

# ΣΤΑΤΙΣΤΙΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΓΙΑ ΤΗΝ ΑΠΟΤΙΜΗΣΗ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

- Στατιστική ανάλυση του γεωχημικού δείγματος → μας δίνει πληροφορίες για τον γεωχημικό πληθυσμό που μελετάμε.
- Συνυπολογισμός σφαλμάτων
- Πειραματικά δεδομένα → οργανωμένα με τρόπο που επιτρέπει την εύκολη εισαγωγή και διαχείρισή τους στον υπολογιστή.
- Στατιστική εκτίμηση με υπολογισμούς παραμέτρων και εξέταση υποθέσεων.

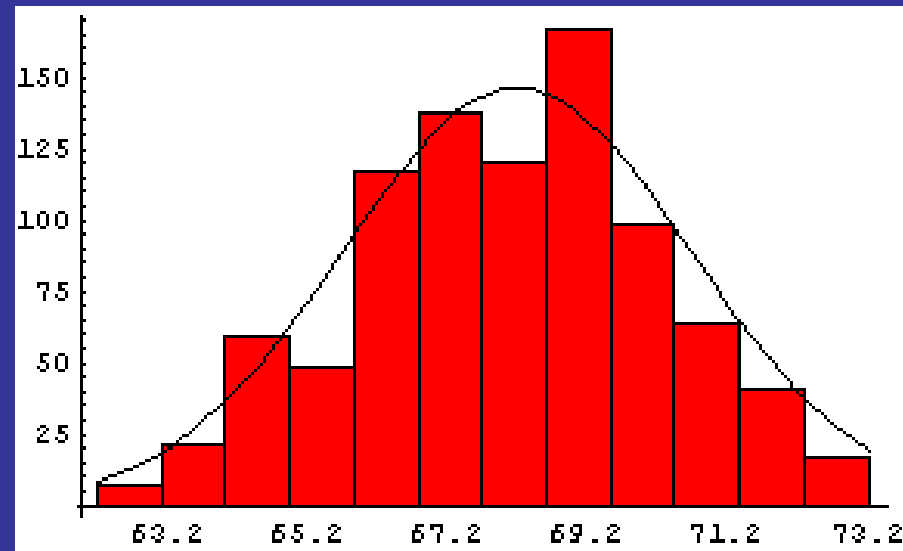
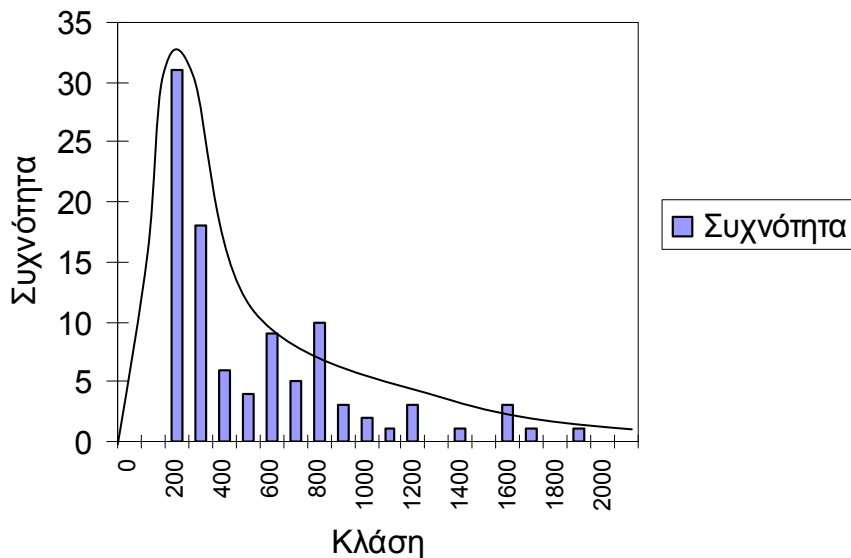
# ΒΗΜΑΤΑ ΒΑΣΙΚΗΣ ΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

1. Απεικόνιση και μελέτη της στατιστικής κατανομής των δεδομένων.
2. Υπολογισμός περιγραφικών στατιστικών παραμέτρων.
3. Αναγνώριση απόμακρων τιμών.
4. Προσδιορισμός τιμών «γεωχημικού πλαισίου».
5. Αναγνώριση «γεωχημικής ανωμαλίας» που πιθανά σχετίζεται με ρύπανση.

# ΣΤΑΤΙΣΤΙΚΗ ΚΑΤΑΝΟΜΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

- Έλεγχος με γραφικές μεθόδους (π.χ. ιστόγραμμα) ή με εφαρμογή στατιστικών δοκιμών.
- Ο υπολογισμός πολλών στατιστικών παραμέτρων προϋποθέτει κανονική κατανομή δεδομένων → πιθανή αναγκαιότητα η μετατροπή (π.χ. με λογαρίθμηση) για κανονικοποίηση.

Ιστόγραμμα Zn ppm



# ΠΕΡΙΓΡΑΦΙΚΑ ΣΤΑΤΙΣΤΙΚΑ – (Α) ΕΓΓΥΤΗΤΑ ΣΤΗΝ ΚΕΝΤΡΙΚΗ ΤΙΜΗ

ο **αριθμητικός μέσος** ή μέση τιμή (arithmetic mean or average)  $n$  μετρήσεων:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

**διάμεσος** (median), δηλαδή η τιμή για την οποία το ήμισυ των δειγμάτων στην κατανομή έχει τιμές μικρότερες και το ήμισυ μεγαλύτερες. Η διάμεση τιμή αποτελεί πιο σταθερή παράμετρο από τη μέση τιμή σε περιπτώσεις λίγων μετρήσεων ( $n < 5$ ).

**τιμή μέγιστης συχνότητας** (mode). Τα δεδομένα υποδιαιρούνται σε ειδικά διαστήματα (κλάσεις) και λαμβάνεται το κέντρο της κλάσης με την υψηλότερη συχνότητα. Πολλές γεωχημικές κατανομές έχουν περισσότερες από μία τιμή μέγιστης συχνότητας.

Ο **γεωμετρικός μέσος** (geometric mean). Η μέση τιμή του δεκαδικού λογαρίθμου των δεδομένων.

# ΠΕΡΙΓΡΑΦΙΚΑ ΣΤΑΤΙΣΤΙΚΑ – (B) ΔΙΑΣΠΟΡΑ

**Περιοχή ή εύρος τιμών** (range). Η διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής των δεδομένων.

**Διακύμανση** (variance). Το τετράγωνο της μέσης διαφοράς μεταξύ τιμών δεδομένων και της μέσης τιμής τους:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

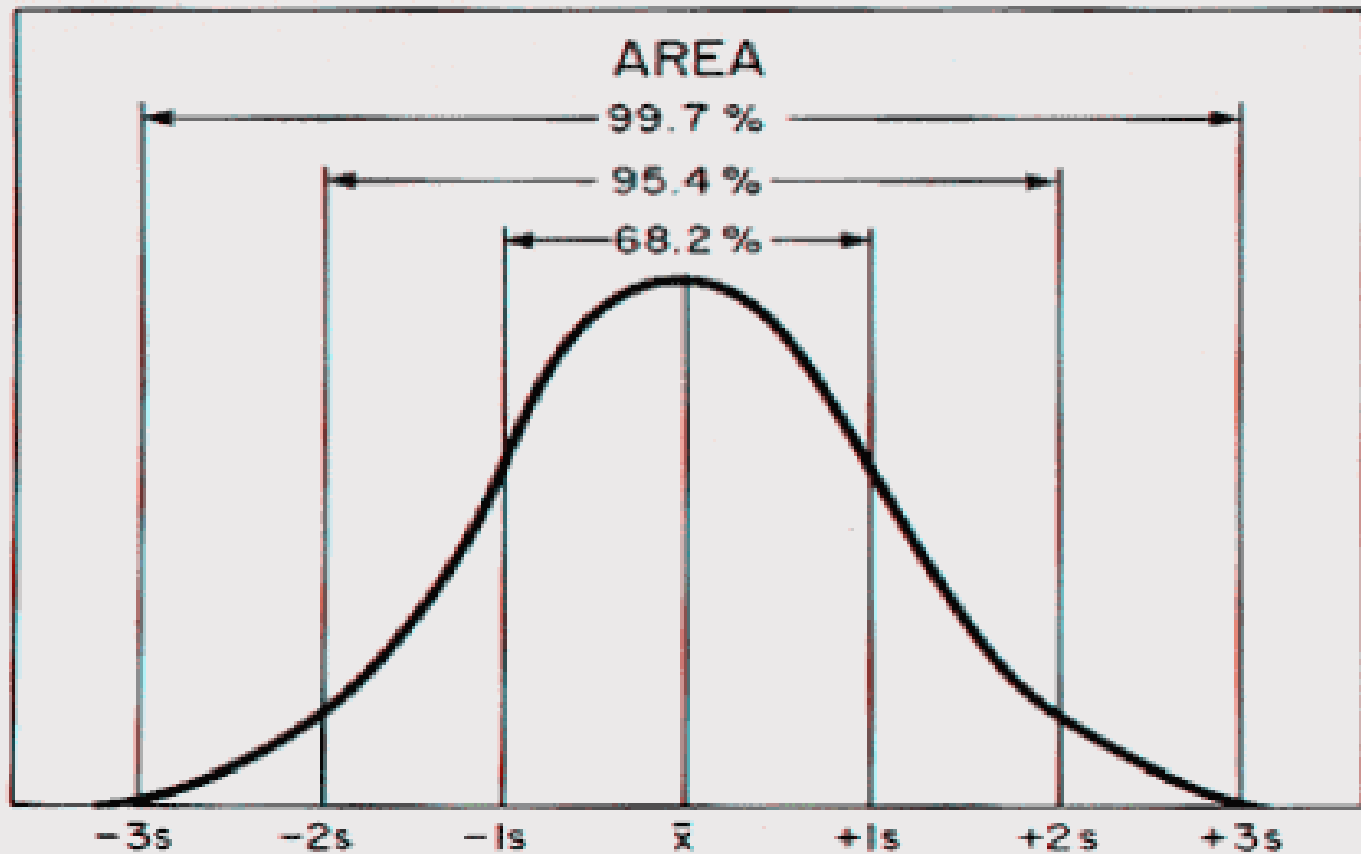
**Τυπική απόκλιση** (standard deviation). Η τετραγωνική ρίζα της διακύμανσης. Είναι η συνηθέστερη παράμετρος έκφρασης της διασποράς και σε αντιδιαστολή με τη διακύμανση εκφράζεται με τις μονάδες μέτρησης της μέσης τιμής.

**Σχετική τυπική απόκλιση** (coefficient of variation). Ο λόγος:  
τυπική απόκλιση / μέση τιμή

# ΚΑΜΠΥΛΗ ΚΑΝΟΝΙΚΗΣ ΚΑΤΑΝΟΜΗΣ

## Gaussian Curve

ΣΥΧΝΟΤΗΤΑ



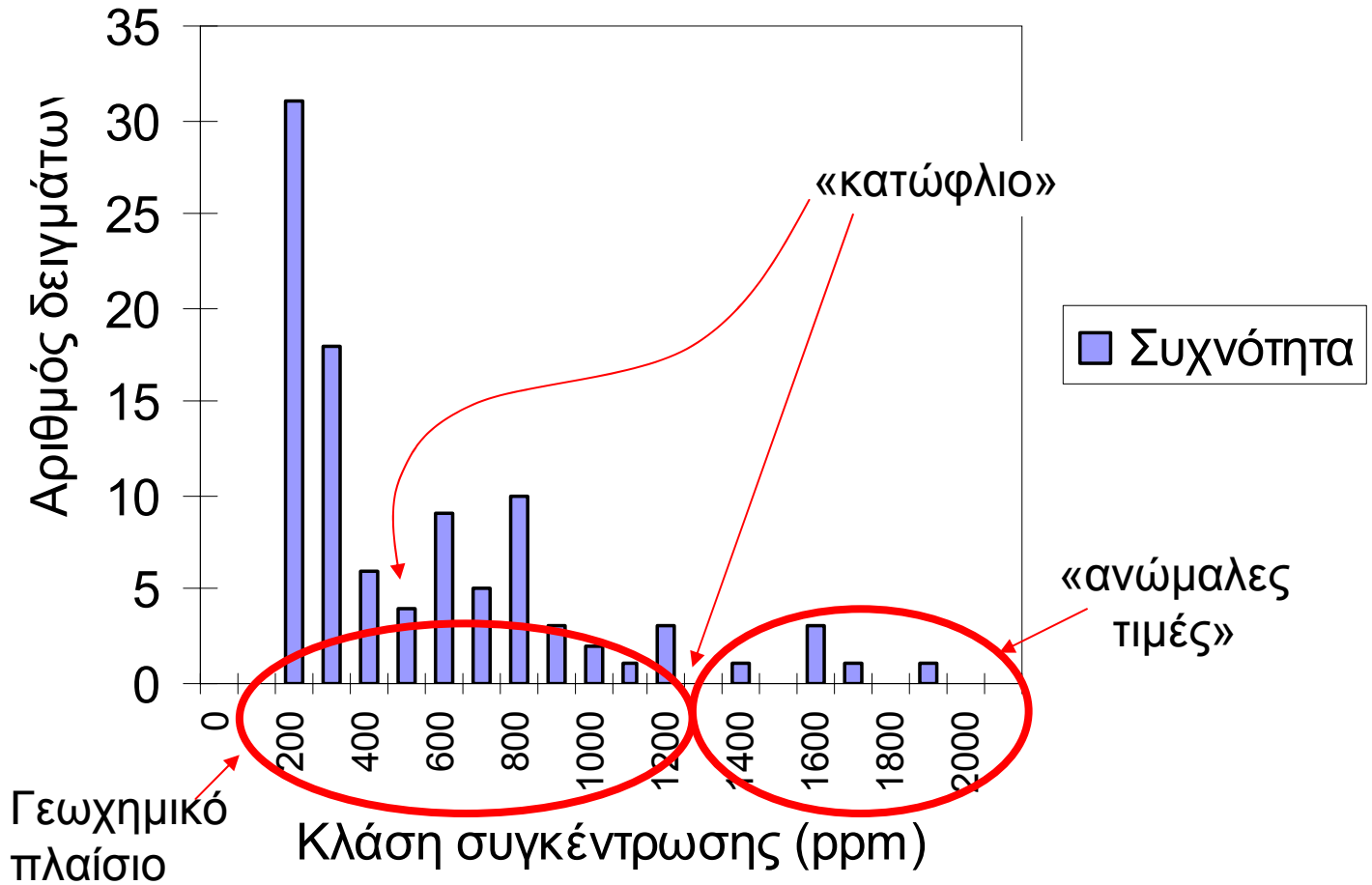
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ

# ΚΑΤΑΣΚΕΥΗ ΙΣΤΟΓΡΑΜΜΑΤΟΣ

1. Έστω ότι έχουμε  $n$  μετρήσεις μιας γεωχημικής παραμέτρου.
2. Επιλέγουμε τον αριθμό των κλάσεων ταξινόμησης της παραμέτρου. Κατά κανόνα 7-11 κλάσεις καλύπτουν το συνολικό εύρος τιμών των μετρήσεων. Οι κλάσεις προβάλλονται στον άξονα  $x$  του γραφήματος.
3. Υπολογίζουμε τον αριθμό των μετρήσεων κάθε κλάσης.
4. Διαιρούμε τον αριθμό μετρήσεων κάθε κλάσης με το  $n$  ώστε να υπολογίσουμε το ποσοστό μετρήσεων κάθε κλάσης. Τα ποσοστά προβάλλονται στον άξονα  $y$  του γραφήματος.
5. Για κάθε κλάση σχεδιάζουμε μια στήλη με ύψος όπως υπολογίστηκε στο προηγούμενο βήμα.

# ΑΠΕΙΚΟΝΙΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΣΕ ΓΡΑΦΗΜΑΤΑ

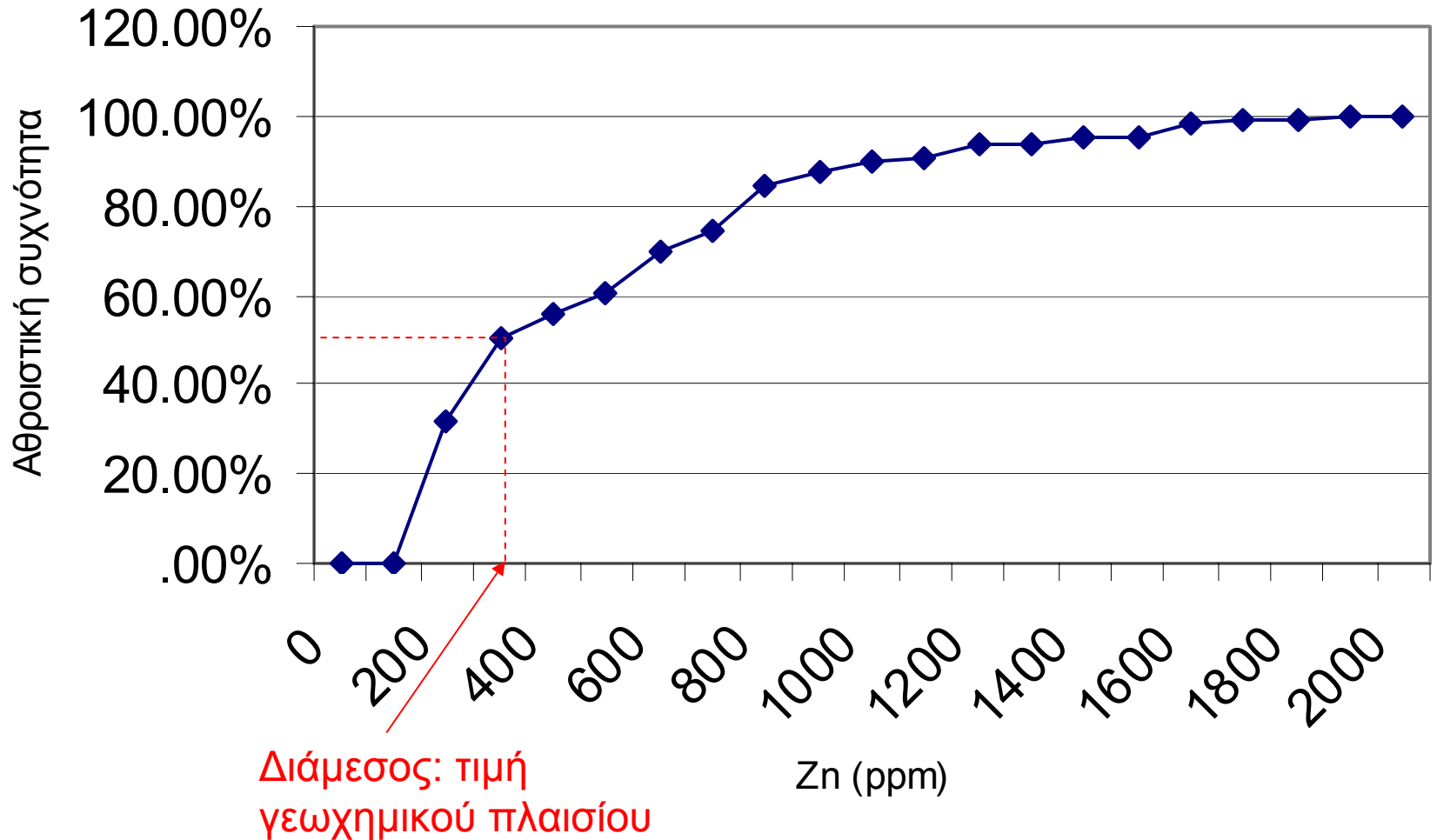
## Ιστόγραμμα Zn ppm



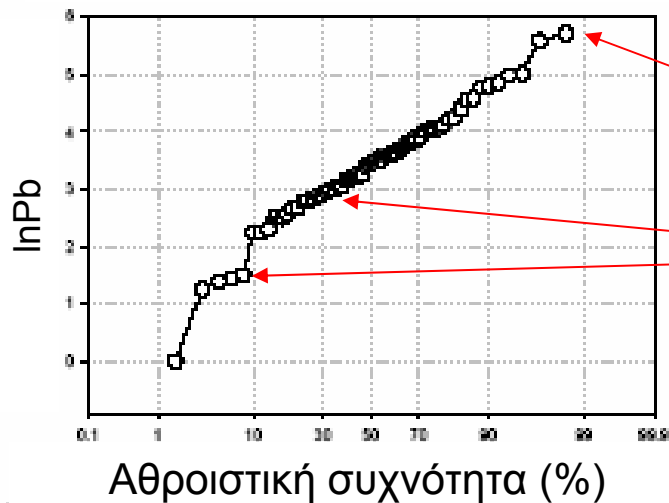
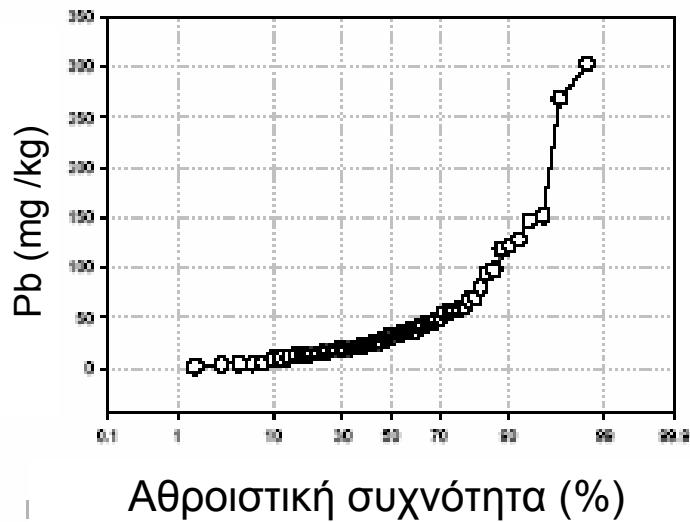


# ΓΡΑΦΗΜΑ ΑΘΡΟΙΣΤΙΚΗΣ ΣΥΧΝΟΤΗΤΑΣ

Αθροιστική συχνότητα Zn



# ΓΡΑΦΗΜΑ ΠΙΘΑΝΟΤΗΤΩΝ

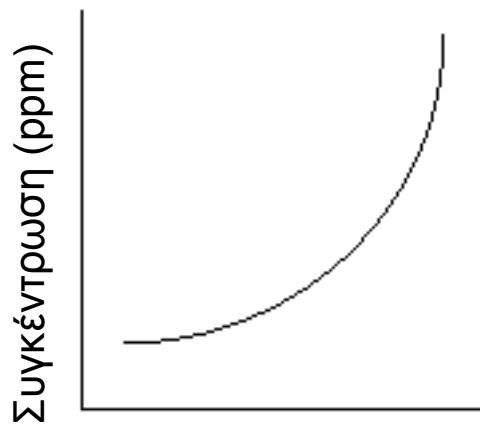


Προβολή αθροιστικής  
συχνότητας σε χαρτί  
αριθμητικών πιθανοτήτων.

Επιτρέπει την αναγνώριση :

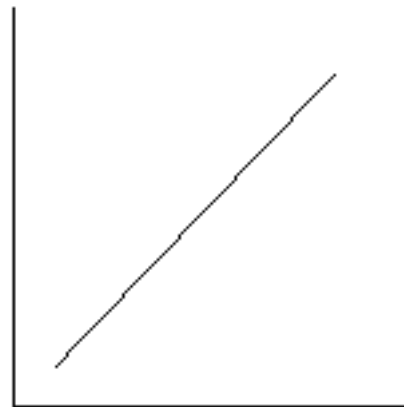
1. είδους κατανομής των δεδομένων (π.χ. κανονική, λογαριθμική).
2. Απόμακρων τιμών.
3. Πολλαπλών πληθυσμών στα δεδομένα.

# ΓΡΑΦΗΜΑ ΠΙΘΑΝΟΤΗΤΩΝ – ΕΙΔΟΣ ΚΑΤΑΝΟΜΗΣ

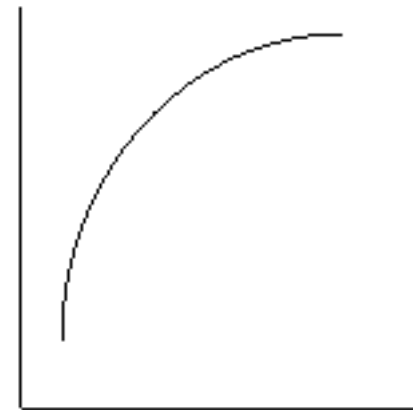


Αθροιστική  
συχνότητα (%)

Λοξή κατανομή  
με «ουρά» προς  
τα δεξιά

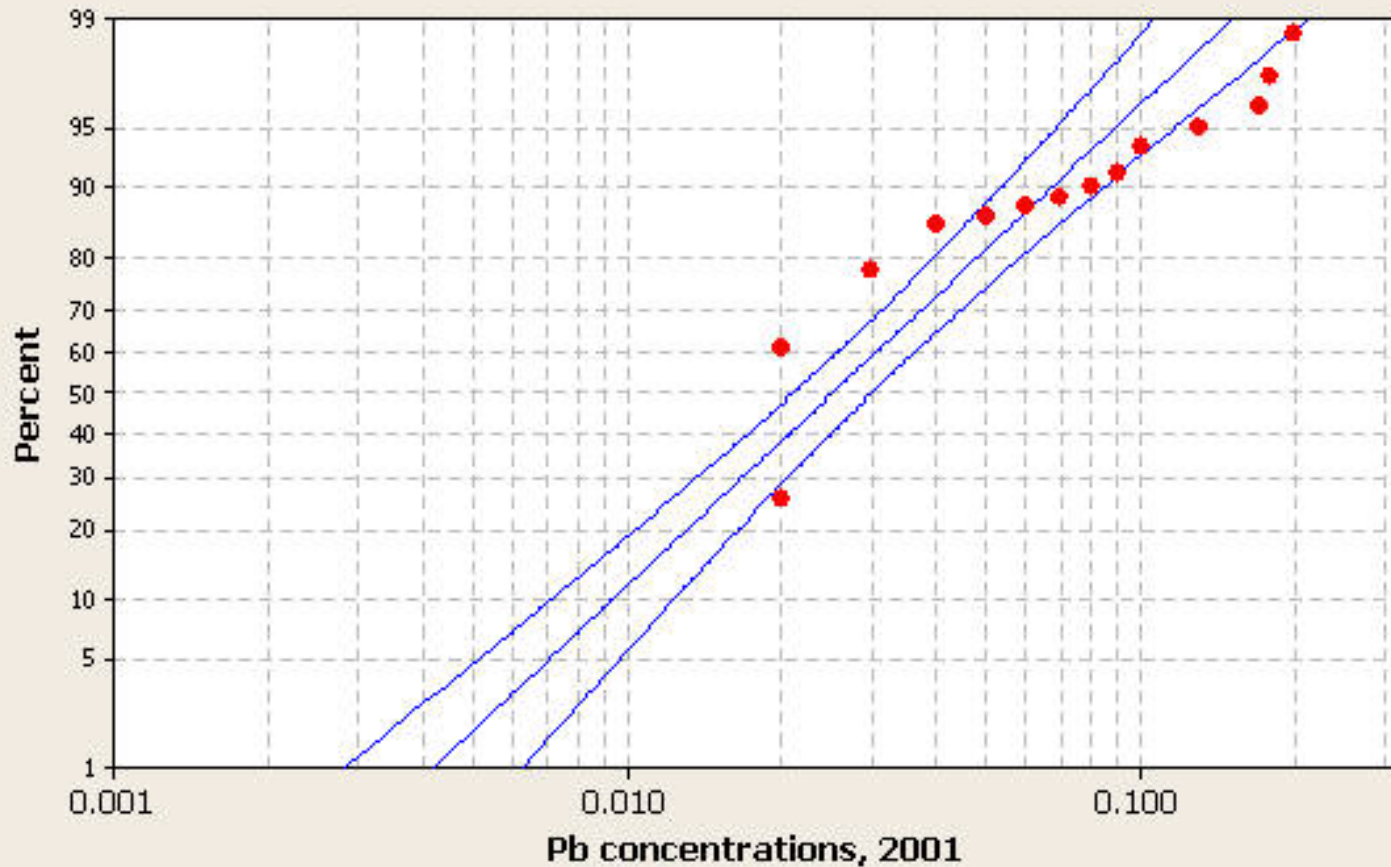


Κανονική  
κατανομή



Λοξή κατανομή  
με «ουρά» προς  
τα αριστερά

Percentiles of Lead Concentrations in Soils, 2001



Πώς κρίνετε την εφαρμογή του μοντέλου πιθανοτήτων του γραφήματος στα συγκεκριμένα δεδομένα συγκεντρώσεων μολύβδου (κόκκινοι κύκλοι στο γράφημα);

# ΣΥΝΤΕΛΕΣΤΗΣ ΣΥΣΧΕΤΙΣΗΣ

Συντελεστής συσχέτισης (correlation coefficient)  $r$ .

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Ο συντελεστής αυτός παίρνει τιμές από 1 έως -1 με την τιμή 1 να σημαίνει τέλεια θετική συσχέτιση, -1 τέλεια αρνητική συσχέτιση και 0 καμία συσχέτιση. Γενικά τιμές  $r >$  από 0.5 σημαίνουν καλή στατιστική συσχέτιση μεταξύ παραμέτρων.

Υψηλοί συντελεστές συσχέτισης στοιχείων σε έρευνες περιβαλλοντικής γεωχημείας μπορεί να σημαίνουν κοινές πηγές ρύπανσης.

Υψηλοί συντελεστές συσχέτισης σε μη ρυπασμένες περιοχές μπορεί να σημαίνουν κοινή πρωτογενή πηγή προέλευσης των στοιχείων (π.χ. Cr-Ni σε υπερβασικά πετρώματα).

	<i>Pb</i>	<i>Zn</i>	<i>Cu</i>	<i>As</i>	<i>Cd</i>	<i>Mn</i>	<i>Fe</i>	<i>Ni</i>	<i>Co</i>
<i>Zn</i>	0.77 0.000								
<i>Cu</i>	0.33 0.019	0.50 0.000							
<i>As</i>	0.87 0.000	0.77 0.000	0.38 0.000						
<i>Cd</i>	0.66 0.000	0.86 0.000	0.54 0.000	0.71 0.000					
<i>Mn</i>	0.51 0.000	0.79 0.000	0.32 0.023	0.58 0.000	0.75 0.000				
<i>Fe</i>	0.08 0.592	0.23 0.118	0.22 0.135	0.21 0.142	0.29 0.044	0.36 0.010			
<i>Ni</i>	-0.10 0.480	-0.16 0.279	-0.28 0.055	-0.14 0.355	-0.04 0.781	-0.04 0.799	0.22 0.123		
<i>Co</i>	-0.33 0.020	-0.22 0.122	-0.23 0.112	-0.29 0.045	-0.10 0.538	-0.10 0.481	0.46 0.001	0.80 0.000	
<i>Cr</i>	-0.11 0.471	-0.22 0.137	-0.26 0.072	-0.15 0.308	-0.10 0.591	-0.14 0.339	0.24 0.096	0.97 0.000	0.75 0.000

**ΠΙΝΑΚΑΣ ΣΥΣΧΕΤΙΣΗΣ**

# ΓΡΑΦΗΜΑ ΣΥΣΧΕΤΙΣΗΣ

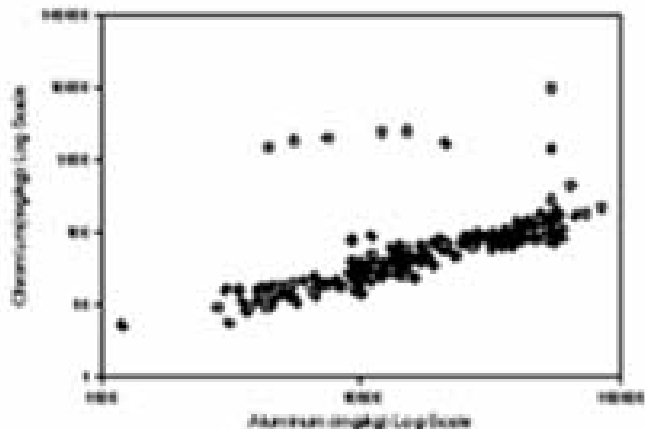


FIGURE 3-4. Scatter plot of aluminum vs. chromium concentrations (log scale) showing unclear association

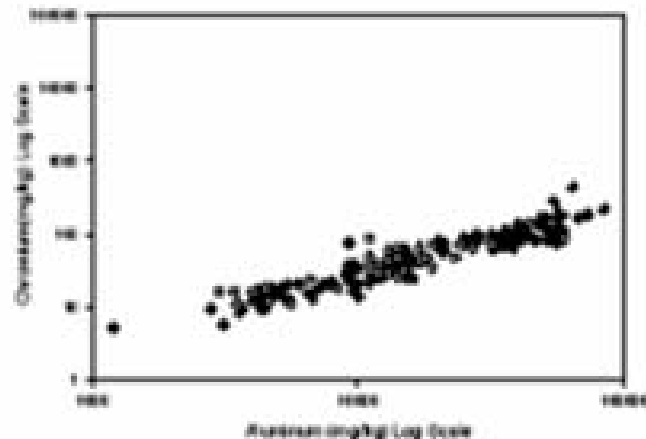


FIGURE 3-5. Scatter plot of aluminum vs. chromium concentrations (log scale) with outliers removed

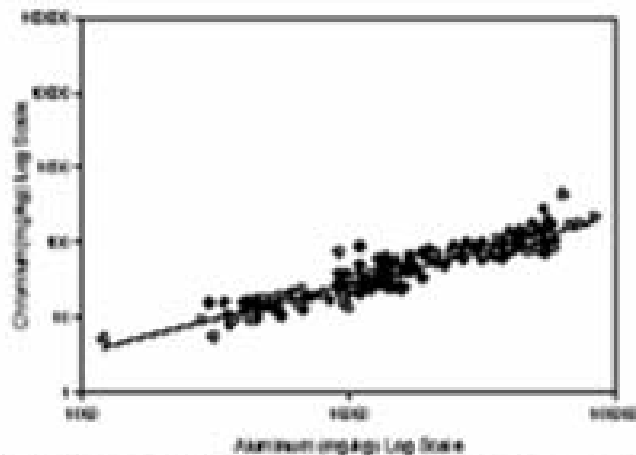


FIGURE 3-6. Scatter plot of aluminum vs. chromium concentrations showing least-squares regression line

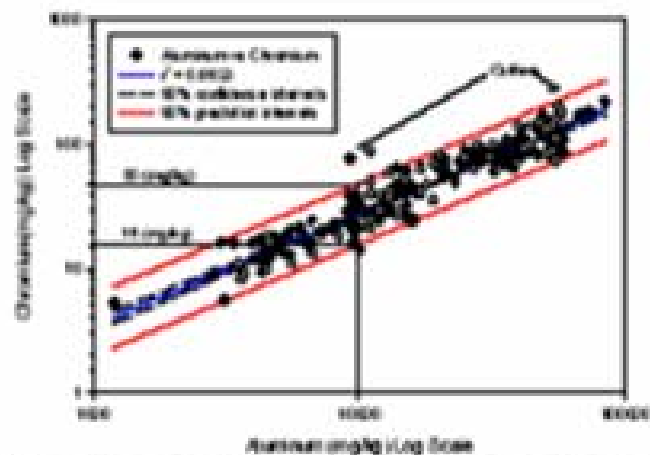


FIGURE 3-7. Geochemical regression: aluminum vs. chromium concentrations (log scale)

# ΕΡΜΗΝΕΙΑ ΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

Ερμηνεία μιας απλής στατιστικής ανάλυσης γεωχημικών δεδομένων → αναγνώριση σημαντικών γεωχημικών ανωμαλιών.

Τιμές πλαισίου και κατωφλίου καθορίζονται από το ιστόγραμμα ή το διάγραμμα αθροιστικής συχνότητας της κατανομής των δεδομένων μας.

Σύγκριση των στατιστικών παραμέτρων με θεσμοθετημένα όρια περιεκτικότητας χημικών στοιχείων.

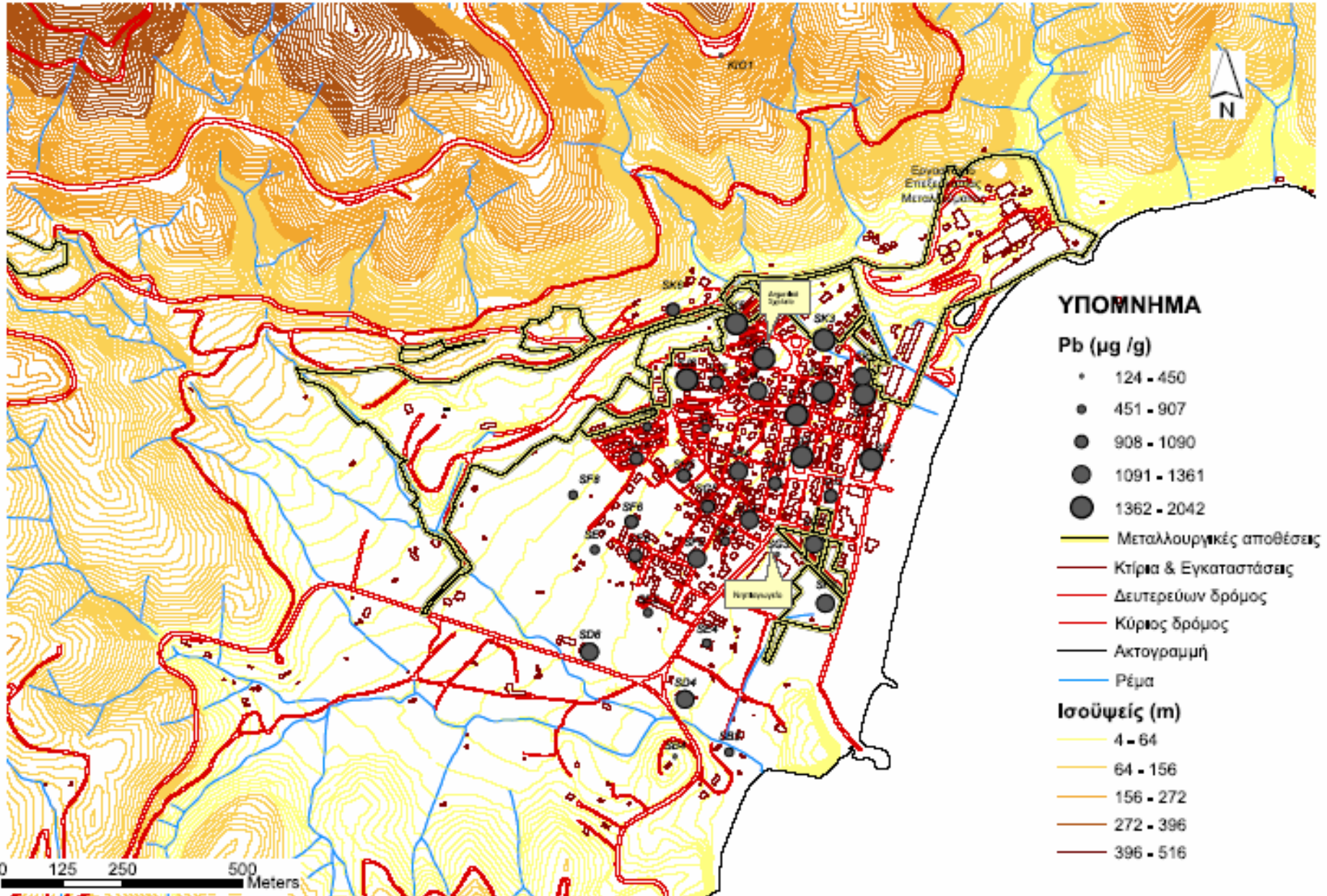
Χρήσιμο να εξετάσουμε και τους συσχετισμούς μεταξύ των προσδιοριζόμενων παραμέτρων.

Τα αποτελέσματα γεωχημικών διασκοπίσεων απεικονίζονται συνήθως σε γεωχημικούς χάρτες συγκέντρωσης στους οποίους γραφικά σύμβολα δείχνουν την προσδιοριζόμενη συγκέντρωση αναλυτή σε κάθε θέση δειγματοληψίας.



# ΑΠΕΙΚΟΝΗΣΗ ΔΕΔΟΜΕΝΩΝ ΣΕ ΧΑΡΤΗ

Pb στο έδαφος (n = 38)



# ΠΕΡΙΕΧΟΜΕΝΑ ΓΕΩΧΗΜΙΚΩΝ ΧΑΡΤΩΝ

- ✓ Χημικό στοιχείο που αναλύθηκε και μονάδες μέτρησης
- ✓ Αναλυτική μέθοδος
- ✓ Τύπος δείγματος (βάθος ορίζοντα, μέγεθος κλάσματος)
- ✓ Θέσεις δειγμάτων
- ✓ Αναλυτικά αποτελέσματα
- ✓ Τοπογραφία
- ✓ Υδρογραφικό δίκτυο
- ✓ Πιθανές πηγές ρύπανσης
- ✓ Κλίμακα
- ✓ Ένδειξη Βορρά
- ✓ Τοποθεσία διασκόπησης
- ✓ Ημερομηνία
- ✓ Ονόματα δειγματοληπτών και του υπεύθυνου του έργου.

# ΓΕΩΧΗΜΙΚΕΣ ΕΚΘΕΣΕΙΣ

- Τίτλος
- Περίληψη
- Πίνακας Περιεχομένων
- Εισαγωγή
- Περιγραφή της περιοχής μελέτης
- Μεθοδολογία υπαίθριας δειγματοληψίας, προετοιμασίας δειγμάτων και χημικής ανάλυσης
- Μεθοδολογία επεξεργασίας των αποτελεσμάτων
- Ποιοτικός έλεγχος αποτελεσμάτων
- Περιγραφή των αποτελεσμάτων
- Συζήτηση των αποτελεσμάτων
- Συμπεράσματα
- Προτάσεις
- Βιβλιογραφία
- Παραρτήματα