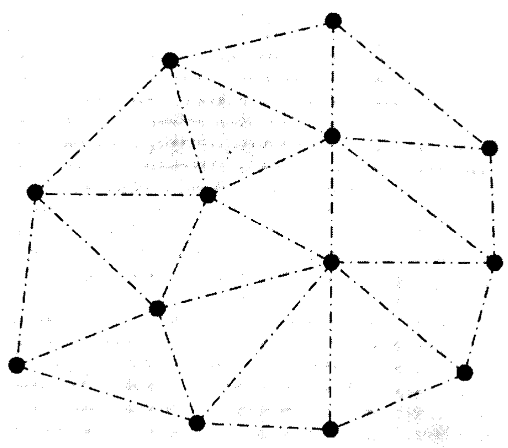
Οι μέσες περιεκτικότητες g_i των δειγμάτων κατά τμήματα ζυγίζεται με το πάχος t_i του τμήματος και δίνονται σε μια περιοχή που ορίζεται στην τομή. Η περιεκτικότητα αυτής τομής υπολογίζεται από έναν ζυγισμένο μέσο:

$$z^* = \frac{\sum_{i=1}^{N} g_i \cdot t_i}{\sum_{i=1}^{N} t_i}$$

Μέθοδοι Τριγώνων

Πρόγονοι της ζύγισης αντιστρόφου αποστάσεως σε δύο παραλλαγές:

- Η μέση περιεκτικότητα ενός τριγώνου που ορίζεται από τα τρία γωνιακά δείγματα εκτιμάται από τη μέση τιμή τους (σχήμα 4)



Σχήμα 4. Τριγωνική Μέθοδος

- Εκτίμηση μικρών μπλοκ, κάθε μπλοκ εντός του τριγώνου εκτιμάται από ένα γραμμικό συνδυασμό των γωνιακών γεωτρήσεων. Τα βάρη που χρησιμοποιούνται είναι αντιστρόφως ανάλογα της απόστασης του δείγματος από το μπλοκ.

Μέθοδοι Αντιστρόφου Αποστάσεως (ΑΑ)

Πιο πρόσφατες από τις προηγούμενες τεχνικές και με την εξάπλωση των υπολογιστών έγιναν πιο δημοφιλείς. Η ένταση του φαινομένου διερευνάται με διαίρεση του χώρου σε μπλοκ ίδιου μεγέθους. Η μέση περιεκτικότητα ενός μπλοκ εκτιμάται από έναν ζυγισμένο γραμμικό συνδυασμό γειτονικών δειγμάτων. Οι παράγοντες βαρύτητας δίνονται από:

$$\lambda_i = \frac{\frac{1}{d_i^2}}{\sum_{i=1}^{N} \frac{1}{d_i^2}}$$

όπου d_i είναι η απόσταση του δείγματος είναι η απόσταση του δείγματος *I* από το κέντρο του μπλοκ που υπολογίζεται.

Η μέθοδος ΑΑ είναι μια γενίκευση της τριγωνικής μεθόδου και εφαρμόζεται εύκολα στον υπολογιστή.

Οι κλασσικές μέθοδοι που αναφέρθηκαν παραπάνω μπορούν να «βελτιωθούν» μόλις αρχίσει η παρέμβαση, αλλά δεν υπάρχει τρόπος να πάρουμε τις «βέλτιστες» εκτιμήσεις από τις τεχνικές αυτές *πριν* από τη χωρική παρέμβαση.[h]

Kriging [4]

Το kriging είναι ένας τρόπος για να βρίσκουμε τα βάρη λ_i έτσι ώστε να αποδίδουν την μεταβλητότητα των περιεκτικοτήτων στο χώρο:

- Με την διόρθωση του Krige ο εκτιμητής αυτός ζυγίζει ένα δείγμα σύμφωνα με την θέση του σε σχέση με το τμήμα
- Το kriging δίνει τα βάρη κατά τέτοιο τρόπο που να είναι μαθηματικά βέλτιστος
- Το kriging επιτρέπει να αποδώσουμε το μέσο σφάλμα που προέκυψε στην εκτίμηση ενός τμήματος ορισμένης γεωμετρίας σε ένα φαινόμενο χρησιμοποιώντας μια συγκεκριμένη διάταξη δειγμάτων
- Το kriging είναι στατιστική μέθοδος, βασίζεται στις ιδέες της θεωρίας των πιθανοτήτων και χρησιμοποιεί την ιδέα των χωρομεταβλητών.[i]

---

[4] Στις αρχές της δεκαετίας του ’60, ο George Matheron ανέπτυξε μια γενική λύση στο πρόβλημα της τοπικής εκτίμησης η οποία στηρίχθηκε σε μια εμπειρική λύση που αναπτύχθηκε από τον Νοτιο-Αφρικανό D.G. Krige. Για να τιμήσει την συνεισφορά του Krige σε αυτό το πεδίο ο Matheron ονόμασε την νέα τεχνική που ανέπτυξε *kriging*.

Μέθοδοι Τοπικών Εκτιμήσεων

Οι μέθοδοι αυτές αναφέρονται στην εκτίμηση της τιμής ενός χαρακτηριστικού σε μια συγκεκριμένη θέση, με βάση στοιχεία τα οποία προέρχονται από σημεία που βρίσκονται στην άμεση γειτονική περιοχή του. Η διαδικασία απαιτεί τα εξής βήματα:

- Τον ορισμό της τοπικής περιοχής εκτίμησης ή την αποδεκτή «γειτονία» γύρω από το υπό εκτίμηση σημείο.
- Την εύρεση του αριθμού των σημείων που οι τιμές τους θα πρέπει να ληφθούν υπόψη για την εκτίμηση.
- Την επιλογή των σημείων αυτών από το σύνολο των σημείων της περιοχής μελέτης.
- Την επιλογή της μαθηματικής συνάρτησης που αντιπροσωπεύει τη διαφοροποίηση της τιμής του χαρακτηριστικού, δηλαδή τη διαδικασία εκτίμησης.

Θα αναφερθούμε σε δύο μεθόδους τοπικών εκτιμήσεων που είναι οι εξής:

Χωρικός Κινητός Μέσος

Η μέθοδος εκτίμησης αυτή βασίζεται στην εύρεση του μέσου όρου των τιμών που παρατηρούνται γύρω από το σημείο που έχει επιλεγεί.

Στατιστικά θεωρείται ότι:

- Υπάρχει μια σειρά από παρατηρήσεις Zi $i=1,\ldots,n$ για ένα χωρικά συνεχές χαρακτηριστικό.
- Οι μετρήσεις Zi είναι παρατηρήσεις μιας χωρικής στοχαστικής διαδικασίας $\{Zi, s\in A\}$,οι οποίες έχουν καταγραφεί σε αντίστοιχες χωρικές θέσεις si στην περιοχή μελέτης A.
- Υποτίθεται ότι η εκτίμηση της απλής (χωρίς βάρη) μέσης τιμής, είναι μια ισοτροπική διαδικασία (δεν παρατηρούνται «τάσεις» προς καμία κατεύθυνση).

Η υπόθεση της ισοτροπίας δεν είναι πάντοτε σωστή και δεν επιτρέπει τη χωρική διαφοροποίηση στην κατανομή των σημείων που έχουν επιλεγεί.

Για την αντιμετώπιση του προβλήματος αυτού χρησιμοποιείται ο μέσος όρος των τιμών των παρατηρήσεων με βάρη. Πιο συγκεκριμένα:

$$\hat{\mu}(s) = \sum_{i=1}^{n} w_i(s) Z_i \quad \underline{\text{ΟΠΟΥ}} \quad \sum_{i=1}^{n} w_i(s) = 1$$

$$w_i(s) = h_i^{-a} \qquad \underline{\text{ΚΑΙ}} \qquad w_i(s) = e^{-ah_i}$$

Όπου:

h_i = η απόσταση μεταξύ του υπό εκτίμηση σημείου **s** και του σημείου $\mathbf{s_i}$.

α = είναι μια παράμετρος που παίρνει τιμές, ώστε η εκτιμηθείσα επιφάνεια να είναι όσο το δυνατό πιο ομαλή.

Συνήθως, η $\mathbf{w_i(s)}$ ορίζεται να παίρνει τιμή μηδέν πέρα από μια συγκεκριμένη απόσταση.

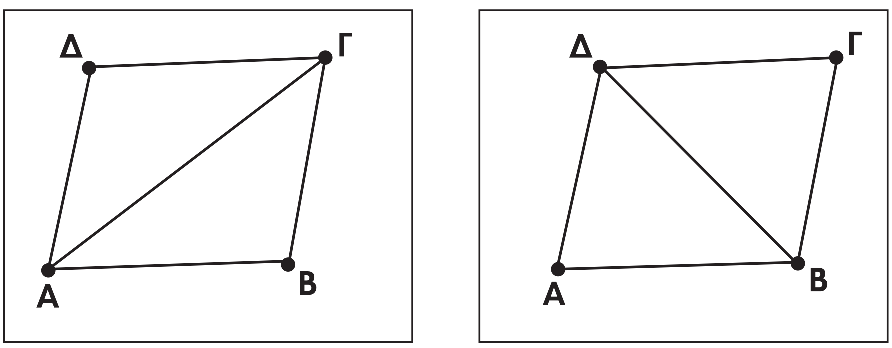
Ψηφιδοποίηση σε TIN (Triangulated Irregular Network)

Η διαδικασία της δημιουργίας του TIN αφορά:

- Στη δημιουργία του πλέγματος των τριγώνων (όλα τα σημεία ενώνονται μεταξύ τους, μετατρεπόμενα σε ένα σύνολο πλευρών τριγώνων).
- Στον καθορισμό της συνάρτησης της χωρικής διαφοροποίησης των τιμών (η τιμή του χαρακτηριστικού μεταξύ των δύο κορυφών της πλευράς μεταβάλλεται με έναν καθορισμένο και σταθερό τρόπο).

Η διαδικασία «τριγωνοποίησης» μπορεί να επιτευχθεί με διαφορετικούς τρόπους ανάλογα με το κριτήριο σύνδεσης των σημείων.

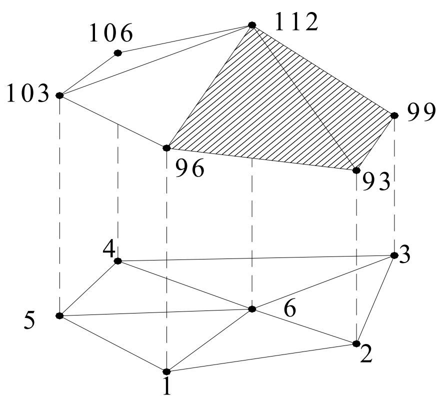
Από τις μεθόδους αυτές η πλέον γνωστή είναι η **μέθοδος Delannay**, γνωστή και ως κριτήριο μέγιστη-ελάχιστη γωνία.



α β

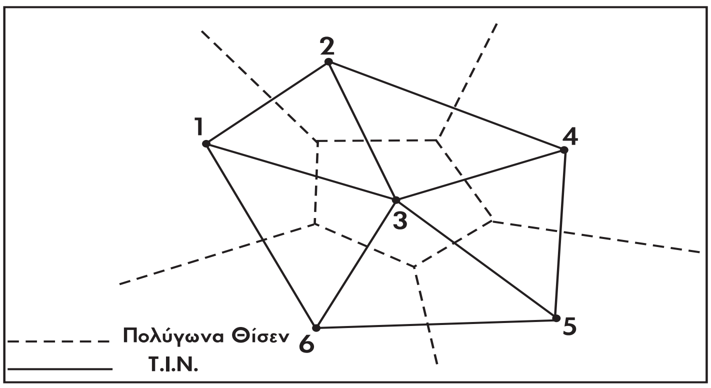
Με την ολοκλήρωση της «τριγωνοποίησης», ο ορισμός της χωρικής συνάρτησης, μπορεί να αρχίσει και μπορεί να πάρει πολλές μορφές.

Η πιο απλή: γραμμική παρεμβολή $\mathbf{Z_i = \alpha X_i + \beta Y_i + \gamma}$



Πολύγωνα Θίσσεν

Αν υπάρχουν n γνωστά σημεία σε μια περιοχή μελέτης Α, σε κάθε σημείο si κατανέμεται ένα τμήμα της Α, έτσι ώστε κάθε σημείο του τμήματος αυτού να είναι πλησιέστερο στο si περισσότερο από κάθε άλλο σημείο sj. Οι γραμμές μεταξύ δύο σημείων αποτελούν το γεωμετρικό τόπο των σημείων που ισαπέχουν από τα σημεία αυτά. Κάθε πλευρά ενός πολυγώνου Θίσσεν είναι ένα γραμμικό τμήμα που τέμνει κάθετα τη γραμμή που ενώνει το σημείο του πολυγώνου με το γειτονικό του, που ενώνει δηλαδή την πλευρά ενός τριγώνου του TIN



Γενικευμένες Μέθοδοι

Οι **γενικευμένες μέθοδοι** εκτίμησης χρησιμοποιούν όλα τα υπάρχοντα στοιχεία, επιτυγχάνοντας εκτιμήσεις για το σύνολο της περιοχής που ενδιαφέρει.

Οι **γενικευμένοι εκτιμητές** χρησιμοποιούνται συνήθως εμμέσως για χωρικές παρεμβολές (αποτελούν εργαλεία εξέτασης και απομάκρυνσης των γενικευμένων χωρικών διαφοροποιήσεων).

Οι **διαδικασίες εκτίμησης** των γενικευμένων μεθόδων υπολογίζονται σχετικά απλά και βασίζονται κυρίως σε συνηθισμένες στατιστικές μεθόδους ανάλυσης διασποράς.

Θα αναφερθούμε σε δύο γενικευμένες μεθόδους εκτιμήσεων που είναι οι εξής:

Επιφάνεια Τάσης

Όταν η διαφοροποίηση σε ένα χαρακτηριστικό είναι συνεχής στο χώρο, υπάρχει ανάγκη για τη δημιουργία ενός μοντέλου, ώστε να μπορεί:

- Να εξηγηθεί αυτή η διαφοροποίηση.
- Να εκτιμηθεί η τιμή του χαρακτηριστικού σε θέσεις εκτός αυτών για τις οποίες υπάρχουν στοιχεία.

Η μέθοδος της **ανάλυσης επιφάνειας τάσης** επιτυγχάνει το διαχωρισμό των παρατηρήσεων μιας χωρικά κατανεμημένης μεταβλητής σε:

- Ένα τμήμα που σχετίζεται με τις γενικευμένες τάσεις που υπάρχουν.
- Ένα τμήμα που είναι το αποτέλεσμα των τοπικών επιδράσεων.

Επομένως, κάθε τιμή **Z(s)** ενός φαινομένου μπορεί να διαχωριστεί σε δύο μέρη, που το καθένα είναι αποτέλεσμα μιας ξεχωριστής, σε διαφορετική κλίμακα, χωρικής διαδικασίας.

- Το πρώτο τμήμα είναι αποτέλεσμα μιας διαδικασίας **μεγάλης κλίμακας**, λειτουργεί δηλαδή σε μια μεγάλη περιοχή και δημιουργεί την επιφάνεια τάσης.
- Το δεύτερο τμήμα συνδυάζει τις τυχαίες μεταβολές και τα σφάλματα μέτρησης και είναι αποτέλεσμα μιας χωρικής διαδικασίας που επιδρά σε μια σημαντικά **μικρότερη περιοχή** από την περιοχή μελέτης.

Για τη χωρική διαφοροποίηση σε ένα χαρακτηριστικό ισχύει η σχέση:

παρατηρούμενη τιμή σε ένα σημείο	=	τιμή επιφάνειας τάσης στο σημείο αυτό	+	υπόλοιπο στο σημείο αυτό

Η βασική εξίσωση κάθε χωρικού προτύπου να δίνεται από τη σχέση:

$Z_i = f(X_i, Y_i) + \varepsilon_i$

Όπου:

$\mathbf{Z_i}$ = η παρατηρούμενη τιμή της μεταβλητής Ζ στο σημείο s_i (X_i, Y_i)

$\mathbf{f(X_i, Y_i)}$ = η τιμή της επιφάνειας τάσης στο σημείο s_i (X_i, Y_i)

$\boldsymbol{\varepsilon}_i$ = το υπόλοιπο

Η μορφή της συνάρτησης **f** μπορεί να ποικίλλει από απλή γραμμική μέχρι και πολύπλοκου τετάρτου ή μεγαλύτερου βαθμού.

Μοντέλα Ταξινόμησης

Οι τάσεις θεωρούνται ότι είναι το αποτέλεσμα της ύπαρξης ενός υπόβαθρου αποτελούμενου από μια ομάδα περιοχών οι οποίες διακρίνονται μεταξύ τους από τις μέσες τιμές του υπό εξέταση χαρακτηριστικού.

Υποθέτει ότι η παρατηρούμενη δομή μιας συγκεκριμένης χωρικής διαφοροποίησης καθορίζεται από εξωγενώς καθοριζόμενες χωρικές ενότητες.

Στη χρησιμοποιούμενη μέθοδο της ανάλυσης διασποράς, θεωρείται ότι εντός κάθε χωρικής ενότητας η διασπορά των τιμών είναι μικρότερη από τη διασπορά των τιμών μεταξύ των ενοτήτων.

Επομένως, σημαντικές αλλαγές στις τιμές της παρατηρούμενης μεταβλητής παρατηρούνται στα όρια μεταξύ των χωρικών μονάδων.

Το στατιστικό μοντέλο ταξινόμησης έχει ως εξής:

$Z(s) = \mu + v_{\kappa} + \varepsilon$

όπου:

Z(s) = η τιμή του χαρακτηριστικού στη θέση **s**.

μ = η συνολική, για ολόκληρη την περιοχή μελέτης, μέση τιμή.

$\mathbf{v_{\kappa}}$ = η απόκλιση από τη **μ**, της μέσης τιμής κάθε χωρικής μονάδας **κ**, και

ε = το υπόλοιπο, σφάλμα, που είναι γνωστό και ως «θόρυβος».

Το μοντέλο θεωρεί ότι για κάθε ομάδα κ, οι τιμές του χαρακτηριστικού έχουν κανονική κατανομή που η μέση τιμή της είναι ίση με $\mathbf{\mu + v_{\kappa}}$

Ο δείκτης που έχει σημασία είναι η σχετική διασπορά

$$\sigma^2_{\varepsilon} / \sigma^2_{\sigma}$$

Όπου:

σ^2_{σ} = η συνολική διασπορά που είναι ίση με $\sigma^2_{\varepsilon} + \sigma^2_{\mu}$

σ^2_{μ} = η διασπορά μεταξύ κατηγοριών, που είναι η ίδια για όλες τις κατηγορίες.[j]

Το Βαριόγραμμα

Το βαριόγραμμα είναι το βασικό διαγνωστικό εργαλείο για το χωρικό χαρακτηρισμό μιας μεταβλητής και είναι επίσης κεντρικό στην γεωστατιστική εκτίμηση ή της μεθόδους παρεμβολής (kriging).

Ορισμός: Το βαριόγραμμα μιας εσωτερικής τυχαίας συνάρτησης ορίζεται ως εξής

$$\gamma(h) = \frac{1}{2} Var[Z(x+h) - Z(x)]$$

Με την υπόθεση εσωτερικής στασιμότητας (παραπάνω) το βαριόγραμμα μπορεί να οριστεί ως:

$$\gamma(h) = \frac{1}{2} E[\{Z(x+h-Z(x)\}^2]$$

Στην πράξη χρησιμοποιείται η παρακάτω εξίσωση για τον υπολογισμό του πειραματικού βαριογράμματος από τα διαθέσιμα δεδομένα:

$$\hat{\gamma}(h) = \frac{1}{2N} \sum_{i=1}^{N} [\{Z(x_i + h) - Z(x_i)\}^2]$$

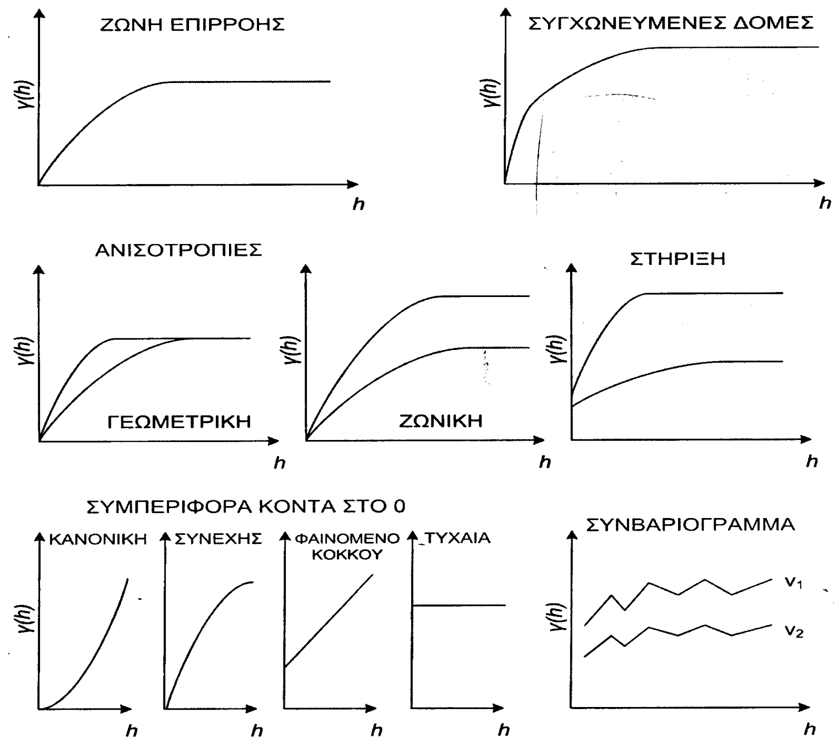
Έτσι έχουμε N ζεύγη δεδομένων.

Κύρια Χαρακτηριστικά: Το γράφημα του γ(h) παρουσιάζει τα παρακάτω χαρακτηριστικά (σχήμα):

- Το γ(h) είναι μια μη-αρνητική συνάρτηση, δηλαδή $\gamma(h) \geq 0$
- Ξεκινά στο 0 για h=0
- Γενικά αυξάνει με το h
- Το γ(h) μπορεί να αυξάνει μέχρι μια ορισμένη τιμή του h (που ονομάζεται οριακή τιμή) και μετά οριζοντιώνεται – σταθεροποιείται.
- Εναλλακτικά το γ(h) μπορεί να συνεχίσει να αυξάνει για αυξανόμενα διαστήματα h.

Εύρος και «Ζώνη Επιρροής»: Ο ρυθμός αύξησης του βαριογράμματος με την αύξηση του *h* είναι ενδεικτικός του πόσο γρήγορα μειώνεται η ‘επιρροή’ των δειγμάτων με την απόσταση. Το *γ(h)* δίνει μια ακριβή σημασία στην συμβατική ιδέα της ‘ζώνης επιρροής’ μιας τιμής.

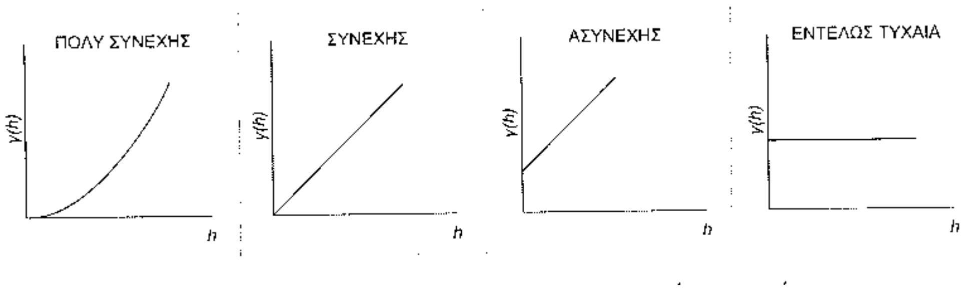
Τα παρακάτω διαγράμματα δείχνουν τη διαφορά των βαριογραμμάτων όσον αφορά την οριακή συμπεριφορά όταν *h* τείνει σε μεγάλες τιμές.



Σχήμα 6. Κύρια χαρακτηριστικά του βαριογράμματος

Η συμπεριφορά του βαριογράμματος για πολύ μικρά διαστήματα αποδίδει την κανονικότητα και συνέχεια της ίδιας της χωρομεταβλητής.

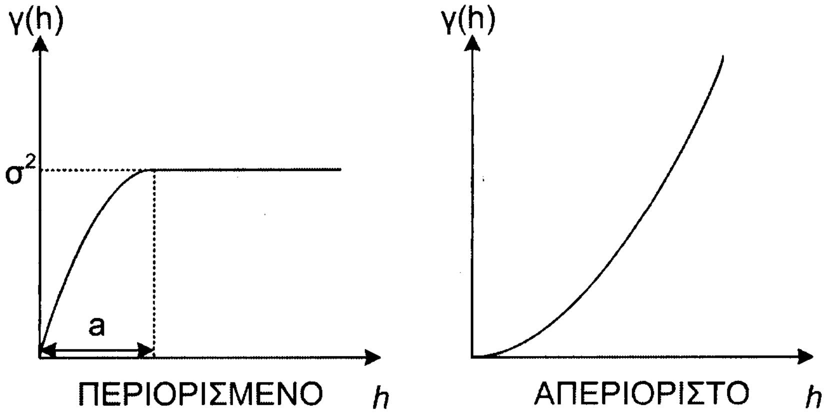
Συμπεριφορά κοντά στην αρχή (Πολύ συνεχής, Συνεχής, Ασυνεχής, Τυχαία Συμπεριφορά).



Σχήμα 7. Συμπεριφορά του βαριογράμματος κοντά στην αρχή

Πολύ συνεχής μπορεί να είναι μια παραβολική συμπεριφορά κοντά στην αρχή. Πράδειγμα στην φύση μπορεί να είναι το ‘υψόμετρο μιας πολύ ομαλής, ελαφρά κυματώδους τοπογραφικές επιφάνειας’. Στην περίπτωση αυτή περιμένουμε οι κοντινές τιμές να είναι, κατά μέσο όρο, *πολύ*, όμοιες και έτσι εμφανίζεται το παραβολικό σχήμα.

Συνεχής – Η γραμμική συμπεριφορά του *γ(h)* κοντά στην αρχή είναι ενδεικτική μέτριας συνέχειας μικρού εύρους. Η συνέχεια είναι έντονα μικρότερη από την παραβολική συμπεριφορά. Μερικά κοιτάσματα βασικών μετάλλων εμφανίζουν αυτόν το τύπο βαριογράμματος.



Σχήμα 8. Περιορισμένα και απεριόριστα ή μεταβατικά και μη-μεταβατικά βαριογράμματα

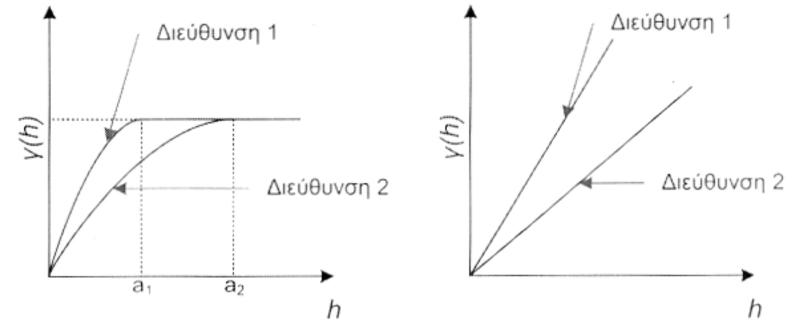
Ασυνεχής - Όταν το *γ(h)* δεν τείνει στο μηδέν όταν το *h* τείνει στο μηδέν. Η ασυνεχής συμπεριφορά κοντά στην αρχή δείχνει μια ιδιαίτερα ασυνεχή συμπεριφορά της χωρομεταβλητής σε μικρές αποστάσεις. Παράδειγμα είναι η συμπεριφορά των αναλύσεων χρυσού σε μικρές αποστάσεις λόγω της φυσικής κατανομής του χρυσού σε ψήγματα. Έτσι το **φαινόμενου συσσωματώσεως** εφαρμόστηκε για να περιγράψει αυτό το αυτόματο άλμα στην αρχή.

Τυχαία συμπεριφορά - Το οριζόντιο γράφημα του *γ(h)* δείχνει ακραία ασυνέχεια της χωρομεταβλητής. Αυτό δείχνει πλήρη τυχαιότητα – 'λευκός θόρυβος'

Ανισοτροπία (Γεωμετρική Ανισοτροπία, Ζωνική Ανισοτροπία)

Η ανισοτροπία παρουσιάζεται όταν υπάρχουν διαφορές στην συμπεριφορά του βαριογράμματος όταν υπολογίζεται σε διαφορετικές διευθύνσεις.

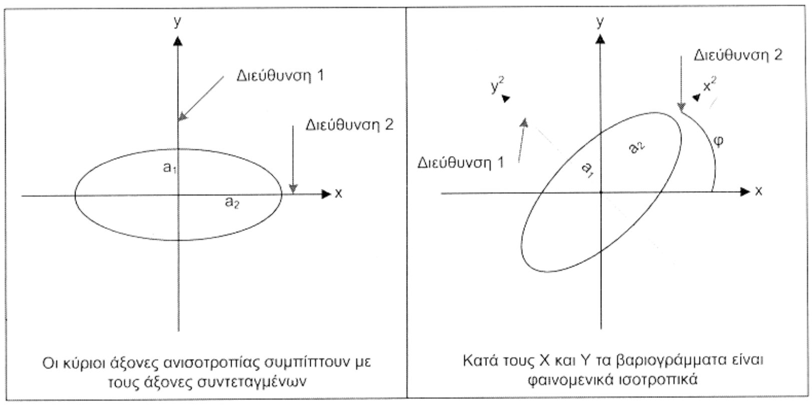
Γεωμετρική ανισοτροπία



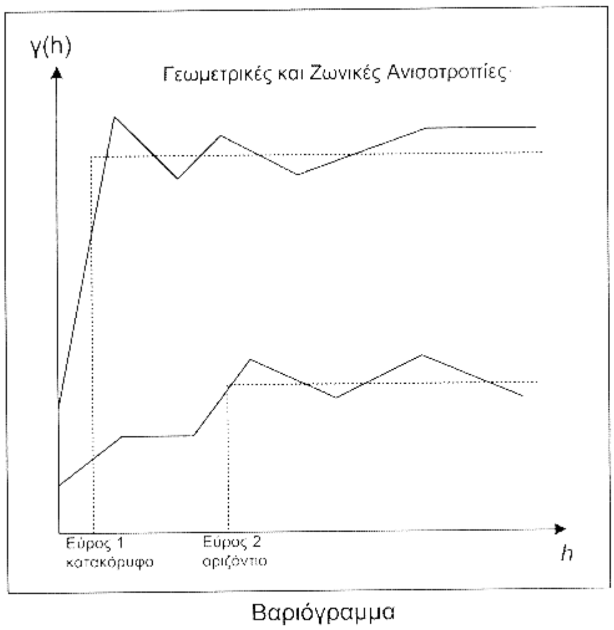
Σχήμα 9. Γεωμετρική ανισοτροπία περιορισμένων και απεριόριστων περιπτώσεων

Διαφορά μόνο στο εύρος. Διορθώνεται με ένα γραμμικό μετασχηματισμό των συντεταγμένων

Ζωνική ανισοτροπία – διάφορα είδη.



Σχήμα 10. Γεωμετρική ανισοτροπία



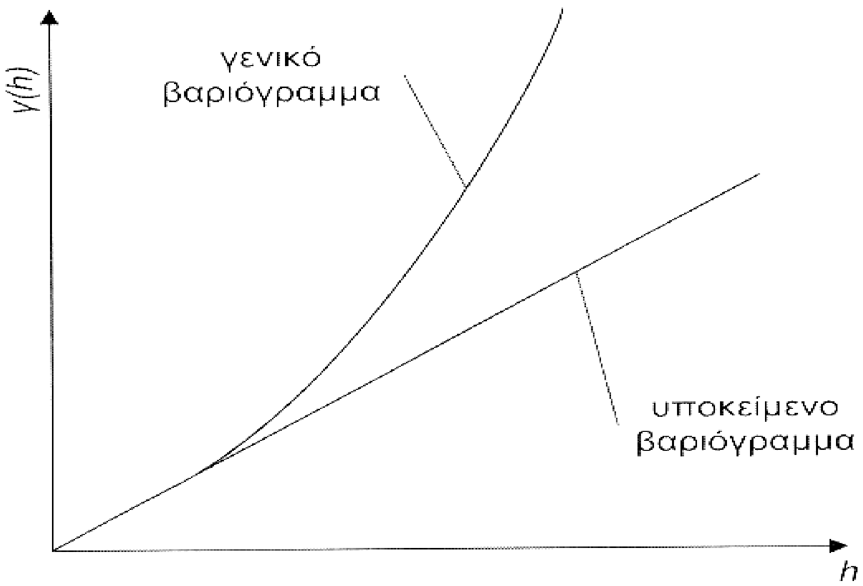
Σχήμα 11. Ανισοτροπία και ‘ζώνη επιρροής’

Παρουσία Εκτροπής

Θεωρητικά ισχύει ότι για υψηλές τιμές του h, το γ(h) πρέπει να αυξάνει πιο αργά από μια παραβολή:

$\frac{\gamma(h)}{h^2} \rightarrow 0$ καθώς h→ ∞

Στην πράξη βλέπουμε βαριογράμματα που αυξάνουν πιο γρήγορα από το h2. Αυτή η εκτροπή διορθώνεται με την πρόσθεση ενός διορθωτικού όρου:



Σχήμα 12. Γενικό και υποκείμενο βαριόγραμμα

$\gamma(h)_{\gamma\varepsilon\nu\iota\kappa o} = \gamma(h)_{\upsilon\pi o\kappa\varepsilon\iota\mu\varepsilon\nu o} + \alpha^2 h^2$

Αναλογικό Φαινόμενο

Ένα βαριόγραμμα λέγεται να έχει αναλογικό φαινόμενο όταν η τιμή γ(h) είναι αναλογική της τοπικής μέσης περιεκτικότητας. Τα βαριογράμματα διαφορετικών ζωνών έχουν το ίδιο σχήμα, αλλά διαφορετικές οριακές τιμές. Συχνά συμβαίνει η οριακή τιμή να είναι ανάλογη στο τετράγωνο του τοπικού μέσου. Το υποκείμενο βαριόγραμμα μπορεί να βρεθεί διαιρώντας τα τοπικά βαριογράμματα με το τετράγωνο του τοπικού μέσου και στην συνέχεια βρίσκουμε την μέση τιμή τους πριν προσαρμόσουμε κάποιο μοντέλο.

Συγχωνευμένες Δομές, Φαινόμενο Οπής, Περιοδικότητα

Το σχήμα δείχνει παραδείγματα των φαινομένων αυτών.

Σχήμα 13. Συγχωνευμένες δομές, περιοδικότητα και φαινόμενο οπής[k]

Μερικές εφαρμογές[l]

Για να κατανοήσουμε καλύτερα τις διάφορες έννοιες θα εξετάσουμε την πρακτική υπολογισμού του πειραματικού βαριογράμματος

$$\hat{\gamma}(h) = \frac{1}{2N}\sum_{i=1}^{N}[\{Z(x_i + h) - Z(x_i)\}^2]$$

που είναι μια εκτίμηση του γενικού βαριογράμματος

$$\gamma(h) = \frac{1}{2}E[\{Z(x + h - Z(x)\}^2]$$

Στη γενική περίπτωση το βαριόγραμμα πρέπει να υπολογιστεί (θεωρητικά και στην πράξη) σε τρεις διαστάσεις. Θα εξετάσουμε πρώτα τις δύο περιπτώσεις, κατά μήκος μιας γραμμής και στο επίπεδο. Παρακάτω θα δούμε στις μελέτες που έχουν γίνει την γενικότερη περίπτωση.

Κατά μήκος μιας γραμμής

Στην περίπτωση που ο υπολογισμός του πειραματικού βαριογράμματος δειγμάτων κατά μήκος μιας γραμμής μπορεί να θεωρηθεί ως μονοδιάστατος.

Θεωρούμε ότι τα δείγματα βρίσκονται στο κέντρο κάθε διαστήματος. Εάν τα σημεία είναι σε κανονικές αποστάσεις κατά μήκος της γραμμής, το βαριόγραμμα μπορεί να υπολογιστεί από κάθε διάστημα h χρησιμοποιώντας την εξίσωση:

$$\hat{\gamma}(h) = \frac{1}{2N}\sum_{i=1}^{N}[\{Z(x_i + h) - Z(x_i)\}^2]$$

όπου:

$Z(x_i)$ είναι τα δεδομένα

x_i είναι σε θέσεις στις οποίες υπάρχουν δεδομένα και για *xi* και για (x_i+h).

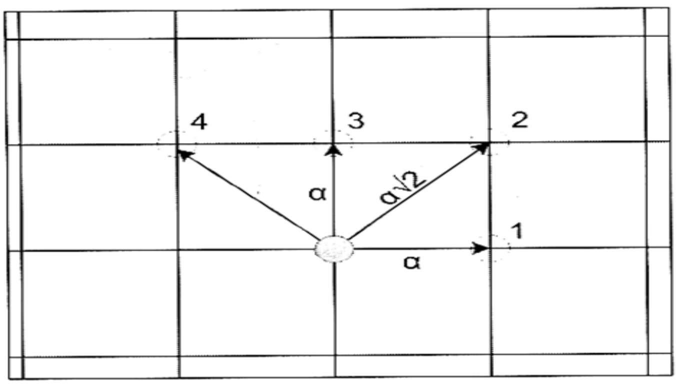
N(h) είναι το πλήθος των x_i δηλαδή το πλήθος των ζευγών σημείων που λαμβάνονται στο άθροισμα όταν υπολογίζεται το *γ(h)*. Εάν λείπουν δεδομένα, το ζεύγος απλά αγνοείται.

Στο επίπεδο – δύο διαστάσεις

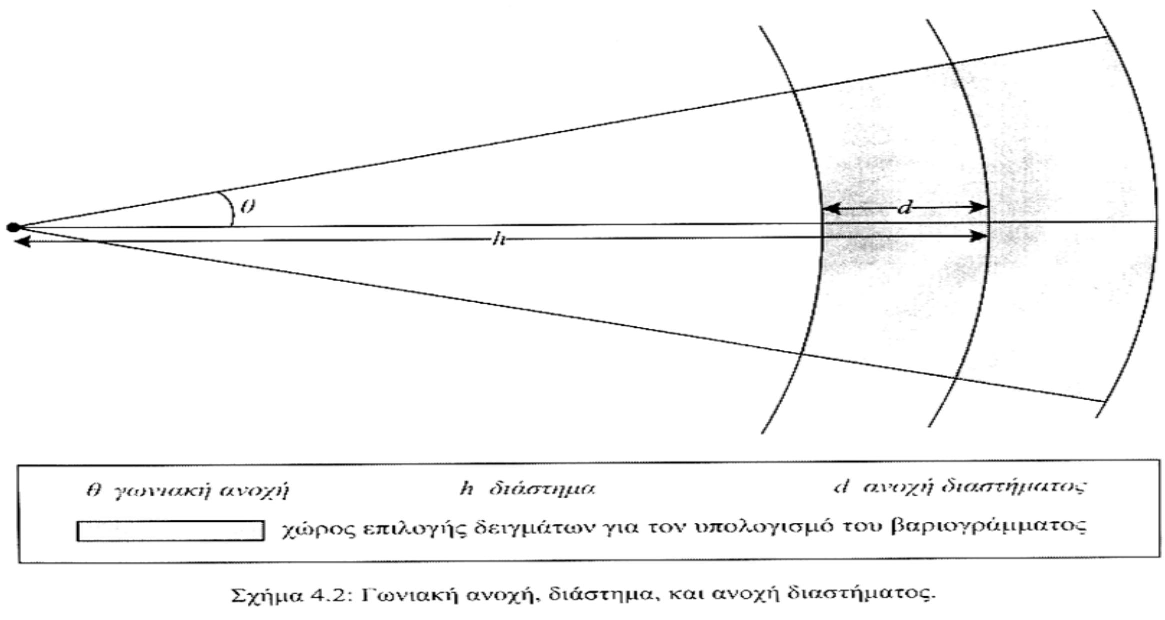
Σε πολλές εφαρμογές απαιτείται ο υπολογισμός των βαριογραμμάτων σε δύο διαστάσεις:

- Για την μοντελοποίηση τοπογραφικών δεδομένων ή γεωγραφικών επιφανειών.
- Όταν εξετάζονται δεδομένα που έχουν συλλεχθεί σε σημεία μιας επιφάνειας, π.χ. εξερευνητική γεωχημεία εδάφους.

Πρέπει να χρησιμοποιούνται τουλάχιστον τέσσερις διευθύνσεις για τον έλεγχο ανισοτροπίας. (δείτε σχήμα)



Σχήμα 4.1: Τα διαστήματα είναι διαφορετικά, ακόμα και για ένα ισοτροπικό πλέγμα.



Σχήμα 4.2: Γωνιακή ανοχή, διάστημα, και ανοχή διαστήματος.

Σχήμα. 13. Τα διαστήματα είναι διαφορετικά ακόμα και για ένα ισοτροπικό πλέγμα

Η ανίχνευση γίνεται σε τομέα ενός κύκλου (σχήμα 13)

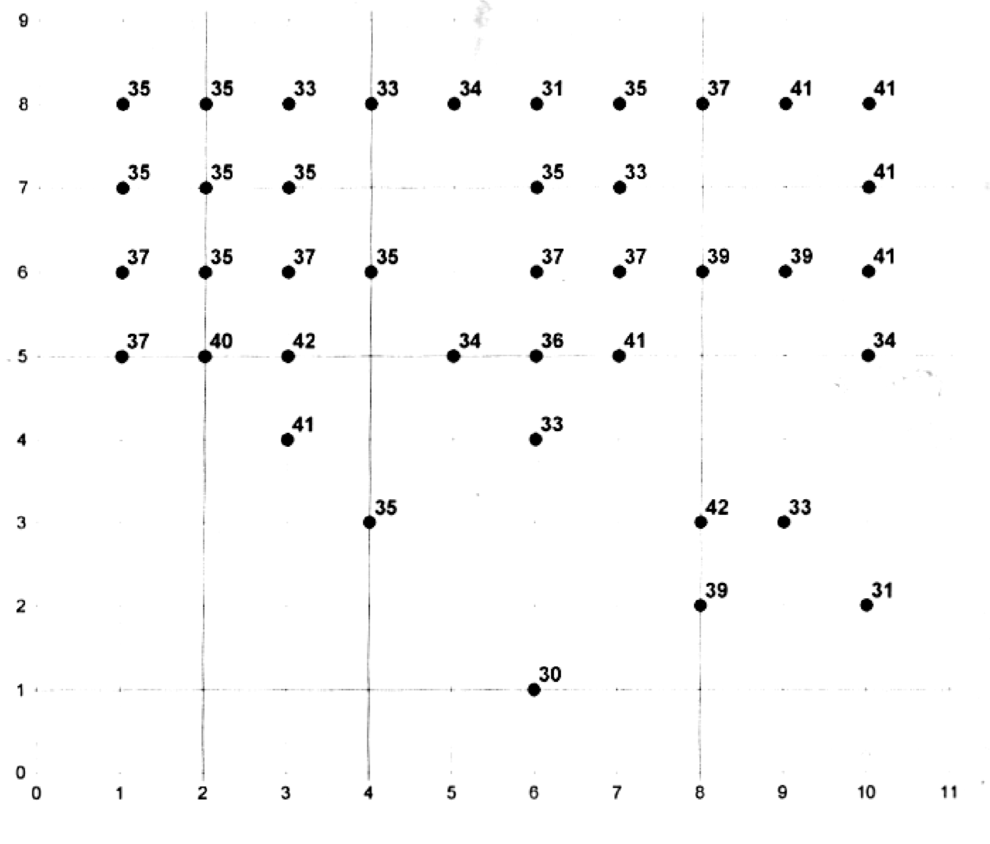
Στο χώρο – σε τρεις διαστάσεις

Σε τρεις διαστάσεις γενικεύουμε την προσέγγιση δύο διαστάσεων ώστε η ανίχνευση να γίνεται σε ένα κωνικό τμήμα μιας σφαίρας.

Παράδειγμα – εφαρμογή υπολογισμού ενός Πειραματικού Βαριογράμματος

Θα χρησιμοποιήσουμε τα δεδομένα και τεχνική του υπολογισμού στην Εισαγωγή στη Γεωστατιστική (Καπαγερίδης) με μια προσπάθεια να συμπληρώσουμε τον πίνακα που έχει γίνει μόνο μια αρχή, έτσι ώστε να κατανοήσουμε καλύτερα το πώς γίνεται ο υπολογισμός ενός πραγματικού βαριογράμματος.

Τα δεδομένα παρουσιάζονται στο παρακάτω σχήμα



Σχήμα 14. Δεδομένα παραδείγματος

Είναι σε τετραγωνικό πλέγμα διαστήματος α. Το βαριόγραμμα θα πρέπει να υπολογιστεί στις τέσσερις κύριες διευθύνσεις α=1,2,3,4 και για τιμές h=1,2, και 3.

Στις διευθύνσεις 3 και 4 το διάστημα θα είναι ίσο με $a\sqrt{2}$.

Χρησιμοποιούμε την εξίσωση

$$\hat{\gamma}_a(h) = \frac{1}{2N}\sum_{i=1}^{N}[\{Z(x_i+h)-Z(x_i)\}^2]$$

για να υπολογίσουμε το βαριόγραμμα σε κάθε διεύθυνση

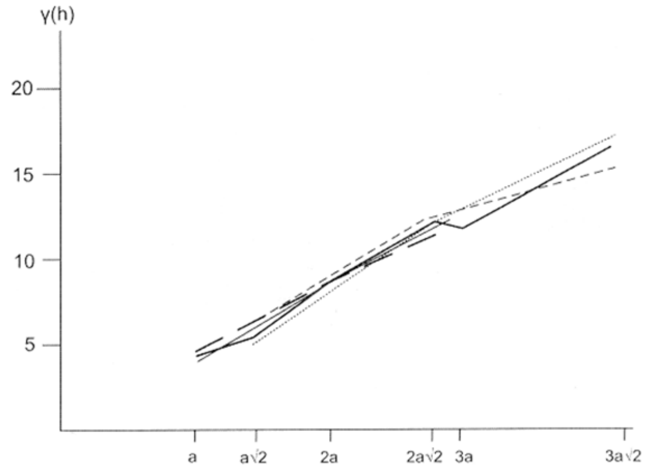
Τα ζεύγη των κόμβων με μόνο μια (ή καθόλου) τιμές θεωρούνται άγνωστα και δεν εισάγονται στον υπολογισμό:

Υπολογίζουμε τα κενά στον παρακάτω πίνακα (δείτε παράρτημα)

Διεύθυνση	h = 1 διάστημα		h = 2 διαστήματα		h = 3 διαστήματα	
	N(1)	*γ(1)*	N(2)	*γ(2)*	N(3)	*γ(3)*
1	24	4.10				
2						
3						
4						

Τα αποτελέσματα συνοψίζονται στον πίνακα:

Διεύθυνση	h=1(Διάστημα)		h=2 (Διαστήματα)		h=3 (Διαστήματα)	
	N(1)	γ(1)	N(2)	γ(2)	N(3)	γ(3)
1	24	4,10	17	7,94	13	12,73
2	17	5,03	12	13,08	7	17,36
3	19	4,82	15	9,20	12	10,83
4	19	5,03	17	11,32	9	12,28



Σχήμα 4.4: Παράδειγμα πειραματικού βαριογράμματος: ισοτροπική συμπεριφορά.

Μοντέλα Βαριογραμμάτων – Μερικά κοινά μοντέλα[m]

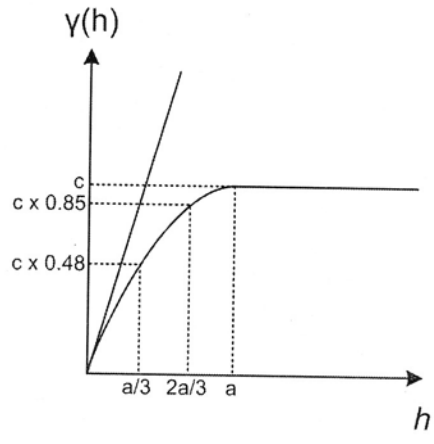
Σφαιρικό

Το σφαιρικό μοντέλο είναι το πιο κοινά χρησιμοποιούμενο μοντέλο βαριογράμματος, και συχνά προσαρμόζεται μαζί με ένα μοντέλο φαινόμενου κόκκου. Το σφαιρικό μοντέλο έχει μια απλή πολυονυμική έκφραση και ορίζεται ως εξής:

$$\gamma(h) = C\left[\frac{3}{2}\frac{|h|}{a} - \frac{1}{2}\frac{|h|^3}{a^3}\right] \quad \gamma\iota\alpha \ |h| \le a$$

$$\gamma(h) = C \qquad \gamma\iota\alpha \ |h| > a$$

όπου a είναι το εύρος του βαριογράμματος και C είναι η οριακή τιμή.



Σχήμα 15. Το σφαιρικό μοντέλο

Το σφαιρικό μοντέλο ταιριάζει καλά σε ότι βλέπουμε συνήθως στις μεταλλευτικές μεταβλητές: μια ημι-γραμμική συμπεριφορά κοντά στην αρχή ακολουθούμενη από σταθεροποίηση στην οριακή τιμή.

Μοντέλο Δύναμης

Τα μοντέλα συνάρτησης δύναμης χρησιμοποιούνται συχνά στο γεωστατιστικό λογισμικό. Αυτά τα μοντέλα δεν έχουν οριακή τιμή (δεν είναι μεταβατικά) και έχουν την μορφή:

$$\gamma(h) = \varpi |h|^{\lambda}$$

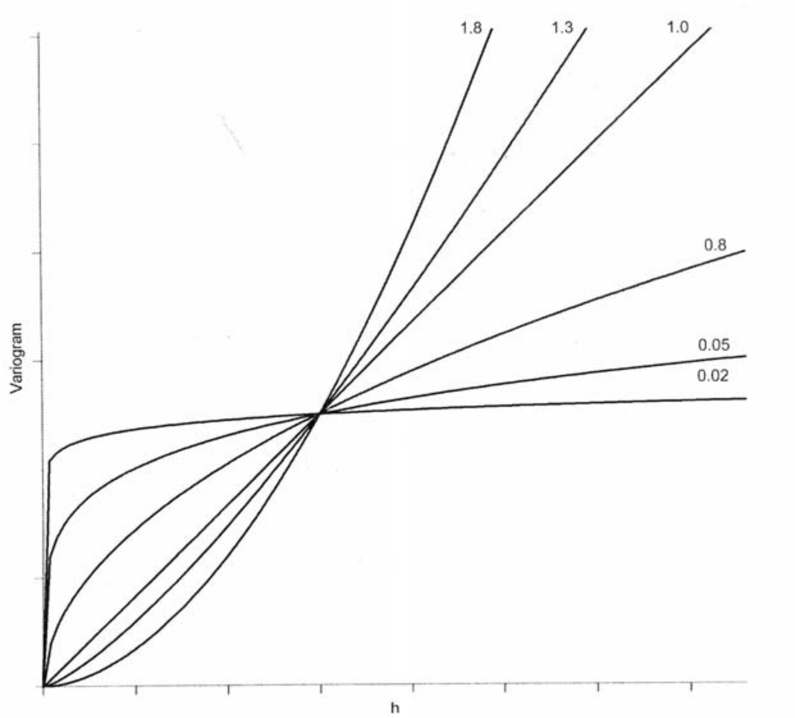
όπου λ είναι μια δύναμη μεταξύ του 1 και 2.

Συγκεκριμένη περίπτωση είναι το γραμμικό μοντέλο όπου το λ ισούται με μηδέν.

$$\gamma(h) = \varpi |h|$$

όπου ω είναι η κλίση του βαριογράμματος. Για το γραμμικό μοντέλο η τιμή του βαριογράμματος είναι απλά ανάλογη του h.

Το γραμμικό μοντέλο μπορεί να χρησιμοποιηθεί κοντά στην αρχή γιατί εκεί το βαριόγραμμα παρουσιάζει γραμμική συμπεριφορά.



Σχήμα 16. Μοντέλα δύναμης

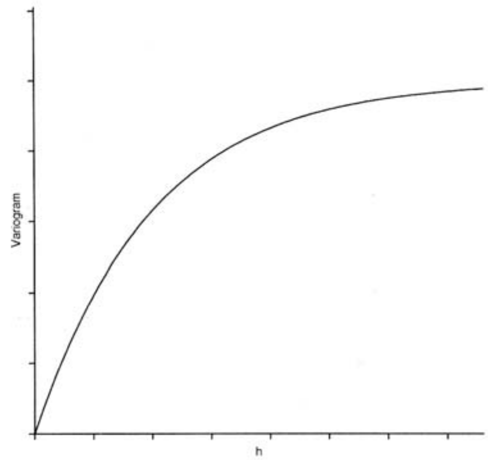
Εκθετικό Μοντέλο

Το εκθετικό μοντέλο έχει την εξής μορφή:

$$\gamma(h) = C\left[1 - e^{-\frac{h}{a}}\right]$$

Όπου C είναι η ασύμπτωτη της καμπύλης του μοντέλου. Εδώ το a δεν είναι το εύρος αλλά μια παράμετρος απόστασης που ελέγχει την χωρική έκταση της συνάρτησης.

Παρόλο που μερικές φορές εφαρμόζεται σε μεταλλευτικά δεδομένα, το μοντέλο αυτό βρίσκει περισσότερη εφαρμογή σε μη-μεταλλευτικές εφαρμογές, για παράδειγμα στην χημεία εδάφους.[5]



Σχήμα Εκθετικό μοντέλο

Gaussian

Το μοντέλο του Gauss εφαρμόζεται μερικές φορές στο γεωστατιστικό λογισμικό. Η μορφή του είναι η εξής:

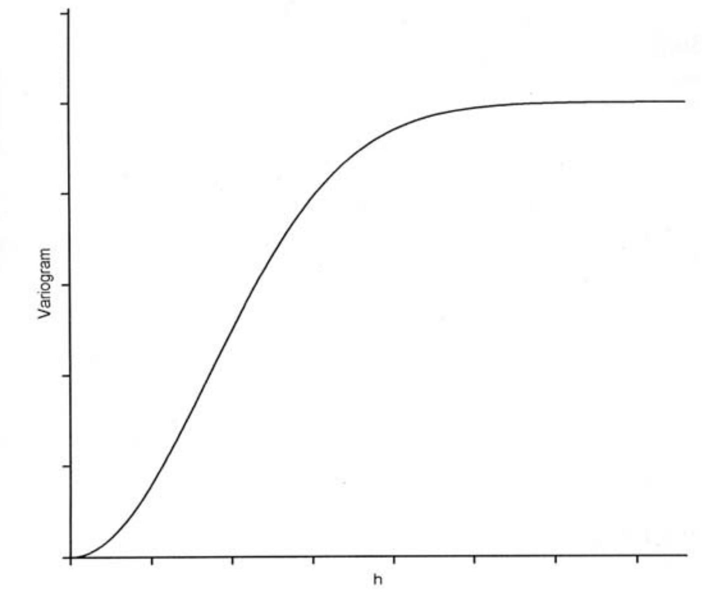
$$\gamma(h) = C\left[1 - e^{-\frac{|h|^2}{a^2}}\right]$$

Όπου, όπως και με το εκθετικό μοντέλο, το C είναι η ασύμπτωτη της καμπύλης του μοντέλου. Πάλι το a δεν είναι το εύρος αλλά μια παράμετρος της απόστασης που ελέγχει την χωρική έκταση της συνάρτησης.

---

[5] Webster και Oliver, 1990

Το Gaussian αντιπροσωπεύει υπερβολικά συνεχή συμπεριφορά στην αρχή. Στην πράξη ένα τέτοιο βαριόγραμμα είναι ακατάλληλο για μια μεταβλητή περιεκτικότητας. Η μόνη του εφαρμογή είναι για πολύ ομαλές, συνεχείς μεταβλητές όπως η τοπογραφία. Μερικές γεωλογικές επιφάνειες επίσης μπορεί να μοντελοποιηθούν επαρκώς από ένα Gaussian βαριόγραμμα.



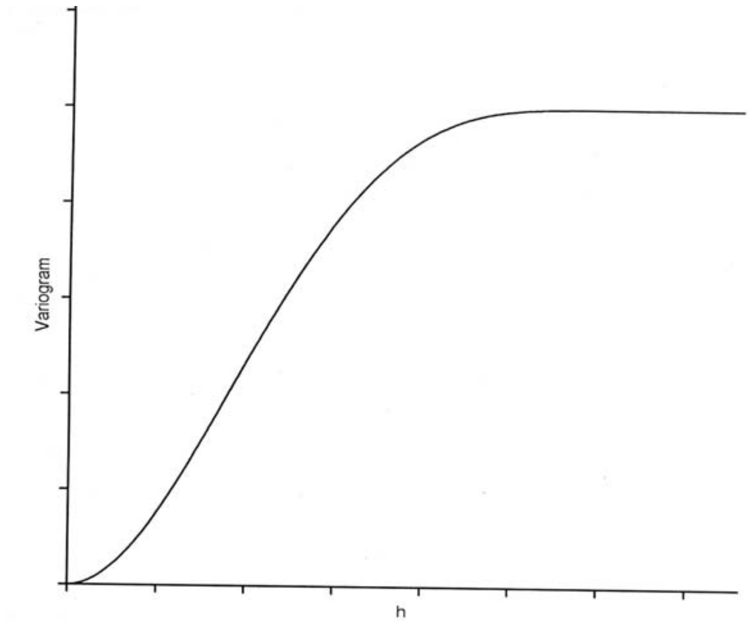
Σχήμα 17 Μοντέλο του Gauss

Είναι σημαντικό ότι το μοντέλο αυτό, εάν χρησιμοποιηθεί για kriging, θα πρέπει να πάντα να συνδυάζεται με ένα μικρό φαινόμενο κόκκου για να αποφεύγονται αριθμητικές αστάθειες στο kriging.

Κυβικό Μοντέλο

Το κυβικό μοντέλο ορίζεται ως εξής:

$$\gamma(h) = C\left[7\frac{h^2}{a^2} - \frac{35}{4}\frac{h^3}{a^3} + \frac{7}{2}\frac{h^5}{a^5} - \frac{3}{4}\frac{h^7}{a^7}\right], |h| \le a$$
$$\gamma(h) = C, |h| > a$$



Σχήμα 18 Κυβικό μοντέλο

Συνδυασμός Μοντέλων [n]

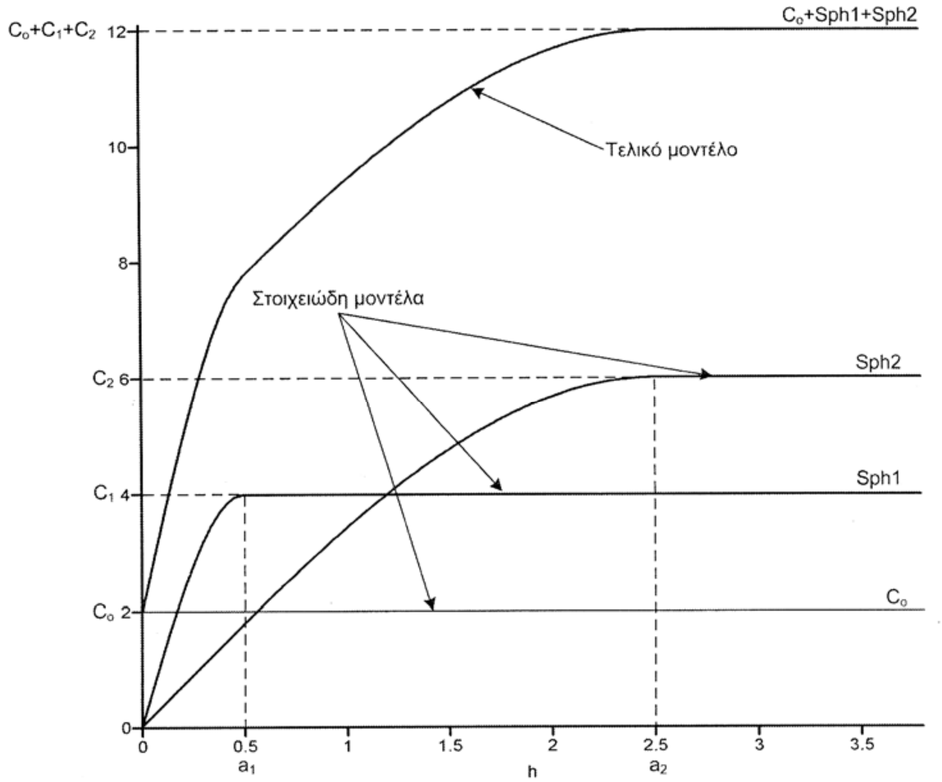
Όλα τα μοντέλα που αναφέραμε (μαζί με τα μοντέλα για το φαινόμενα κόκκου) περιγράφουν απλές καμπύλες ή ευθείες γραμμές. Τα βαριογράμματα που συναντούμε στην πράξη στις μεταλλευτικές (και σε άλλες) εφαρμογές είναι συχνά πιο πολύπλοκα και κανένα από τα παραπάνω μοντέλα δεν φαίνεται κατάλληλο.

Η προσαρμογή μοντέλων σε τέτοια βαριογράμματα γίνεται καλύτερα με συνδυασμούς δύο ή περισσοτέρων από τα παραπάνω μοντέλα.

Τα μοντέλα αυτά προσθέτονται μεταξύ τους

$$\gamma(h) = \gamma_1(h) + \gamma_2(h) + \gamma_3(h) + \ldots$$

Ένα παράδειγμα φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 19 Παράδειγμα συνδυασμού μοντέλων

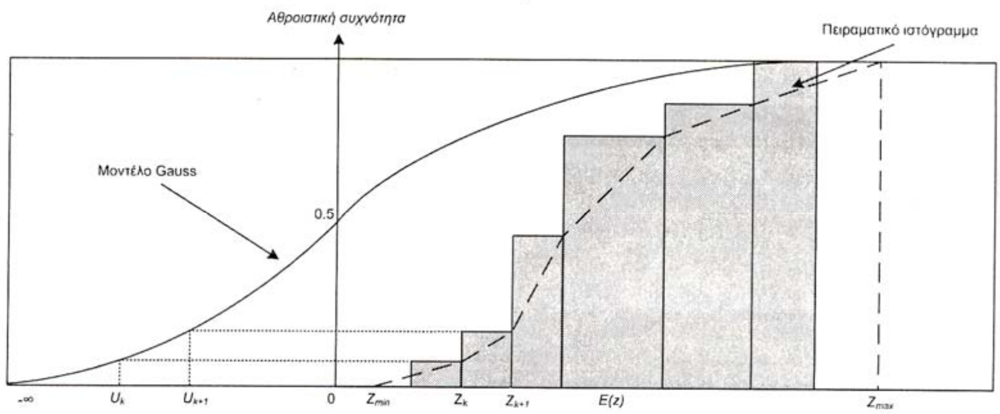
Βαριογραφία Μετασχηματισμών

Παίρνοντας τον λογάριθμο κάθε δείγματος πριν τον υπολογισμό του βαριογράμματος οδηγεί σε σημαντικά καλύτερη βαριογραφία.

Μετασχηματισμός Gauss[o]

Ένας μετασχηματισμός Gauss είναι ένας μετασχηματισμός δεδομένων που οδηγεί σε ένα κανονικό ιστόγραμμα.

Υπάρχουν δύο τρόποι για να γίνει. Γραφικά (δείτε σχήμα) και με την πολυονυμική επέκταση Hermite.



Σχήμα 20. Γραφικός τρόπος μετασχηματισμού Gauss

Η μαθηματική σχέση δίνεται από

$$\gamma_Z = \sum_{n=1}^{N} \frac{\Psi_n^2}{n!} \left[1 - \left\{ 1 - \gamma_\gamma(h) \right\}^n \right]$$

όπου Ψn είναι τα πολυώνυμα Hermite, γZ(h) είναι το βαριόγραμμα σε πραγματικές τιμές και γγ(h) είναι το βαριόγραμμα των μετασχηματισμένων κατά Gauss τιμών6.

[6] Guidal (1987)

Μέθοδος των ελάχιστων τετραγώνων στο νέφος του βαριογράμματος Werner G. Muller [p]

Εισαγωγή

Η τεχνική αυτή ξεκινά από το ορισμό του βαριογράμματος (όπως συνηθίζεται να ονομάζεται) που δίνεται από τον τύπο:

$$2\gamma(h) = Var[Z(\mathbf{s}+h) - Z(\mathbf{s})] \qquad \forall \mathbf{s}, s+h \in \Re^2$$

Στην απλούστερη μορφή ισχύουν οι παρακάτω ιδιότητες

$$E\left\{\left[Z(s_i) - Z(s_j)\right]\right\} = 0$$

όπου τα $Z(s_1), Z(s_2), ...Z(s_n)$ είναι τα δεδομένα ενός τυχαίου πεδίου με δείγματα σε n σημεία. και

$$E\left\{\left[Z(s_i) - Z(s_j)\right]^2\right\} = 2\gamma(h_{ij}, \theta)$$

Όπου $h_{ij} = \left\| s_i - s_j \right\|$ είναι η απόσταση που χωρίζει τις τοποθεσίες si και sj

Μία συνήθης προσέγγιση του βαριογράμματος (ονομάζεται και ημιδιασπορά) είναι:

$$2\hat{\gamma}_k = \frac{1}{N_{H_k}} \sum_{s_i, s_j \in H_k} \left[Z(s_i) - Z(s_j)\right]^2$$

Το σύνολο Hk περιέχει όλα τα ζεύγη σημείων που με αποστάσεις έτσι που hij = hk. Το Nk είναι τότε το πλήθος των ζευγών που ανήκουν σε αυτή την κλάση.

Με βάση αυτές τις υποθέσεις το άρθρο συνεχίζει με μία αναφορά στη χρήση τηε γενικευμένης μεθόδου των ελαχίστων τετραγώνων (GLS) για μια παραμετρική προσέγγιση του βαριογράμματος από διάφορους ερευνητές όπως οι Matheron (1963), Cressie(1985) κα.

Απευθείας προσέγγιση από το νέφος του βαριογράμματος.

Με τη δύναμη που έχουν οι σημερινοί υπολογιστές ο συγγραφέας εκφράζει την άποψη ότι μπορεί να γίνει χρήση όλων των δεδομένων χωρίς την χρήση ομαδοποίησης και κλάσεων (εφόσον το πλήθος των δεδομένων δεν είναι υπερβολικά μεγάλο).

Θεωρεί ότι έτσι διαλέγουμε κάθε Hk να περιέχει μόνο ένα στοιχείο – το πλήθος των ζευγαριών ονομάζεται τότε το νέφος του βαριογράμματος και προσπαθούμε να προσεγγίσουμε το $\gamma(h,\hat{\theta})$ απευθείας από το νέφος του (ημι-) βαριογράμματος χωρίς ομαδοποίηση και χωρίς να παίρνουμε μέσους όρους.

Ο G. Muller χρησιμοποιεί την απόδειξη συνέπειας (proof of consistency) του Fedorov (1974) για να βρεί μία απλή διόρθωση στην $\gamma(h,\hat{\theta})$. Χωρίζει της «παρατηρήσεις»

$\hat{\gamma}_k = \gamma(h_k,\theta_t) + \varepsilon_\kappa$ όπου θt είναι η πραγματική παράμετρος.

Τότε το άθροισμα που πρέπει να ελαχιστοποιηθεί είναι:

$$\sum_{k=1}^{\infty} \gamma^{-2}(h_k,\theta)\left[\gamma(h_k,\theta) - \gamma(h_k,\theta_t)\right]^2 + \sum_{k=1}^{\infty} \gamma^{-2}(h_k,\theta)\varepsilon_k{}^2$$

Αυτό λόγω του νόμου των μεγάλων αριθμών δίνει

$$\sum_{k=1}^{\infty}\left[\frac{\gamma(h_k,\theta_t)}{\gamma(h_k,\theta)} - 1\right]^2 + \sum_{k=1}^{\infty}\frac{2\gamma^2(h_k,\theta_t)}{\gamma^2(h_k,\theta)}$$

Που σημαίνει (μετά από πράξεις) ότι καταλήγουμε στην ελαχιστοποίηση του:

$$\sum_{k=1}^{\infty}\left[\frac{\gamma(h_k,\theta_t)}{\gamma(h_k,\theta)} - \frac{1}{3}\right]^2$$

και έτσι το $\gamma(h,\hat{\theta})$ είναι μία συνεπή προσέγγιση του $3\gamma(h,\theta_t)$ αντί για του $\gamma(h,\theta_t)$

Επαναλαμβανόμενη επαναστάθμιση – "Iterative reweighting"

Εάν το πλήθος των παρατηρήσεων είναι σχετικά μικρό και οι υπολογιστική δύναμη επιτρέπει, μπορεί να βρεθεί ένας ποιο λεπτομερής εκτιμητής.

Η μέθοδος αυτή έχει προταθεί από διάφορους ερευνητές αλλά εδώ παρουσιάζεται του Fedorov (1989) που ξανά σωστά περιέγραψε έναν επαναληπτικό αλγόριθμο για την περίπτωση όπου:

$$Cov(\hat{\gamma}) = diag(2\gamma^2(\theta)/\mathrm{N}_{\mathrm{H}_k})$$

Ο Cressie δίνει ότι για ένα πεδίο τύπου Gauss και $N_{H_k} = 1$ έχουμε:

$$Cov(\hat{\gamma})_{kl} = \frac{1}{2}\left[\gamma\left(\left\|s_j - s_{i'}\right\|\right) + \gamma\left(\left\|s_i - s_{j'}\right\|\right) - \gamma\left(\left\|s_i - s_{i'}\right\|\right) - \gamma\left(\left\|s_j - s_{j'}\right\|\right)\right]^2$$

όπου $\left\| s_i - s_{j'} \right\| = h_l$ και οι τιμές της $Cov\left[(\gamma(\hat{\theta}_r) \right]$ μπορούν να υπολογιστούν από μια παραμετρική μορφή της παραπάνω εξίσωσης.

Στην πράξη μαντεύουμε μία αρχική τιμή και οι επαναλήψεις ξεκινούν από αυτήν (ας πούμε $\hat{\theta}_0 = \hat{\theta}_{OLS}$) (Original Least – Square). Η διαδικασία αποδίνει τότε συνεπείς και ασυμπτοτικά αποδοτικές προσεγγίσεις.

Αποτελέσματα προσομοίωσης

Για να αξιολογηθεί η συμπεριφορά σε μικρά δείγμα των διάφορων εκτιμητών που χρησιμοποιούνται ένας μεγάλος αριθμός πειραμάτων προσομοίωσης έχει εκτελεσθεί.

Για πρακτικούς λόγους έχει χρησιμοποιηθεί ή προσέγγιση:

$$Cov\left[\gamma(h_k, \theta), \gamma(h_{k'}, \theta) \right] \propto \begin{cases} \gamma^2(h_k, \theta_t) & k = k' \\ 0 & \alpha\lambda\lambda\iota\omega\varsigma \end{cases}$$

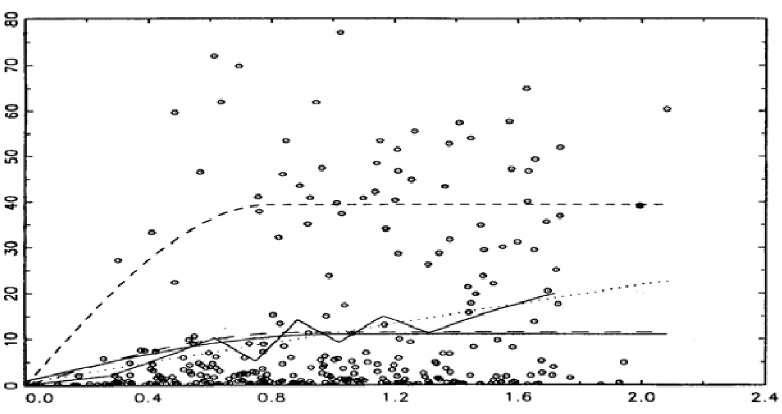
που φαίνεται λογικό για όλα τα συγκρινόμενα θ.

Για την προσέγγιση χρησιμοποιήθηκε ίσως το ποιο ευρείας χρήσης παραμετρικό μοντέλο βαριογράμματος (το λεγόμενο σφαιρικό (ημι) βαριόγραμμα)

$$\gamma_s(h, \theta) = \begin{cases} 0, & h = 0 \\ \theta_1 + \theta_2 \left[\frac{3}{2}\left(\frac{h}{\theta_3}\right) - \frac{1}{2}\left(\frac{h}{\theta_3}\right)^2 \right], & 0 < h \le \theta_3 \\ \theta_1 + \theta_2 & h > \theta_3 \end{cases}$$

Το διάγραμμα που ακολουθεί δείχνει μια τυπική περίπτωση

W.G. Müller / Statistics & Probability Letters 43 (1999) 93–98



Το νέφος του ημι - βαριογραμματος και μια σειρά από προσεγγίσεις για δεδομένα μιας τυπικής προσομοίωσης από 25 παρατηρήσεις (οριζόντιος άξονας είναι το h και κάθετος το $\hat{\gamma}$.

Table 1
Mean (upper line) and median (lower line) of the performance criterion of various variogram estimators ($\theta_t = (1, 10, 1)^T$; $K = 10$ for binned estimators)

n	Binned OLS	OLS	Binned WLS	WLS	Corrected WLS	Iterated GLS
5	[illegible]	0.778	[illegible]	3.301	0.393	0.617
		0.340		0.563	0.310	0.257
10	0.303	0.301	0.568	3.146	0.204	0.230
	0.162	0.167	0.219	1.581	0.130	0.158
15	0.248	0.252	0.317	3.281	0.188	0.179
	0.135	0.131	0.167	1.683	0.125	0.125
20	0.239	0.237	0.262	2.947	0.178	0.165
	0.095	0.091	0.113	1.680	0.109	0.093
25	0.184	0.180	0.216	3.027	0.133	0.136
	0.079	0.082	0.098	1.607	0.100	0.055

Ο πίνακας δείχνει συγκριτικά το κριτήριο απόδοσης του κάθε εκτιμητή για τις μέθοδες που περιγράψαμε (OLS – ελαχίστων τετραγώνων και WLS – Σταθμισμένη Ελαχίστων Τετραγώνων - ομαδοποιημένα και μη δεδομένα, και WLS με τη διόρθωση και με την επαναλαμβανόμενη επαναστάθμιση)

Τα αποτελέσματα δείχνουν ότι είναι προτιμότερη η μέθοδος WLS με επαναλαμβανόμενη επαναστάθμιση αφού παρατηρούμε ότι όσο μεγαλώνει

το n, μικραίνουν οι διαφορές.

Kriging[q]

Το kriging είναι μια μέθοδος εκτίμησης που λαμβάνει υπόψη τους παρακάτω παράγοντες ως επιθυμητούς σε έναν εκτιμητή:

1. Ο αριθμός των δειγμάτων και η ποιότητα των δεδομένων σε κάθε δείγμα. Ο εκτιμητής πρέπει να προσαρμόζεται ανάλογα με το δείγμα.

2. Η γεωμετρία των δειγμάτων στο κοίτασμα.

3. Η απόσταση μεταξύ των δειγμάτων στο κοίτασμα.

4. Η χωρική συνέχεια των μεταβλητών παρεμβολής.

Οι ιδιότητες αυτές προέρχονται από το γεγονός ότι το kriging χρησιμοποιεί τον μοντελοποιημένο χωρικό συσχετισμό για την απόδοση βαρών.

BLUE (Best Linear Unbiased Estimator - Βέλτιστος Γραμμικός Αμερόληπτος Εκτιμητής)

Best (Βέλτιστος)

Το kriging είναι ο ‘βέλτιστος’ με την έννοια ότι έχει το ελάχιστο μέσο τετραγωνισμένο σφάλμα, δηλαδή η αναμενόμενη τετραγωνισμένη διαφορά μεταξύ της εκτίμησης ${Z^*}_0$ και της πραγματικής τιμής Z_0.

$$E[{Z^*}_0 - Z_0]^2$$

Είναι η ελάχιστη για όλους τους πιθανούς γραμμικούς εκτιμητές.

Linear (Γραμμικός)

Γραμμικοί εκτιμητές είναι εκείνοι που σχηματίζονται με γραμμική ζύγιση των διαθέσιμων δειγμάτων, δηλαδή:

$$Z_0^* = \sum_{i=1}^{N} \lambda_i Z(x_i)$$

Όπου τα λ_i είναι τα βάρη, και η εκτίμηση Z^*_0 είναι ένα ζυγισμένο άθροισμα όλων των δεδομένων τιμών ($Z(x_i)$) σε κάθε σημείο x_i.

Unbiased (αμερόληπτος)

Η συνθήκη αμεροληψίας είναι σημαντική. Καθορίζει ότι το αναμενόμενο σφάλμα:

$$E[Z_0^* - Z_0]$$

είναι ίσο με μηδέν.

Estimator (εκτιμητής)

Το kriging είναι ένας εκτιμητής.

Έτσι το kriging είναι ο βέλτιστος γραμμικός αμερόληπτος γραμμικός αμερόληπτος εκτιμητής της ποσότητας που εκτιμάται. Το πετυχαίνει δίνοντας στα δείγματα που οδηγούν στην εκτίμηση, βάρη σχεδιασμένα να δίνουν μια εκτίμηση με το ελάχιστο μέσο τετραγωνισμένο σφάλμα.

Λειτουργία του Kriging

Την εκτίμηση την παίρνουμε χρησιμοποιώντας ένα γραμμικό ζυγισμένο τρόπο:

$$Z^*_0 = \sum_{i=1}^{N} \lambda_i Z(x_i)$$

Επιλέγουμε τα βάρη λi με τον καλύτερο δυνατό τρόπο για να πετύχουμε τις επιθυμητές ιδιότητες που περιγράψαμε παραπάνω.

σημ. για N δείγματα υπάρχουν N βάρη: $\lambda_i = \lambda_1, \lambda_2, \lambda_3, ..., \lambda_N$

Επιλέγοντας τα 'Βέλτιστα' βάρη

Η Συνθήκη Αμεροληψίας

$$E[Z_0^* - Z_0] = 0$$
$$\mathrm{E}\left[\left\{\sum_{i=1}^{N} \lambda_i Z(x_i)\right\} - Z_0\right] = 0$$

Εάν η Z(x) είναι στάσιμη, η προσδοκία για το Z(x) είναι ίση με το μέσο:

$$E[Z(x_i)] = m$$

και η προσδοκία Z_0 είναι ίση με το μέσο:

$$E[Z_0] = m$$

Κάνοντας τις αντικαταστάσεις και με την προϋπόθεση ότι η Z(x) είναι στάσιμη καταλήγουμε στο συμπέρασμα ότι (βλέπε π.χ. Εισαγωγή στη Γεωστατιστική, Καπαγερίδης) :

$$\sum_{i=1}^{N} \lambda_i = 1$$

Έτσι η συνθήκη αμεροληψίας ικανοποιείται όταν το άθροισμα των βαρών ισούται με 1.

Ο όρος του Βέλτιστου (best)

Όπως αναφέραμε θα πρέπει

$$E[Z^*{}_0 - Z_0]^2$$

να είναι ελάχιστο. Αυτό σημαίνει ότι με άλλα λόγια η Διακύμανση σφάλματος ή διακύμανση kriging να είναι ελάχιστη.

Η διακύμανση σφάλματος μπορεί μαθηματικά να αποδοθεί σε όρους της συνάρτησης βαριογράμματος (δείτε σχήμα).

$$E[Z^*{}_0 - Z_0]^2 = Var[Z^*{}_0 - Z_0]$$
$$= 2\sum \lambda_i \overline{\gamma}(x_i, V) - \sum\sum \lambda_i \lambda_j \gamma(x_i, x_j) - \overline{\gamma}(V, V)$$

όπου $\overline{\gamma}(x_i, V)$ είναι η μέση τιμή του βαριογράμματος υπολογισμένου μεταξύ x_i και του όγκου V, δηλαδή

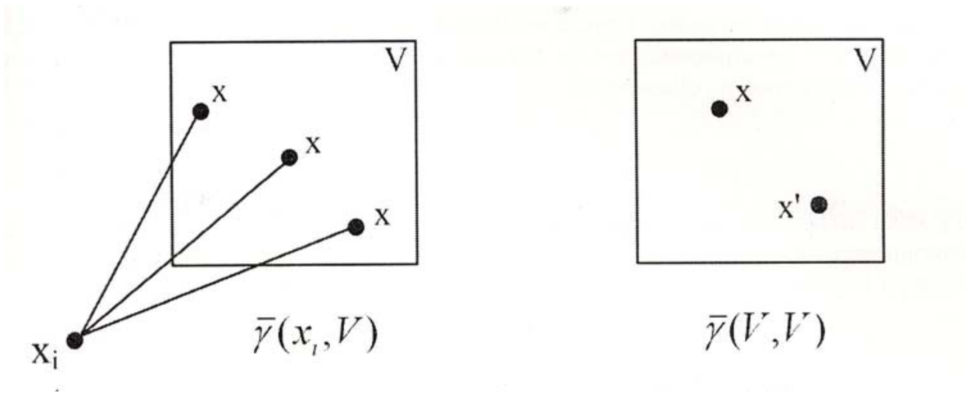
$$\overline{\gamma}(x_i, V) = \frac{1}{V}\int_V \gamma(x_i - x)dx$$

- γ(xi,xj) είναι η τιμή του βαριογράμματος μεταξύ των σημείων x_i και x_j δηλαδή μεταξύ των δειγμάτων.

-$\overline{\gamma}(V, V)$ είναι η μέση τιμή του βαριογράμματος μεταξύ δύο σημείων x και x' που κινούνται ανεξάρτητα μέσα στον όγκο V, δηλαδή

$$\overline{\gamma}(V, V) = \frac{1}{V^2}\iint \gamma(x - x')dxdx' = F(V)$$

Αυτή είναι η διακύμανση διασποράς των σημείων μέσα στον όγκο *V* (δείτε σχήμα)



Σχήμα: Διακυμάνσεις επέκτασης και διασποράς στις εξισώσεις kriging

Θέλουμε την ελαχιστοποίηση της διακύμανσης της εκτίμησης με τον όρο ότι το άθροισμα των βαρών να ισούται με 1. Πρέπει να λύσουμε ένα σύστημα *n+1* εξισώσεων με *n* αγνώστους κάθε μία με ένα από τα βάρη.

Η μέθοδος Lagrange

Η μέθοδος των *παραμέτρων Lagrange* ταιριάζει στη λύση αυτού του προβλήματος.

Για να ελαχιστοποιήσουμε κάποια συνάρτηση ως προς τα λ_i χρειάζεται να θέσουμε τις μερικές παραγώγους ως προς τα λ_i να είναι μηδέν.

Εισάγουμε την παράμετρο Lagrange ως εξής:

$$\phi = \sigma_e^2 - 2\mu\left(\sum_{i=1}^{N}\lambda_i - 1\right) = Var(z_0^* - z_0) - 2\mu\left(\sum_{i=1}^{N}\lambda_i - 1\right)$$

Θέτοντας τις μερικές παραγώγους

$$\frac{\partial\phi}{\partial\lambda_i}, \quad \frac{\partial\phi}{\partial\lambda_i}$$

ίσον με μηδέν βρίσκουμε τα λ_i που ικανοποιούν τα κριτήρια ελάχιστης διακύμανσης.

Καταλήγουμε στο ακόλουθο σύστημα γραμμικών εξισώσεων, γνωστό ως *κανονικό σύστημα kriging:*

$$\sum_{i=1}^{N}\lambda\gamma(x_i - x_j) + m = \bar{\gamma}(x_i,V), \quad i = 1,2,\ldots,N$$

$$\sum_{i=1}^{N}\lambda_i = 1$$

Το σύστημα kriging συχνά αναπαρίσταται με μορφή πινάκων

$$AX = B$$

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots\gamma_{1N} & 1 \\ \gamma_{21} & \gamma_{22} & \cdots\gamma_{2N} & 1 \\ . & . & . & . \\ \gamma_{N1} & \gamma_{N2} & \cdots\gamma_{NN} & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ . \\ \lambda_N \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{\gamma}(x_1, V) \\ \bar{\gamma}(x_2, V) \\ . \\ \bar{\gamma}(x_N, V) \\ 1 \end{bmatrix}$$

Όπου:

Α είναι ένας πίνακας των συνδιακυμάνσεων μεταξύ δειγμάτων (σε όρους βαριογράμματος)

Χ είναι ένας πίνακας που περιέχει τα βάρη που θέλαμε να υπολογίσουμε

Β είναι ένας πίνακας που πειέχει τις συνδιακυμάνσεις δείγματος – μπλοκ (πάλι σε όρους βαριογράμματος)

Για μια αποδεκτή συνάρτηση βαριογράμματος η λύση είναι

Χ= Β/Α

Εφαρμογές Kriging[r]

Σημειακό Kriging

Το σημειακό Kriging είναι μια διαδικασία κατά την οποία εκτιμάται η τιμή σημείων χρησιμοποιώντας τις παραμέτρους του επιλεγμένου βαριογράμματος.

Με την εισαγωγή των παραμέτρων του βαριογράμματος στο πρόγραμμα εκτίμησης, κάθε τιμή δείγματος Z_i αφαιρείται με τη σειρά και γίνεται εκτίμηση του σημείου με kriging. Για κάθε σημείο υπολογίζονται οι παρακάτω παράμετροι:

Α) Το μέσο μαθηματικό σφάλμα

$$\frac{\sum_{i=1}^{N}(Z_i - Z^*_i)}{N}$$

όπου Z_i είναι η πραγματική τιμή σε κάθε σημείο και N είναι το πλήθος των σημείων.

Β) Το μέσο απόλυτο σφάλμα

$$\frac{\sum_{i=1}^{N}\left|Z_i - Z^*_i\right|}{N}$$

Είναι ο μέσος των διαφορών αλλά αγνοώντας το πρόσημο.

Γ) Τη μέση διακύμανση

$$\sigma_\kappa{}^2 = \frac{\sum_{i=1}^{N}\sigma^2{}_{\kappa_i}}{N}$$

Δ) Το μέσο τετραγωνισμένο σφάλμα εκτίμησης

$$\sigma^2{}_{\mathrm{E}} = \frac{\sum_{i=1}^{N}(Z_i - Z^*_i)^2}{N}$$

Ε) Το πλήθος σημείων που χρησιμοποιούνται στο σημειακό Kriging *N*.

ΣΤ) Το λόγο του μέσου τετραγωνισμένου προς τη μέση διακύμανση kriging $\frac{\sigma^2{}_{\mathrm{E}}}{\sigma^2{}_\kappa}$

Το μοντέλο είναι καλή προσέγγιση του θεωρητικού βαριογράμματος όταν:

1. Η η τιμή του μέσου μαθηματικού σφάλματος τείνει στο μηδέν και δεν είναι μεγαλύτερο από 1% των τιμών Z_i για συνολική αμεροληψία.

2. Το μέσο απόλυτο σφάλμα είναι μεγαλύτερο από το μέσο μαθηματικό σφάλμα και όσο δυνατόν μικρότερο.

3. Ο λόγος $\frac{\sigma^2{}_{\mathrm{E}}}{\sigma^2{}_\kappa}$ πρέπει να πλησιάζει την μονάδα, και

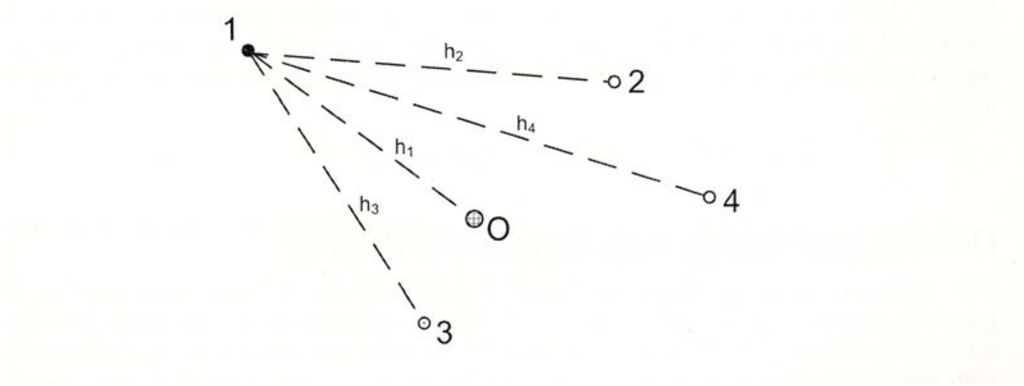
4. Το *N* πρέπει να είναι όσο το δυνατό μεγαλύτερο.

Μαθηματική Βάση του Σημειακού Kriging

Το σημείο ‘Ο’ πρόκειται να εκτιμηθεί χρησιμοποιώντας τα τέσσαρα γειτονικά σημεία. Έτσι η πρώτη από τις εξισώσεις του kriging που θα πρέπει να επιλυθούν για την λήψη των συντελεστών K_i είναι οι εξής:

$$K_1\sigma_{1,1} + K_2\sigma_{1,2} + K_3\sigma_{1,3} + K_4\sigma_{1,4} + \mu = \sigma_{\mathrm{O},1}$$

όπου $\sigma_{1,1}$ είναι η συνδιακύμανση μεταξύ δείγματος 1 και του εαυτού του (C_0), $\sigma_{1,2}$ είναι η συνδιακύμανση μεταξύ δείγματος 1 και 2, κλπ. μ είναι ο συντελεστής Lagrange και $\sigma_{O,1}$ είναι η συνδιακύμανση μεταξύ του σημείου εκτίμησης και του δείγματος 1.



Σχήμα Εκτίμηση σημείου με σημειακό kriging και χρήση τεσσάρων γειτονικών δειγμάτων

Μπορούμε να γράψουμε άλλες τρεις εξισώσεις για τα δείγματα 2, 3, 4 όπως κάναμε και για το δείγμα 1. Η $(n + 1)^{\eta}$ εξίσωση (η πέμπτη στην περίπτωσή μας) είναι επομένως:

$K_1+K_2+K_3+K_4 = 1$

Οι συνδιακυμάνσεις των εξισώσεων του kriging μπορούν να αντικατασταθούν από τις τιμές του μοντέλου του βαριογράμματος που έχουμε στο πειραματικό βαριογράμμα. Έτσι το πλήρες σετ εξισώσεων του kriging είναι το εξής:

$$
\begin{aligned}
&K_1\gamma_{1,1} + K_2\gamma_{1,2} + K_4\gamma_{1,3} + K_5\gamma_{1,4} + \mu = \gamma_{\mathrm{O},1} \\
&K_1\gamma_{2,1} + K_2\gamma_{2,2} + K_4\gamma_{2,3} + K_5\gamma_{2,4} + \mu = \gamma_{\mathrm{O},2} \\
&K_1\gamma_{3,1} + K_2\gamma_{3,2} + K_4\gamma_{3,3} + K_5\gamma_{3,4} + \mu = \gamma_{\mathrm{O},3} \\
&K_1\gamma_{4,1} + K_2\gamma_{4,2} + K_4\gamma_{4,3} + K_5\gamma_{4,4} + \mu = \gamma_{\mathrm{O},4} \\
&K_1 + K_2 + K_3 + K_4 = 1
\end{aligned}
$$

Οι εξισώσεις αυτές μετατρέπονται σε πίνακα για επίλυση με κάποιο κατάλληλο πρόγραμμα:

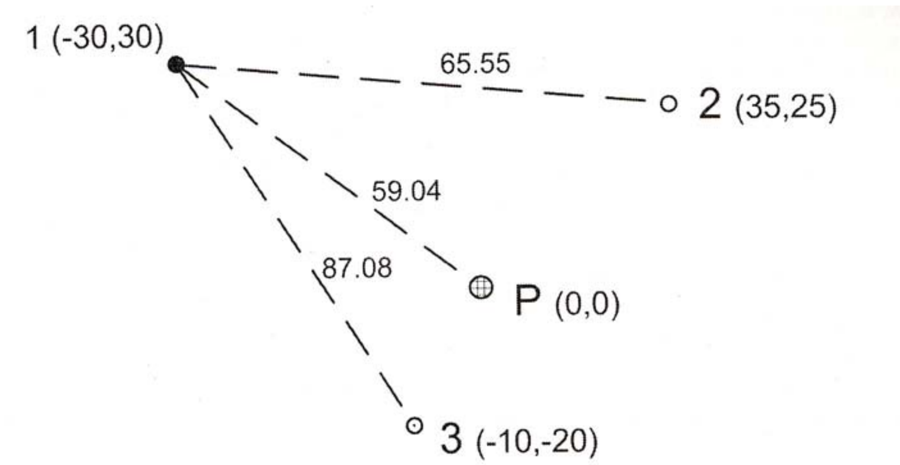
$$
\begin{pmatrix}
\gamma_{1,1} & \gamma_{1,2} & \gamma_{1,3} & \gamma_{1,4} & 1 \\
\gamma_{2,1} & \gamma_{2,2} & \gamma_{2,3} & \gamma_{2,4} & 1 \\
\gamma_{3,1} & \gamma_{3,2} & \gamma_{3,3} & \gamma_{3,4} & 1 \\
\gamma_{4,1} & \gamma_{4,2} & \gamma_{4,3} & \gamma_{4,4} & 1 \\
1 & 1 & 1 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
K_1 \\ K_2 \\ K_3 \\ K_4 \\ \mu
\end{pmatrix}
=
\begin{pmatrix}
\gamma_{\mathrm{O},1} \\ \gamma_{\mathrm{O},2} \\ \gamma_{\mathrm{O},3} \\ \gamma_{\mathrm{O},3} \\ 1
\end{pmatrix}
$$

Εφόσον υπολογιστούν οι τιμές των K_i λύνοντας αυτές τις εξισώσεις, η τιμή στο σημείο 'O' καθορίζεται αθροίζοντας τα προϊόντα κάθε τιμής δείγματος με τον συντελεστή βάρους τους.

Η διακύμανση kriging, το σφάλμα που συμβαίνει κατά την εκτίμηση σημειακού δείγματος, υπολογίζεται με τη βασική εξίσωση:

$$\sigma_{\kappa}^{2} = \sigma_{0,0}^{2} - \sum K_i x \sigma_{0,\iota} + \mu$$

Μια εφαρμογή του σημειακού kriging



Σχήμα Διάταξη δειγμάτων παραδείγματος σημειακού kriging (Καπαγερίδης, Εισαγωγή στην Γεωστατιστική σελ 171),

Για να εκτιμήσουμε το σημείο *P* του σχήματος (με συντεταγμένες 0,0) χρησιμοποιώντας τις τιμές των δειγμάτων 1, 2 και 3 των οποίων οι συντεταγμένες είναι (-30, 30), (35, 25) και (-10, -20) αντίστοιχα, θα πρέπει να κατασκευάσουμε μια σειρά εξισώσεων η κάθε μια εκ των οποίων εκφράζει την σχέση κάθε δείγματος προς κάθε άλλο στην περιοχή / όγκο ανίχνευσης και το υπό εκτίμηση σημείο. Επειδή τα δείγματα βρίσκονται εντός της ακτίνας ανίχνευσης (μόλις μικρότερη από το εύρος) έχουν κάποια σχέση μεταξύ τους – υπάρχει συνδιακύμανση μεταξύ κάθε ζεύγους δειγμάτων ή μεταξύ ενός δείγματος και του σημείου *P*.

Η επιρροή του καθενός στην τιμή του *P* θα είναι μοιρασμένη και επομένως θα πρέπει να σχηματίσουμε εξισώσεις που ορίζουν αυτήν την επιρροή.

Πρώτα καθορίζουμε την 'γεωστατιστική απόσταση' του δείγματος 1 από το *P* και από τα δείγματα 2 και 3 με τη χρήση του συντελεστή διευθυντικής ανισοτροπίας *λ*. Λαμβάνοντας υπόψη μόνο τη διαφορά των συντεταγμένων *x* και *y*, *dx* και *dy* αντίστοιχα:

$$r = \left[dx^2 + (\lambda dy)^2 \right]^{1/2}$$

Υποτίθεται ότι ο κύριος άξονας ανισοτροπίας είναι παράλληλα στο άξονα *x*.

Στο παράδειγμα θεωρούνται οι παρακάτω παράμετροι βαριογράμματος:

$C_o = 2{,}0$ $C = 3{,}1$ $a_x = 100m$ $a_y = 59m$

Έτσι λ = 1,695

Από τα παραπάνω μπορούμε να υπολογίσουμε τις απαραίτητες αποστάσεις ως εξής:

$$r_{1,P} = [(-30)^2 + (30 \times 1{,}695)^2]^{1/2} = 59{,}04m$$
$$r_{1,2} = 65{,}55m$$
$$r_{1,3} = 87{,}08m$$

Οι συντελεστές *kriging* καθορίζονται λύνοντας όλες τις εξισώσεις του πίνακα (δείτε παραπάνω) οι οποίες περιλαμβάνουν και τον πολλαπλασιαστή Lagrange. Έτσι για το παράδειγμα παραπάνω έχουμε:

$$k_1\sigma_{1,1} + k_2\sigma_{1,2} + k_3\sigma_{1,3} + \mu = \sigma_{1,P}$$
$$k_1\sigma_{2,1} + k\sigma_{2,2} + k\sigma_{2,3} + \mu = \sigma_{2,P}$$
$$k_1\sigma_{3,1} + k\sigma_{3,2} + k\sigma_{3,3} + \mu = \sigma_{3,P}$$
$$k_1 + k_2 + k_3 = 1$$

Η τελευταία εξίσωση προκύπτει από την συνθήκη αμεροληψίας. Έτσι για *ν* δείγματα τότε θα έχουμε *ν+1* εξισώσεις. Οι τιμές συνδιακύμανσης μπορούν να οριστούν από την εξίσωση μοντέλου βαριογράμματος βάζοντας ως h τις αποστάσεις $r_{i,j}$. Το αποτέλεσμα της εξίσωσης αυτής είναι η ημι-διακύμανση $\gamma(r_{i,j})$ όμως, καθώς συνδιακύμανση = οριακή τιμή + ημι - διακύμανση μπορούμε να αφαιρέσουμε την $\gamma(r_{i,j})$ από την $C_o + C$ για να ορίσουμε την συνδιακύμανση για την συγκεκριμένη απόσταση δειγμάτων. Επιστρέφοντας στο παράδειγμα μας και το δείγμα 1, $\gamma(r_{1,1}) = C_o = 2{,}0$ καθώς η ημι – διακύμανση στο διάστημα 0 είναι ίση με C_o και επομένως η συνδιακύμανση $\sigma_{1,1} = C = 3{,}10$.

$$\gamma(r_{1,2}) = 2,0 + 3,1[1,5(65,55/100) - 0,5(65,55/100)^3]$$
$$= 2,0 + 2,611 \qquad \text{(σφαιρικό μοντέλο)}$$
$$= 4,61$$

Επομένως $\sigma_{1,2}$ = 2,0 +3,1 – 4,611 = 0,489, $\sigma_{1,3}$ = 0,0743, $\sigma_{1,P}$ = 0,6737.

Η πρώτη εξίσωση σημειακού kriging μπορεί να γραφεί ως εξής:

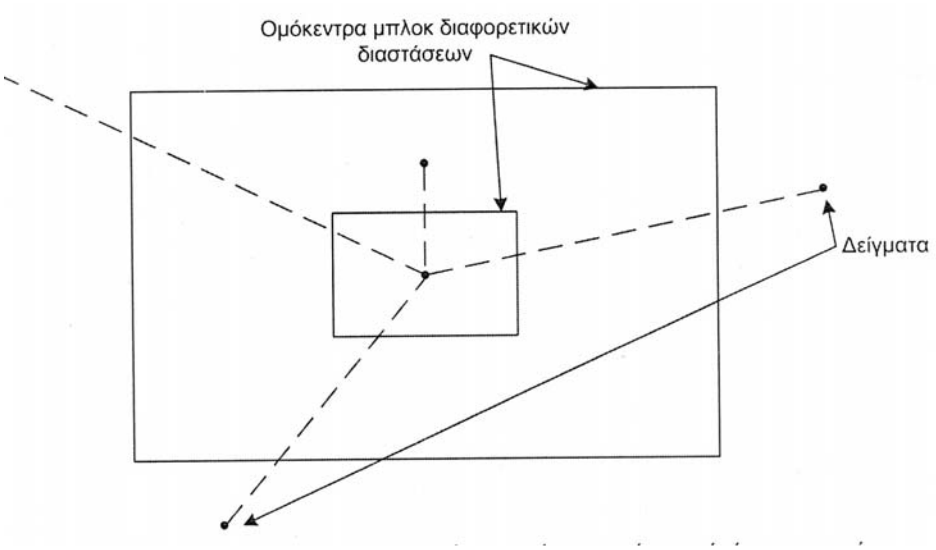
$$k_1 3,1 + k_2 0,489 + k_3 0,0743 + \mu = 0,6737$$

Εφόσον υπολογιστούν όλες οι συνδιακυμάνσεις για τις άλλες εξισώσεις και καθοριστούν οι τιμές των συντελεστών kriging και μ, λαμβάνεται η ζυγισμένη εκτίμηση σημειακού kriging από την εξίσωση

$$P = \sum^{n=3} k_i Z_i$$

Μπλοκ kriging

Η εκτίμηση μπλοκ μεταλλεύματος μπορεί να γίνει με οποιαδήποτε παραλλαγή της μεθόδου μπλοκ kriging. Οι τεχνικές αυτές αντικαθιστούν μεθόδους όπως η αντιστρόφου αποστάσεως (γραμμική, τετραγώνου, κυβική) ή άλλες τεχνικές πλεγμάτωσης και υπολογισμού ισοκαμπύλων αναλύσεων περιεκτικότητας, πάχους και συγκέντρωσης.



Σχήμα. Το πρόβλημα μη-προσέγγισης του όγκου εκτίμησης από τεχνικές όπως αντιστρόφου αποστάσεως

Στο μπλόκ kriging, ορίζεται μια ζώνη ανίχνευσης γύρω από κάθε μπλοκ μεταλλεύματος με διάμετρο ίση ή ελάχιστα μικρότερη από το διπλάσιο του εύρους σε κάθε διεύθυνση. Στην περίπτωση ισοτροπίας, μπορεί να χρησιμοποιηθεί μια σφαιρική περιοχή ανίχνευσης με ακτίνα ελάχιστα μικρότερη από το εύρος. Σε πολλές περιπτώσεις, τα μπλοκ διαιρούνται σε πολλαπλά σημεία, τα οποία εκτιμώνται με τα ίδια δείγματα ξεχωριστά και ο μέσος όρος τους αποδίδεται στο κέντρο βάρους του μπλοκ.

Κάθε δείγμα γεώτρησης που πέφτει εντός της περιοχής ανίχνευσης λαμβάνει ένα βέλτιστο συντελεστή βάρους (συντελεστής kriging) όπως αναφέρθηκε νωρίτερα, και υπολογίζεται η συνολική εκτίμηση των μπλοκ μεταλλεύματος.

Βιβλιογραφία

i)Εισαγωγή στην Γεωστατιστική, Καπαγερίδης, Ιωάννης, Εκδόσεις Ιων, Αθήνα, 2006

ii) Εισαγωγή στη Γεωστατιστική, Διονύσιος Θ. Χριστόπουλος

iii) Εισαγωγή στη Γεωστατιστική, Μαμάσης (Παρουσίαση)

iv) Guidal, 1987. Recoverable reserves estimation at an Australian gold project, in: Matheron, G., and Armstrong, M., (Eds.). Geostatistical case studies (σελ 149 – 168)

v) Practical Geostatistics (1979) by Isobel Clark

vi) Kriging spreadsheet developed by O. Dubrule and P. Delfiner

vii) Webster, R. και Oliver, M.A. 1990. Statistical methods in soil and land resource survey, Oxford Press (New York), σελ 316.

viii) Αναλυση Χώρου: Θεωρία, Μεθοδολογία και Μεθόδοι, Κωστης Κουτσοπουλος , Εκδόσεις Παπασωτηρίου, Αθήνα, 2002

Παράρτημα

Υπολογισμός των τιμών *γ(h)*

Διεύθυνση=1

h=1α

Σημεία	Απόσταση(α)	Ζεύγη	Περιεκτικότητα	Z((x(i)+h)-Z(x(i))	Z((x(i)+h) -Z(x(i))^2
1	1		35		
2	2	1	35	0	0
3	3	2	33	-2	4
4	4	3	33	0	0
5	5	4	34	1	1
6	6	5	31	-3	9
7	7	6	35	4	16
8	8	7	37	2	4
9	9	8	41	4	16
10	10	9	41	0	0
11	1		35		
12	2	10	35	0	0
13	3	11	35	0	0
14	6		35		
16	7	12	33	-2	4

17	10		41		
18	1		37		
19	2	13	35	-2	4
	3	14	37	2	4
	4	15	35	-2	4
	6		37		
	7	16	37	0	0
	8	17	39	2	4
	9	18	39	0	0
	10	19	41	2	4
	1		37		
	2	20	40	3	9
	3	21	42	2	4
	5		34		
	6	22	36	2	4
	7		41	5	25
	10		34		
	8		42		
	9	24	33	-9	81
				Αθροισμα	197
				γ(h)	4,10

Διεύθυνση = 1 h=2α

Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	35	33	2	4
2	33	31	2	4
3	31	37	-6	36
4	37	41	-4	16
5	35	35	0	0
6	37	37	0	0
7	37	39	-2	4
8	39	41	-2	4
9	37	42	-5	25
10	34	41	-7	49
11	39	31	8	64
12	35	33	2	4
13	33	31	2	4
14	31	37	-6	36
15	37	41	-4	16
16	35	35	0	0
17	37	39	-2	4
			άθροισμα	270
			γ(2a)	7,94

Διεύθυνση = 1

h=3α

Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	35	33	2	4
2	33	35	-2	4
3	35	41	-6	36
4	37	35	2	4
5	37	39	-2	4
6	41	34	7	49
7	41	33	8	64
8	35	34	1	1
9	34	37	-3	9
10	37	41	-4	16
11	40	34	6	36
12	33	31	2	4
13	31	41	-10	100
			άθροισμα	331
			γ(3a)	12,73

Διεύθυνση =

2

h=αT_P(2)

Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	35	35	0	0
2	37	35	2	4
3	35	33	2	4
4	37	35	2	4
5	35	35	0	0
6	35	33	2	4
7	40	37	3	9
8	42	35	7	49
9	35	35	0	0
10	34	37	-3	9
11	37	33	4	16
12	33	37	-4	16
13	36	37	-1	1
14	33	41	-8	64
15	41	39	2	4
16	39	41	-2	4
17	39	33	6	36
			άθροισμα	224
			γ(αT_P(2))	6,59

Διεύθυνση = 2

	h=2αT_P(2)			
Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	37	33	4	16
2	37	35	2	4
3	37	34	3	9
4	35	31	4	16
5	34	33	1	1
6	37	41	-4	16
7	33	39	-6	36
8	39	41	-2	4
9	42	34	8	64
10	30	42	-12	144
11	35	33	2	4
12	37	37	0	0
			άθροισμα	314
			γ(2αT_P(2))	13,08

Διεύθυνση

= 2

h=3αT_P(2)

Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	37	33	4	16
2	40	34	6	36
3	42	31	11	121
4	41	35	6	36
5	34	37	-3	9
6	36	41	-5	25
7	41	41	0	0
			άθροισμα	243
			γ(2αT_P(2))	17,36

Διεύθυνση=3 h=1α

Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	35	35	0	0
2	35	37	-2	4
3	37	37	0	0
4	35	35	0	0
5	35	40	-5	25
6	33	35	-2	4
7	35	37	-2	4

8	37	42	-5	25
9	42	41	1	1
10	31	35	-4	16
11	37	36	1	1
12	36	33	3	9
13	35	33	2	4
14	33	37	-4	16
15	37	41	-4	16
16	42	39	3	9
17	41	41	0	0
18	41	41	0	0
19	41	34	7	49
			αθροισμα	183
			γ(1a)	4,8

Διεύθυνση=3	h=2α			
Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	35	37	-2	4
2	35	35	0	0
3	33	37	-4	16
4	33	35	-2	4
5	31	37	-6	36

6	37	33	4	16
7	35	37	-2	4
8	37	39	-2	4
9	41	41	0	0
10	35	37	-2	4
11	35	40	-5	25
12	35	42	-7	49
13	35	36	-1	1
14	33	41	-8	64
15	41	34	7	49
			αθροισμα	276
			γ(2a)	9,2

Διεύθυνση=3	h=3α			
Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	35	37	-2	4
2	35	40	-5	25
3	33	42	-9	81
4	35	35	0	0
5	31	36	-5	25
6	33	30	3	9
7	39	42	-3	9

8	39	42	-3	9
9	41	34	7	49
10	34	31	3	9
11	35	41	-6	36
12	35	33	2	4
			αθροισμα	260
			γ(3a)	10,8

Διεύθυνση = 4

h=1αT_P(2)

Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	37	40	-3	9
2	40	41	-1	1
3	41	35	6	36
4	35	35	0	0
5	35	42	-7	49
6	35	35	0	0
7	35	37	-2	4
8	35	35	0	0
9	35	35	0	0
10	35	34	1	1
11	34	33	1	1

12	37	41	-4	16
13	33	31	2	4
14	34	35	-1	1
15	35	37	-2	4
16	31	33	-2	4
17	33	39	-6	36
18	39	34	5	25
19	41	41	0	0
			άθροισμα	191
			γ(2αT_P(2))	5,03

Διεύθυνση = 4

	h=2αT_P(2)			
Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	37	41	-4	16
2	35	30	5	25
3	35	42	-7	49
4	35	37	-2	4
5	35	35	0	0
6	35	33	2	4
7	33	39	-6	36

8	36	42	-6	36
9	33	37	-4	16
10	41	33	8	64
11	34	37	-3	9
12	31	39	-8	64
13	33	31	2	4
14	35	39	-4	16
15	37	41	-4	16
16	40	35	5	25
17	35	34	1	1
			άθροισμα	385
			γ(2αT_P(2))	11,32

Διεύθυνση = 4

	h=3αT_P(2)			
Ζεύγη			Z((x(i)+h)-Z(x(i))	Z((x(i)+h)-Z(x(i))^2
1	37	35	2	4
2	35	34	1	1
3	34	39	-5	25
4	33	36	-3	9
5	33	41	-8	64

6	37	33	4	16
7	41	31	10	100
8	35	34	1	1
9	41	40	1	1
			άθροισμα	221
			γ(2αT_P(2))	12,28

Σημειώσεις Τέλους (**βιβλιογραφία)

[a] Αναλυση Χώρου: Θεωρία, Μεθοδολογία και Μεθόδοι, Κωστης Κουτσοπουλος , Εκδόσεις Παπασωτηρίου, Αθήνα, 2002 σελ 41-42

[b] Το ίδιο με προηγούμενο σελ 71-73

[c] Εισαγωγή στην Γεωστατιστική, Καπαγερίδης, Ιωάννης, Εκδόσεις Ιων, Αθήνα, 2006 σελ 35-39

[d] Το ίδιο με προηγούμενο σελ 39-45 και 47-49

[e] Αναλυση Χώρου: Θεωρία, Μεθοδολογία και Μεθόδοι, Κωστης Κουτσοπουλος, Εκδόσεις Παπασωτηρίου, Αθήνα, 2002 σελ 207

[f] Εισαγωγή στην Γεωστατιστική, Καπαγερίδης, Ιωάννης, Εκδόσεις Ιων, Αθήνα, 2006 σελ 53-60

[g] Αναλυση Χώρου: Θεωρία, Μεθοδολογία και Μεθόδοι, Κωστης Κουτσοπουλος, Εκδόσεις Παπασωτηρίου, Αθήνα, 2002 σελ 159-173

[h] Εισαγωγή στην Γεωστατιστική, Καπαγερίδης, Ιωάννης, Εκδόσεις Ιων, Αθήνα, 2006 σελ 25-29

[i] Το ίδιο με προηγούμενο σελ 30

[j] Αναλυση Χώρου: Θεωρία, Μεθοδολογία και Μεθόδοι, Κωστης Κουτσοπουλος, Εκδόσεις Παπασωτηρίου, Αθήνα, 2002 σελ 258-272

[k] Εισαγωγή στην Γεωστατιστική, Καπαγερίδης, Ιωάννης, Εκδόσεις Ιων, Αθήνα, 2006 σελ 61-73

[l] Το ίδιο με προηγούμενο σελ 79-84

[m] Το ίδιο με προηγούμενο σελ 87-92

[n] Το ίδιο με προηγούμενο σελ 95-97

[o] Το ίδιο με προηγούμενο σελ 114

[p] Muller, W.G., Zimmerman, D.L., 1995. in: Kitsos, C.P., M uller, W.G. (Eds.), An algorithm for sampling optimization for semi-variogram estimation. Model-Oriented Data Analysis 4. Physica-Verlag, Heidelberg, pp. 173{178.

[q] Εισαγωγή στην Γεωστατιστική, Καπαγερίδης, Ιωάννης, Εκδόσεις Ιων, Αθήνα, 2006 σελ 153

[r] Το ίδιο με προηγούμενο σελ 167