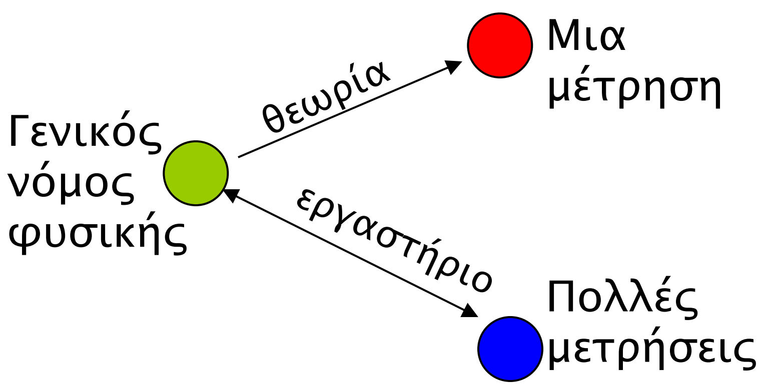
Ανάλυση δεδομένων

Σχέση μεταξύ μετρήσεων και θεωρίας:



Μια και μόνο μέτρηση, για παράδειγμα η θέση του βλήματος που κάνει πλάγια βολή μια χρονική στιγμή δεν είναι αρκετή για να περιγράψει το γενικό φαινόμενο

Για να συνδέσουμε το πείραμα με θεωρία θα πρέπει να έχουμε πολλές μετρήσεις και από το τρόπο κατανομής των δεδομένων να ανακαλύψουμε την θεωρία που κρύβεται πίσω από τα δεδομένα

Στο παράδειγμα του βλήματος, η μελέτη του βεληνεκούς για διάφορες γωνίες ρίψης και διαφορετικές ταχύτητες μπορούν να βοηθήσουν να κατανοήσουμε το φυσικό νόμο που περιγράφει το φαινόμενο αυτό

Όταν πραγματοποιούμε κάποιο πείραμα δεν ενδιαφερόμαστε απλά και μόνο για τις τιμές κάποιων μεγεθών που μετράμε αλλά και για την συσχέτιση που υπάρχει μεταξύ των μεγεθών αυτών

Η συσχέτιση μεταξύ των μεγεθών είναι αυτή που εκδηλώνει την ύπαρξη κάποιου φυσικού νόμου

Το βασικό εργαλείο που χρησιμοποιούμε για να βρούμε κάποιο συσχετισμό είναι οι γραφικές παραστάσεις. Το ερώτημα που γεννάται όμως είναι πως μπορούμε να μειώσουμε το μεγάλο αριθμό μετρήσεων σε ποσότητες που μπορούν να συγκριθούν με τις θεωρητικές προβλέψεις

Μετάδοση σφαλμάτων – Στατιστικά σφάλματα

Χρησιμοποιώντας τους κανόνες για πράξεις μεταξύ μετρούμενων τιμών με διαφορετικό αριθμό σημαντικών ψηφίων μπορούμε να υπολογήσουμε την αβεβαιότητα ενός αποτελέσματος.

Υπάρχουν ωστόσο καλύτεροι και περισσότεροι ακριβείς τρόποι

Θεωρήστε ότι υπολογίζετε μια ποσότητα Q η οποία εξαρτάται από 3 μετρούμενα μεγέθη x, y, z. Η Q είναι επομένως συνάρτηση 3 μεταβλητών

$$Q = f(x,y,z)$$

Κάθε μια από τα μετρούμενα μεγέθη συνοδεύονται με κάποια αβεβαιότητα $x \pm \Delta x$ $(y \pm \Delta y)$ $(z \pm \Delta z)$. Η αβεβαιότητα στο αποτέλεσμα Q δίνεται από

$$\Delta Q = \sqrt{\left(\left(\frac{\partial f}{\partial x}\Delta x\right)^2 + \left(\frac{\partial f}{\partial y}\Delta y\right)^2 + \left(\frac{\partial f}{\partial z}\Delta z\right)^2\right)}$$

Η έκφραση $\dfrac{\partial f}{\partial x}$ παριστάνει τη μερική παράγωγο της f ως προς τη μεταβλητή x

Είναι εύκολο να υπολογισθεί σαν η παράγωγος της συνάρτησης ως προς την εκάστοτε μεταβλητή, κρατώντας τα y και z σταθερά (άρα η παράγωγός τους ως προς x είναι 0).

Προσοχή: Η παραπάνω σχέση ισχύει μόνο για ανεξάρτητες μεταβλητές οι αβεβαιότητες των οποίων είναι τυχαίες (στατιστικές).
Η κατανομή των μεταβλητών περιγράφεται από τη Gaussian κατανομή

Παράδειγμα υπολογισμού σφάλματος

Έστω ότι η συναρτησιακή εξάρτηση της Q από τα x, y και z δίνεται από τη σχέση

$$Q = f(x,y,z) = k\frac{xy}{z^2}$$ Όπου k σταθερά με k = 0.3872

Υποθέτουμε ακόμα ότι

$$x = 21.7 \pm 0.2$$
$$y = 83.9 \pm 0.3$$
$$z = 2.51 \pm 0.04$$

Υπολογίζουμε πρώτα τη τιμή της Q : $Q = k\frac{xy}{z^2} = 0.3872\frac{21.2 \times 83.9}{2.51^2} = 111.8947217$

(Από σημαντικά ψηφία θα έπρεπε να γράψουμε ότι Q = 112±5)

Θα κάνουμε στρογγυλοποίηση αφού πρώτα υπολογήσουμε την αβεβαιότητα

Βρίσκουμε τις μερικές παραγώγους:

$$\frac{\partial Q}{\partial x} = k\frac{y}{z^2} \qquad \frac{\partial Q}{\partial y} = k\frac{x}{z^2} \qquad \frac{\partial Q}{\partial z} = -2k\frac{xy}{z^3}$$

Και χρησιμοποιώντας την εξίσωση της αβεβαιότητας έχουμε:

$$\Delta Q = \sqrt{\left(\left(\frac{\partial f}{\partial x}\Delta x\right)^2 + \left(\frac{\partial f}{\partial y}\Delta y\right)^2 + \left(\frac{\partial f}{\partial z}\Delta z\right)^2\right)} \Rightarrow \Delta Q = \sqrt{\left(\left(k\frac{y}{z^2}\Delta x\right)^2 + \left(k\frac{x}{z^2}\Delta y\right)^2 + \left(-2k\frac{xy}{z^3}\Delta z\right)^2\right)}$$

$$\Rightarrow \Delta Q = \sqrt{k\left(\left(\frac{83.9}{2.51^2}0.2\right)^2 + \left(\frac{21.7}{2.51^2}0.3\right)^2 + \left(-2\frac{83.9 \times 21.7}{2.51^3}0.04\right)^2\right)} = \sqrt{(1.0313)^2 + (0.4001)^2 + (-3.5664)^2} \Rightarrow \Delta Q = 3.734$$

Επομένως ΔQ = 4 (1 σημαντικό ψηφίο) . Το τελικό αποτέλεσμα είναι: $Q = 112 \pm 4$

Μετάδοση σφαλμάτων – Πράξεις

Σαν εφαρμογές του ορισμού της αβεβαιότητας μπορούμε να έχουμε:

(1) $Q = f(x_1, x_2) = a_1 x_1 + a_2 x_2$ με α_1 και α_2 σταθερές

Από το γενικό τύπο της απόκλισης θα έχουμε:

$$\Delta Q = \sqrt{\left(\frac{\partial f}{\partial x_1}\Delta x_1\right)^2 + \left(\frac{\partial f}{\partial x_2}\Delta x_2\right)^2} = \sqrt{(a_1 \Delta x_1)^2 + (a_2 \Delta x_2)^2}$$

Για $\alpha_1 = \alpha_2 = 1$ έχουμε: $\Delta Q = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2}$

Για άθροισμα ή διαφορά η απόκλιση είναι το άθροισμα των τετράγωνων

(2) $Q = f(x_1, x_2) = a_1 x_1 \times a_2 x_2$ με α_1 και α_2 σταθερές

$$\Delta Q = \sqrt{(a_1 a_2 x_2 \Delta x_1)^2 + (a_1 a_2 x_1 \Delta x_2)^2} \Rightarrow \Delta Q = a_1 a_2 \sqrt{(x_2 \Delta x_1)^2 + (x_1 \Delta x_2)^2}$$

Αν προσπαθήσουμε να βρούμε το σχετικό σφάλμα $\dfrac{\Delta Q}{Q}$ τότε παίρνουμε:

$$\frac{\Delta Q}{Q} = \frac{a_1 a_2}{a_1 a_2 x_1 x_2}\sqrt{x_2^2 \Delta x_1^2 + x_1^2 \Delta x_2^2} \Rightarrow \frac{\Delta Q}{Q} = \sqrt{\frac{\Delta x_1^2}{x_1^2} + \frac{\Delta x_2^2}{x_2^2}}$$

(3) $Q = f(x_1, x_2) = a_1 x_1 / a_2 x_2$ με α_1 και α_2 σταθερές

$$\Delta Q = \sqrt{\left(\frac{a_1}{a_2}\frac{\Delta x_1}{x_2}\right)^2 + \left(\frac{a_1}{a_2}\frac{x_1}{x_2^2}\Delta x_2\right)^2} \Rightarrow \Delta Q = \frac{a_1}{a_2}\sqrt{\frac{\Delta x_1^2}{x_2^2} + \frac{x_1^2 \Delta x_2^2}{x_2^4}}$$ με σχετικό σφάλμα $$\Rightarrow \frac{\Delta Q}{Q} = \sqrt{\frac{\Delta x_1^2}{x_1^2} + \frac{\Delta x_2^2}{x_2^2}}$$

Μετάδοση σφαλμάτων – πράξεις

(4) $Q = f(x) = ax^m$ με α και m σταθερές

$$\Delta Q = \sqrt{\left(amx^{(m-1)}\Delta x\right)^2} \Rightarrow \Delta Q = amx^{(m-1)}\Delta x$$ με σχετικό σφάλμα $$\frac{\Delta Q}{Q} = m\frac{\Delta x}{x}$$

(5) $Q = f(x) = ae^{mx}$ με α και m σταθερές

$$\Delta Q = \sqrt{\left(ame^{mx}\Delta x\right)^2} \Rightarrow \Delta Q = ame^{mx}\Delta x$$ με σχετικό σφάλμα $$\frac{\Delta Q}{Q} = m\Delta x$$

(6) $Q = f(x) = a\ln(mx)$ με α και m σταθερές

$$\Delta Q = \sqrt{\left(am\frac{\Delta x}{x}\right)^2} \Rightarrow \Delta Q = am\frac{\Delta x}{x}$$

Προσοχή: Ένα συνηθισμένο λάθος στον υπολογισμό σφαλμάτων προέρχεται από πράξεις μεταξύ των μεταβλητών που ορίζουν μια ποσότητα

Παράδειγμα: Έστω ότι μετράτε 2 αντιστάσεις, R_1 και R_2 οι οποίες είναι συνδεδεμένες παράλληλα και έχουν αβεβαιότητα ΔR_1 και ΔR_2. Θέλετε την αβεβαιότητα της ολικής αντίστασης R_{tot}.

$$\frac{1}{R_{tot}} = \frac{1}{R_1} + \frac{1}{R_2} \Rightarrow R_{tot} = \frac{R_1 R_2}{R_1 + R_2}$$

Αν προσπαθήσετε να εφαρμόσετε τους παραπάνω κανόνες ξεχωριστά για τον αριθμητή και παρονομαστή, η αβεβαιότητα της R_{tot} θα είναι λάθος αφού αριθμητής και παρονομαστής δεν είναι ανεξάρτητοι.

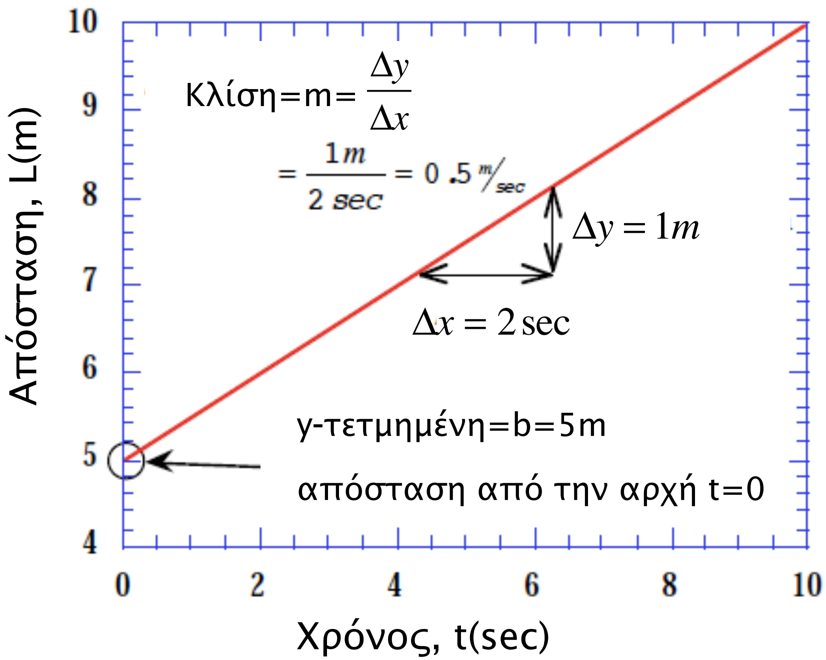
Προσαρμογή σε ευθεία γραμμή

Μια από τις περισσότερες χρήσιμες τεχνικές είναι αυτή της περιγραφής των πειραματικών δεδομένων με μια ευθεία γραμμή.

Υποθέστε ότι έχετε μια σειρά μετρήσεων σε μια γραφική παράσταση y ως προς x και από τη γραφική παράσταση βλέπουμε ότι αντιστοιχεί σε μια ευθεία γραμμή

Για την περιγραφή αυτής της ευθείας χρειάζονται 2 παράμετροι: η κλίση της, m, και η τετμημένη της ευθείας με τον άξονα των y, b

Τη στιγμή που θα προσδιορίσουμε τις δύο αυτές παραμέτρους μπορούμε να υπολογίσουμε την τιμή y που αντιστοιχεί σε οποιαδήποτε τιμή του x.



$$y = ax + b$$

Προσέξτε ότι είναι ακριβώς η κλίση και η τετμημένη b που παίζουν σημαντικό ρόλο στη θεωρία

Η κλίση είναι η ταχύτητα ενώ η τετμημένη μας δίνει την αρχική θέση του σώματος

Επομένως το πρόβλημά μας ανάγεται στην εύρεση των παραμέτρων της ευθείας καθώς και των αβεβαιοτήτων που συνοδεύουν τις εκτιμήσεις αυτών

Εύρεση της ευθείας

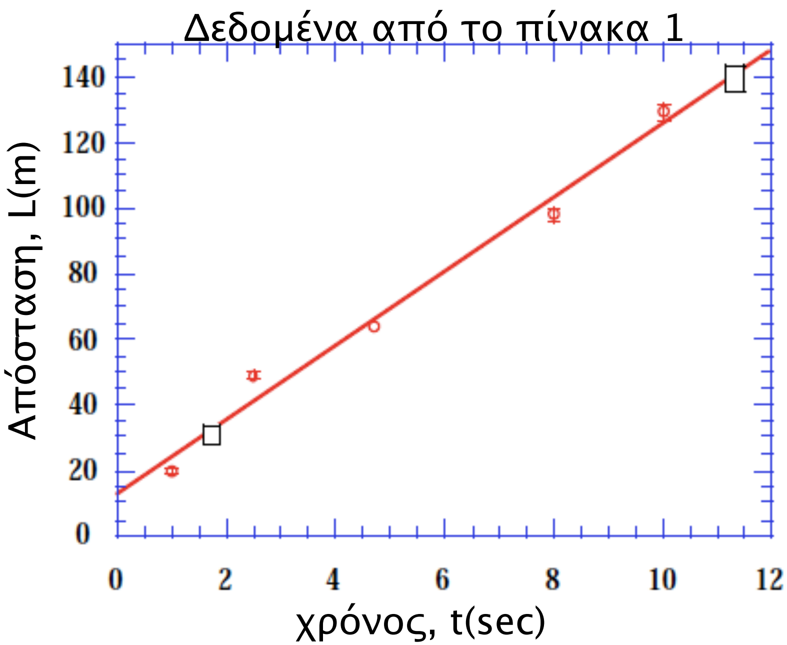
Έστω ότι έχουμε τις μετρήσεις που δίνονται στο παρακάτω πίνακα

Πίνακας 1

Χρόνος (sec)	Απόσταση (m)	Αβεβαιότητα, Δy (m)
1.0	20	1
2.5	59	1
4.7	64	(δεν χρειάζεται να υπολογισθεί)
8.0	98	2
10	129	3

Χρειάζεται να υπολογίσουμε τις αβεβαιότητες για 4 τιμές Τις 2 που βρίσκονται στο κατώτερο όριο τιμών και τις δύο στο υψηλότερο όριο τιμών

Σα 1ο βήμα κάνουμε τη γραφική παράσταση των δεδομένων του πίνακα 1



Χρησιμοποιώντας ένα χάρακα μπορούμε να σχεδιάσουμε τη καλύτερη ευθεία που διέρχεται από όλα τα σημεία

Αυτή η ευθεία λέγεται ευθεία “καλύτερης προσαρμογής (best fit)

Αν οι αβεβαιότητες όλων των σημείων είναι ίσες ή πολύ μικρές τότε η διαδικασία είναι πολύ απλή

Αν οι αβεβαιότητες παρουσιάζουν μεγάλες διακυμάνσεις τότε η διαδικασία είναι πιο πολύπλοκη

Ευθεία γραμμή καλύτερης προσαρμογής

Σημεία με μεγάλες αβεβαιότητες περιέχουν και τη λιγότερο σημαντικότητας πληροφορία και επομένως θα πρέπει να δώσουμε τη λιγότερο σημασία

Η τετμημένη με το y-άξονα μπορεί να βρεθεί διαβάζοντας απλά τη τιμή από τη γραφική παράσταση. Στη περίπτωσή μας είναι περίπου 13m

Για να βρούμε τη κλίση χρησιμοποιούμε το ορθογώνιο τρίγωνο της διαφ. 8 και υπολογίζουμε κάποιο διάστημα Δx και το αντίστοιχο Δy

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

Όπου τα σημεία x_1, x_2, y_1 και y_2 είναι κάποια σημεία της ευθείας γραμμής και όχι απαραίτητα πειραματικά σημεία

Αυτό είναι σημαντικό γιατί από τη στιγμή που σχεδιάσατε τη καλύτερη ευθεία δεν ενδιαφερόμαστε πλέον για τα πειραματικά σημεία αλλά για την κλίση και τη τετμημένη της ευθείας

Προσέξτε ότι για τη περίπτωσή μας κανένα από τα σημεία δεν βρίσκεται ακριβώς πάνω στην ευθεία

Χρησιμοποιούμε επομένως τα δεδομένα για να βρούμε τη καμπύλη και τη καμπύλη για να βρούμε τη θεωρία

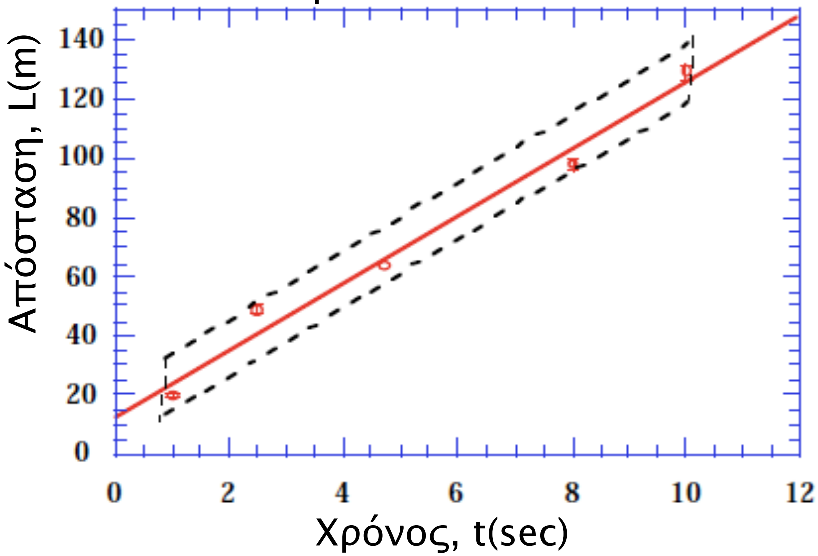
Για το παράδειγμά μας έχουμε: $m = \dfrac{32m - 140m}{1.7s - 11.3s} = 11.25m/s \rightarrow 11m/s$

Στο τελευταίο βήμα γράφουμε το αποτέλεσμα σύμφωνα με τα σημαντικά ψηφία που επιτρέπονται από τις μετρήσεις μας στο πίνακα 1 (δύο σημαντικά ψηφία)

Εύρεση της αβεβαιότητας της κλίσης και τετμημένης

Ξεκινάμε σχεδιάζοντας ένα παραλληλόγραμμο το οποίο περικλείει όλα τα πειραματικά σημεία συμπεριλαμβανομένης της αβεβαιότητάς τους

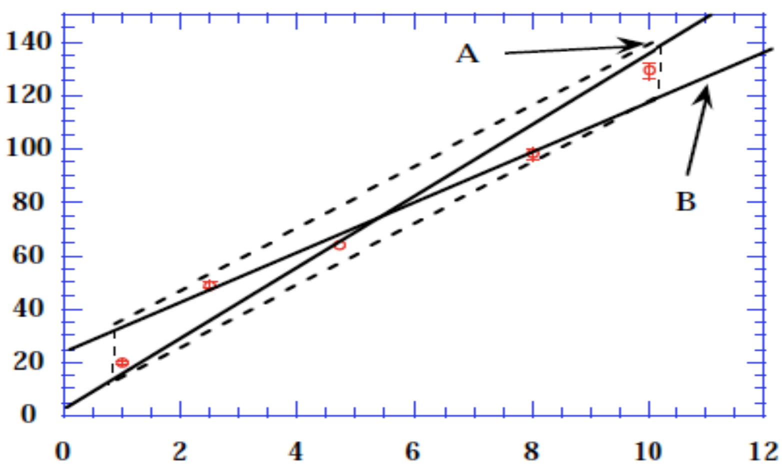
Το παραλληλόγραμμα της αβεβαιότητας φαίνεται στο παρακάτω σχήμα



Η πάνω και κάτω γραμμή αυτού του παρ/μου σχεδιάζονται παράλληλα προς την ευθεία της καλύτερης προσαρμογής

Τα άκρα σχεδιάζονται παρ/λα προς τον y–άξονα

Οποιαδήποτε εκτίμηση της αβεβαιότητας της κλίσης και τετμημένης της ευθείας θα πρέπει να έχει σαν αποτέλεσμα μια ευθεία η οποία περνά από τα άκρα του παρ/μου και δεν τέμνει τις πλάγιες πλευρές



Μπορούμε επομένως να χαράξουμε τις 2 διαγωνίους του παρ/μου και οι κλίσεις και τετμημένες των 2 αυτών ευθειών δίνουν την αβεβαιότητα στη κλίση και τετμημένη της ευθείας της καλύτερης προσαρμογής

Εύρεση της αβεβαιότητας της κλίσης και τετμημένης

Όπως και στην περίπτωση της καλύτερης ευθείας προσαρμογής υπολογίζουμε τη κλίση και τετμημένη των 2 διαγωνίων του παρ/μου

Έχουμε $m_A = \frac{10-140}{0.86-10} = 14m/s$ και $m_B = \frac{31-122}{0.86-10} = 10m/s$

Επομένως η αβεβαιότητα της κλίσης της καλύτερης ευθείας είναι:

$$m_A - m_B = 4m/s$$

Επειδή εκφράζουμε την αβεβαιότητα συνήθως συμμετρικά θα έχουμε ότι η αβεβαιότητα του πειράματος είναι

$$\Delta m = \frac{m_A - m_B}{2} = 2m/s$$

Βλέποντας τις 2 ευθείες της αβεβαιότητας καταλαβαίνουμε ότι είναι αρκετά απίθανο να επιλέξουμε είτε την ευθεία Α ή την ευθεία Β σαν την ευθεία της καλύτερης προσαρμογής. Επομένως έχουμε υπερεκτιμήσει την αβεβαιότητα

Μια προσεκτικότερη ανάλυση δείχνει ότι θα πάρουμε μια πιο καλή εκτίμηση της αβεβαιότητας αν διαιρέσουμε την προηγούμενη εκτίμησή μας με τη τετραγωνική ρίζα του αριθμού των μετρήσεών μας

Γράφουμε: $\Delta m = \frac{m_A - m_B}{2}\frac{1}{\sqrt{N-1}}$

$$\Delta m = \frac{m_A - m_B}{2}\frac{1}{\sqrt{N-1}}$$

Η εξίσωση αυτή δίνει την εκτίμηση της πιο πιθανής αβεβαιότητας.

Στο παρονομαστή, για N=1, Δm=∞ και αυτό είναι λογικό μια και από ένα σημείο μπορούμε να έχουμε οποιαδήποτε ευθεία

Για N=2 μια γραμμή μόνο μπορεί να χαραχθεί από 2 σημεία. Στην περίπτωση αυτή το παρ/μο της αβεβαιότητας προέρχεται μόνο από τις αβεβαιότητες των δύο αυτών σημείων και επομένως είναι λογικό $\Delta m = (m_A - m_B)/2$

Για N>>2 η αβεβαιότητα ελλατώνεται σύμφωνα με τη ρίζα του αριθμού των μετρήσεων N και αυτό είναι το αποτέλεσμα που βρήκαμε όταν υπολογίσαμε του σφάλματος της μέσης τιμής

Αυτό δεν αποτελεί τη λύση του προβλήματος αλλά είναι κάποια πολύ λογική και καλή προσέγγιση

Η εύρεση της αβεβαιότητας της τετμημένης προχωρά σύμφωνα με τα όσα αναπτύξαμε για την αβεβαιότητα της κλίσης της ευθείας.

Κάνοντας τις πράξεις έχουμε ότι $m \pm \Delta m = (11 \pm 1)m/s$ και $b \pm \Delta b = (13 \pm 6)m$

Η αβεβαιότητα της κλίσης είναι ίδιας τάξης με τα δεδομένα αλλά η αβεβαιότητα της τετμημένης είναι περίπου 50%

Στη περίπτωση της κλίσης παίρνουμε μια μέση τιμή ενώ για τη τετμημένη προεκτείνουμε σε περιοχή μακριά από τα δεδομένα και σε χρόνους που δεν έχουμε πειραματικές μετρήσεις και είναι επόμενο να έχουμε μεγαλύτερη αβεβαιότητα

Σε κάποιο σημείο όχι πολύ μακριά από τα δεδομένα η αβεβαιότητα σχετικά με τη θέση του σώματος τη χρονική στιγμή t=0 μπορεί να γίνει ίση με 100% και μεγαλύτερη ακόμα.
Αυτό σημαίνει ότι το πείραμά μας είναι πολύ δύσκολο να προσδιορίσει το τι κάνει το σώμα σε μια προγενέστερη χρονική στιγμή στην οποία δεν υπάρχουν μετρήσεις

Μπορούμε δηλαδή να εκφράσουμε άποψη μόνο σχετικά με τη κίνηση στο διάστημα που μετρήσαμε αλλά όχι πέρα από αυτό

Μπορεί να προσπαθήσει κάποιος να επιχειρηματολογήσει ότι αν το σώμα κινείται με σταθερή ταχύτητα κατά τη διάρκεια του διαστήματος ότι κινείται με τον ίδιο τρόπο έξω από το χρονικό διάστημα της μέτρησής μας.
Αυτό μπορεί να είναι σωστό αλλά δεν έχουμε μετρήσεις οι οποίες μπορούν να δείξουν ότι αυτό συμβαίνει. Για να ελαττώσουμε την αβεβαιότητα χρειαζόμαστε περισσότερες μετρήσεις.

Ένα τελευταίο σημείο. Ακόμα και αν υπήρχε πειραματικό σημείο στον y-άξονα αυτό δεν σημαίνει ότι δεν υπάρχει αβεβαιότητα στη τιμή της τετμημένης. Αυτό γιατί η ευθεία καλύτερης προσαρμογής δεν είναι απαραίτητο να περνά από όλα τα πειραματικά σημεία όπως συμβαίνει και στο παράδειγμά μας

Ημιλογαριθμικό χαρτί

Πολλές φορές τα δεδομένα μας δεν περιγράφονται από μια απλή ευθεία αλλά ο νόμος της φυσικής που περιγράφει το φαινόμενο έχει μια εκθετική μορφή

Για παράδειγμα η ραδιενεργός διάσπαση κάποιων ραδιοισοτόπων. Η διάσπαση ακολουθεί εκθετική μορφή σύμφωνα με τη σχέση

$$A(t) = A_0 e^{-t/\tau}$$

Όπου A_0 είναι η ενεργότητα τη στιγμή t=0 και τ η σταθερά διάσπασης

Μπορούμε και πάλι να χρησιμοποιήσουμε την ίδια τεχνική για να βρούμε τη σταθερά διάσπασης και την αρχική ενεργότητα του δείγματος. Αρκεί να γράψουμε την παραπάνω εξίσωση σε γραμμική μορφή

Η μετατροπή της εκθετικής εξίσωσης σε γραμμική γίνεται εύκολα λογαριθμίζοντας την εξίσωση

$$\ln[A(t)] = \ln[A_0 e^{-t/\tau}] = \ln[A_0] - \frac{t}{\tau} = \ln[A_0] + \lambda t$$

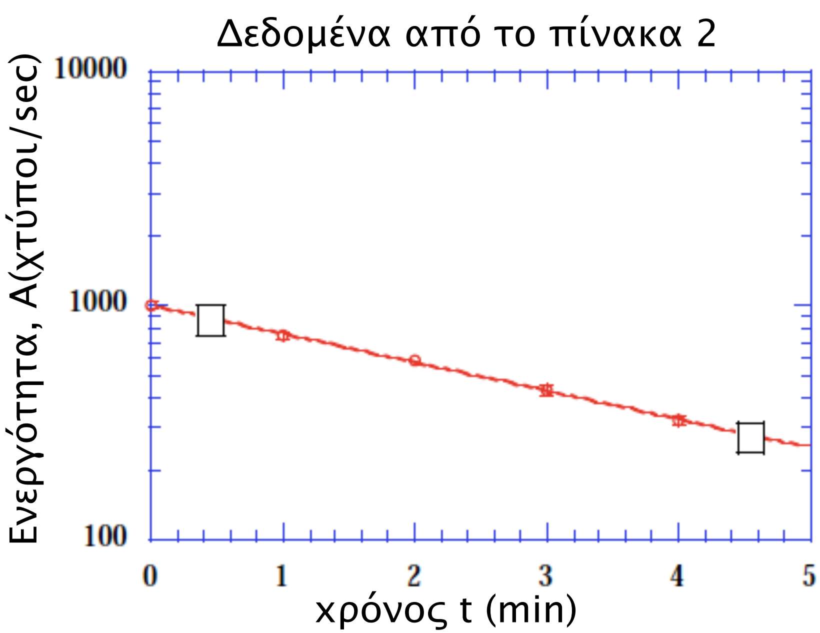
Ο όρος $\ln[A_0]$ αντιπροσωπεύει το σταθερό όρο της εξίσωσης της ευθείας ενώ ο όρος λ=–1/τ την κλίση της. Ο όρος t αποτελεί την ανεξάρτητη μεταβλητή ενώ ο λογάριθμος της ενεργότητας, ln[A(t)], την εξαρτόμενη μεταβλητή

Για την γραφική παράσταση θα μπορούσαμε να υπολογίσουμε τους λογαρίθμους και να χρησιμοποιήσουμε χιλιοστομετρικό χαρτί ή να χρησιμοποιήσουμε το λεγόμενο ημιλογαριθμικό χαρτί

Ημιλογαριθμικό χαρτί

Η χρήση του λογαριθμικού χαρτιού διευκολύνει στη περίπτωση αυτή γιατί μπορούμε να θέσουμε τις μετρήσεις μας απευθείας στο γράφημα χωρίς επιπλέον υπολογισμούς. Το χαρτί περιέχει το κάθετο άξονα με τέτοιο τρόπο ώστε οι Κατακόρυφες αποστάσεις να είναι ανάλογες των επιθυμητών λογαρίθμων

Ο παρακάτω πίνακας περιέχει τα δεδομένα της ραδιενεργούς διάσπασης

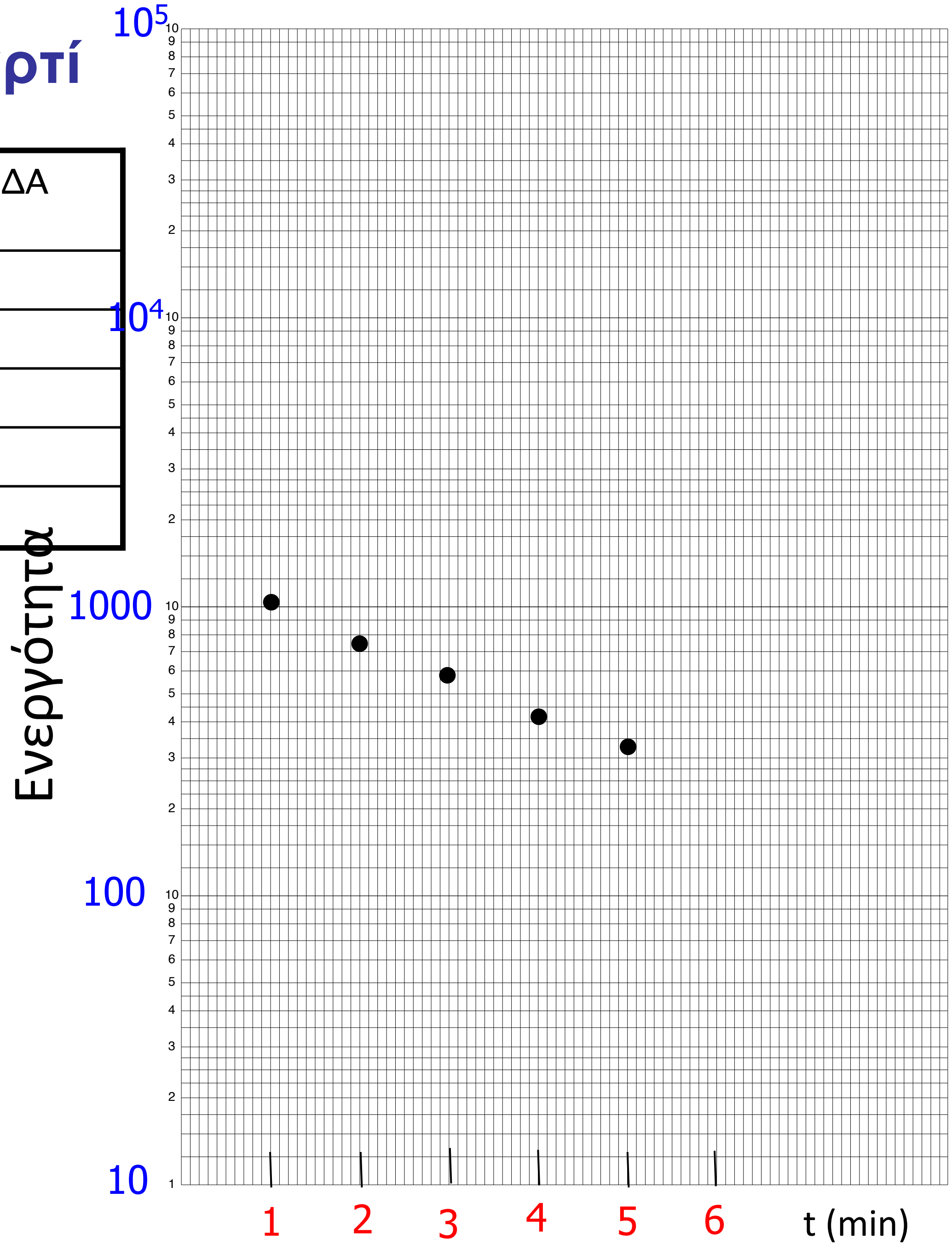


Πίνακας 2. Ενεργότητα ως προς χρόνο για το ραδιενεργό δείγμα μας

Χρόνος (min)	Ενεργότητα A και ΔA (χτύποι/s)
1	1010±31
2	740±28
3	585±24
4	435±21
5	325±18

Ημιλογαριθμικό χαρτί

Χρόνος (min)	Ενεργότητα Α και ΔΑ (χτύποι/s)
1	1010±31
2	740±28
3	585±24
4	435±21
5	325±18



Ημιλογαριθμικό χαρτί

Χρειάζεται ωστόσο κάποια προσοχή, γιατί τα δεδομένα μας δίνονται αυτόματα σε λογαριθμική κλίμακα, αλλά οι λογάριθμοι δεν υπολόγιστηκαν. Απλά το χαρτί παρέχει το κατάλληλο μετασχηματισμό. Όταν όμως πρέπει να υπολογίσουμε τη κλίση στο λογαριθμικό χαρτί θα πρέπει να υπολογίσουμε τους λογαρίθμους.

Για παράδειγμα :

$$\lambda = \frac{\ln(860\,/\,\mathrm{sec}) - \ln(282\,/\,\mathrm{sec})}{0.5\,\mathrm{min} - 4.5\,\mathrm{min}} = -0.27\,/\,\mathrm{min} = -\frac{1}{\tau} \quad \Rightarrow \tau = 3.70\,\mathrm{min}$$

Η τετμημένη θα δίνεται από

$$A_0 = \frac{1.00 \times 10^3}{\mathrm{sec}} \times \frac{60\,\mathrm{sec}}{1\,\mathrm{min}} = 6.00 \times 10^4 \;/\,\mathrm{min}$$

Όπως βλέπουμε χρειάζεται να υπολογίσουμε ένα λογάριθμο για να βρούμε την κλίση ενώ η τετμημένη δίνεται απευθείας από το χαρτί

Ο τρόπος υπολογισμού της αβεβαιότητας της κλίσης και τετμημένης είναι ακριβώς ίδιος όπως και στην περίπτωση της γραμμικής περίπτωσης. Θα πρέπει να σχεδιάσουμε το παρ/μο αβεβαιότητας και να υπολογίσουμε τις κλίσεις των δύο διαγωνίων και τις αντίστοιχες τετμημένες τους.

Log-Log χαρτί (λογαριθμικό – λογαριθμικό)

Πολλές φορές μπορούμε να βρούμε την συναρτησιακή εξάρτηση ενός φυσικού μεγέθους y από το ανεξάρτητο μέγεθος x, θεωρώντας το λογάριθμο των δεδομένων που μετράμε.

Αν το εξαρτόμενο μέγεθος y ειναι ανάλογο κάποιας δύναμης του ανεξάρτητου μεγέθους x, τότε το γράφημα του y ως προς x, σχεδιαζόμενο σε ένα λογαριθμικό-λογαριθμικό χαρτί θα είναι ευθεία η κλίση της οποίας θα είναι ίση με τον εκθέτη στον οποίο είναι υψωμένο το ανεξάρτητο μέγεθος.

Για παράδειγμα, έστω ότι μετρούμε κάποια δεδομένα τα οποία κατανέμονται σύμφωνα με την εξίσωση $y = x^n$

Θεωρώντας το λογάριθμο θα έχουμε: $\log(y) = \log\left(x^n\right) \Rightarrow \log(y) = n\log(x)$

Προφανώς από ένα γράφημα με λογαριθμικούς άξονες μπορούμε να βρούμε αμέσως τη κλίση

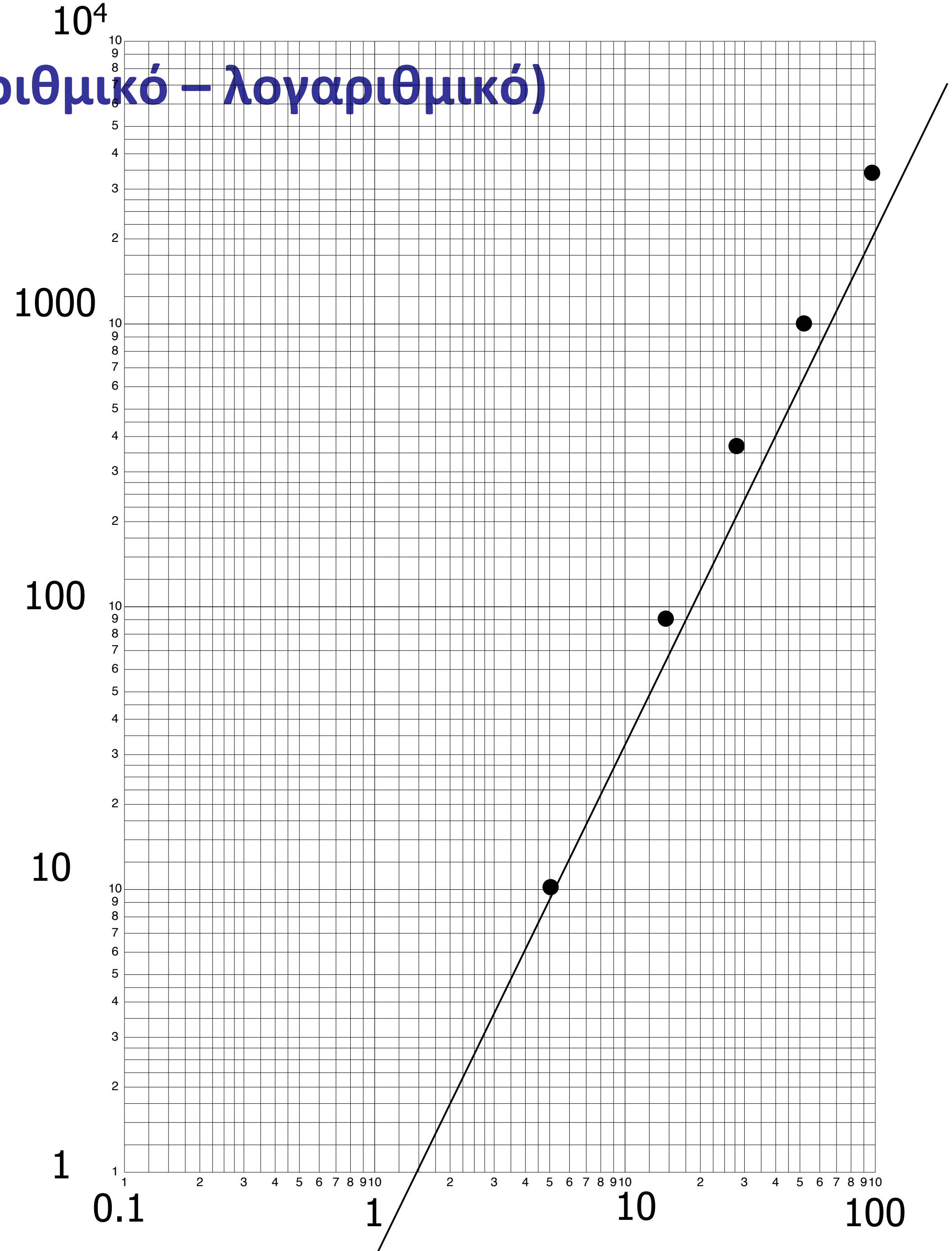
Αν είχαμε περισσότερο πολύπλοκη μορφή:

$$y = Ax^n \Rightarrow \log(y) = \log\left(Ax^n\right) = \log(A) + n\log(x)$$

Έχει διαφορά στο γράφημα αν θεωρήσουμε Log_{10} ή ln ($\log_e$) σχέσεις?

Log-Log χαρτί (λογαριθμικό – λογαριθμικό)

Χ	Υ
5	10
15	90
30	360
50	1000
95	3610



Εύρεση καλύτερης ευθείας προσαρμογής

Μέθοδος ελαχίστων τετραγώνων (μέθοδος χ^2)

Όπως είδαμε στη γραφική εύρεση της καλύτερης ευθείας προσαρμογής, αυτό που ενδιαφερόμαστε είναι ο καθορισμός των παραμέτρων της συνάρτησης (π.χ. μιας ευθείας) που περιγράφει τα δεδομένα

Δύο μεθόδοι: Ελαχίστων τετραγώνων ή χ^2

Μέγιστης πιθανότητας – maximum likelihood

Υποθέτουμε ότι έχουμε μια συνάρτηση y μιας μεταβλητής x και μια σειρά παραμέτρων $\vec{\theta} = (\theta_1, \theta_2, \cdots, \theta_N)$

$$y = y(x;\vec{\theta})$$

Έστω ότι μετρήσαμε διάφορες τιμές του y για κάποιες τιμές της ανεξάρτητης μεταβλητής x.

Επομένως θα έχουμε N ζεύγη τιμών $\left(x_i, y_i \pm \sigma_{y_i}\right)$ όπου $\iota = 1,\ldots,N$

Ορίζουμε σαν "χ^2" τη ποσότητα:

$$\boxed{\chi^2\left(x;\vec{\theta}\right) = \sum_{i=1}^{N}\left[\frac{y_i - y(x_i;\vec{\theta})}{\sigma_i}\right]^2}$$

Η μέθοδος των ελαχίστων τετραγώνων δηλώνει ότι η καλύτερη εκτίμηση των παραμέτρων θ_i επιτυγχάνεται όταν βρεθεί μια ομάδα τιμών θ_i για τις οποίες η συνάρτηση χ^2 είναι ελάχιστη

Μέθοδος ελαχίστων τετραγώνων – χ^2

Αν οι αποκλίσεις κάθε μέτρησης, σ_i, είναι ίσες τότε η σχέση απλουστεύεται

$$\chi^2 = \frac{1}{\sigma^2}\sum_{i=1}^{N}\left[y_i - y\left(x_i;\vec{\theta}\right)\right]^2$$

Είναι σημαντικό να προσέξετε ότι η μέθοδος όπως ορίστηκε χρειάζεται τη γνώση των αβεβαιοτήτων σ_i και ότι υποθέτει ότι δεν υπάρχουν αβεβαιότητες στη γνώση της ανεξάρτητης μεταβλητής x.

Αβεβαιότητες στις τιμές x_i μπορούν να αγνοηθούν εφόσον:

$$\frac{\sigma_{x_i}}{x_i} << \frac{\sigma_{y_k}}{y_k} \quad \mu\varepsilon \;\; i = 1,2,...,N, \;\; k = 1,2,...,N$$

Αγνοώντας τις αβεβαιότητες, η μέθοδος ελαχίστων τετραγώνων είναι απλά το άθροισμα των αποστάσεων κάθε μέτρησης y_i από τα σημεία της θεωρητικής καμπύλης $(x_i,y_i(x_i))$. Ελαχιστοποιώντας τη συνάρτηση ψάχνουμε τα σημεία της καλύτερης καμπύλης για τα οποία αυτή η απόσταση ελαχιστοποιείται

Η εισαγωγή των αβεβαιοτήτων είναι απαραίτητη αν θέλουμε να κάνουμε στατιστική ανάλυση των αποτελεσμάτων διαφορετικά είναι απλό γεωμετρικό πρόβλημα

Εφαρμογή χ^2 – εύρεση παραμέτρων ευθείας

Ας υποθέσουμε ότι η συνάρτηση την οποία θέλουμε να προσαρμόσουμε στα σημεία των μετρήσεων μας είναι της μορφής

$$y(x) = ax + b$$

Θα υποθέσουμε ακόμα ότι οι αβεβαιότητες σ_i των y_i είναι όλες ίσες μεταξύ τους

Επομένως το πρόβλημα εύρεσης των παραμέτρων α και b έγγυται στην ελαχιστοποίησης της συνάρτησης χ^2 ως προς a και b

Η ελαχιστοποίηση γίνεται πέρνοντας τη μερική παράγωγο της χ^2 ως προς a και b και εξισώνοντας με μηδέν:

$$\frac{\partial \chi^2}{\partial a} = 0 \qquad \text{και} \qquad \frac{\partial \chi^2}{\partial b} = 0$$

και λύνοντας το σύστημα των γραμμικών εξισώσεων που προκύπτει προς a και b

$$\left.\begin{array}{l}\dfrac{\partial}{\partial a}\left[\sum_{i=1}^{N}(y_i - ax_i - b)^2\right] = -2\sum_{i=1}^{N}(y_i - ax_i - b)x_i = 0 \\ \dfrac{\partial}{\partial b}\left[\sum_{i=1}^{N}(y_i - ax_i - b)^2\right] = -2\sum_{i=1}^{N}(y_i - ax_i - b) = 0\end{array}\right\} \Rightarrow \left.\begin{array}{l}\sum_{i=1}^{N} y_i x_i = a\sum_{i=1}^{N} x_i^2 + b\sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} y_i = a\sum_{i=1}^{N} x_i + Nb\end{array}\right\} \Rightarrow b = \frac{\sum_{i=1}^{N} y_i}{N} - a\frac{\sum_{i=1}^{N} x_i}{N}$$

Αντικαθιστούμε στην 1η εξίσωση και λύνουμε ως προς a:

$$\sum_{i=1}^{N} y_i x_i = a\sum_{i=1}^{N} x_i^2 + \frac{1}{N}\sum_{i=1}^{N} y_i \sum_{i=1}^{N} x_i - \frac{a}{N}\sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i \Rightarrow a\left[N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2\right] = N\sum_{i=1}^{N} y_i x_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i$$

Εύρεση παραμέτρων ευθείας με μέθοδο χ^2

Από την τελευταία εξίσωση παίρνουμε το α και αντικαθιστώντας στην εξίσωση του b παίρνουμε το b

$$a = \frac{N\sum_{i=1}^{N} y_i x_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\left[N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2\right]} = \frac{N\sum xy - \sum x \sum y}{\Delta}$$

και

$$b = \frac{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i x_i}{\left[N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2\right]} = \frac{\sum x^2 \sum y - \sum x \sum xy}{\Delta}$$

Αν διαιρέσουμε με N^2 τον αριθμητή και παρονομαστή στις παραπάνω σχέσεις

$$a = \frac{\overline{yx} - \overline{x}\ \overline{y}}{\left(\overline{x^2} - \overline{x}^2\right)}$$

και

$$b = \frac{\overline{y}\,\overline{x^2} - \overline{x}\ \overline{xy}}{\left(\overline{x^2} - \overline{x}^2\right)}$$

όπου $\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \overline{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$

Παράδειγμα

Έστω ότι κάποιος φοιτητής θέλει να μετρήσει τις μάζες διαφόρων σωμάτων με μια ζυγαριά ελατηρίου. Θα πρέπει πριν μετρήσει τις μάζες να βαθμονομήσει τη ζυγαριά. Για να το κάνει αυτό χρησιμοποιεί 5 γνωστές μάζες των 2kgr τις οποίες τοποθετεί διαδοχικά πάνω στη ζυγαριά και μετρά κάθε φορά το αντίστοιχο μήκος του ελατηρίου l_k. Υποθέτοντας ότι το ελατήριο υπακούει στο νόμο του Hooke, περιμένει ότι $l = A + Bm$

Η σταθερά Α είναι το φυσικό μήκος τους εκκρεμούς και Β είναι B=g/k, k η σταθερά ελατηρίου

Χρησιμοποιόντας την γραμμική σχέση $l = A + Bm$ μπορεί μετρώντας το μήκος του ελατηρίου για μια άγνωστη μάζα m να βρει τη μάζα. Χρειάζεται επομένως τις τιμές Α και Β.

Μέτρηση	Μάζα m	Μήκος, l	m²	$m_i l_i$
1	2	42.0	4	84
2	4	48.4	16	194
3	6	51.3	36	308
4	8	56.3	64	450
5	10	58.6	100	586
N=5	$\sum m_i = 30$	$\sum l_i = 256.6$	$\sum m_i^2 = 220$	$\sum m_i l_i = 1622$

Επομένως:

$$\Delta = N\sum m^2 - \left(\sum m\right)^2$$

$$\Rightarrow \Delta = 5 \times 220 - 30^2 = 200$$

Η σταθερά Α είναι:

$$A = \frac{\sum m^2 \sum l - \sum m \sum l}{\Delta}$$

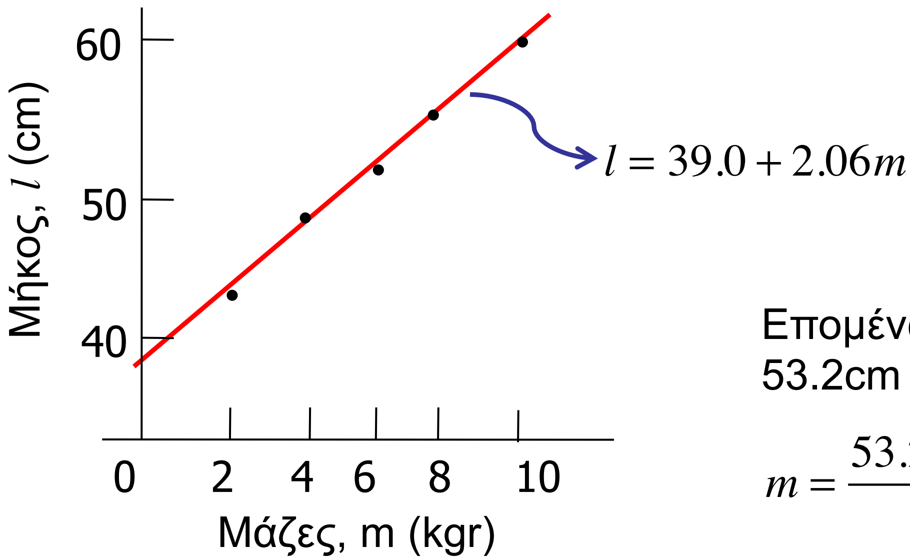
$$A = \frac{220 \times 256.6 - 30 \times 1622}{200}$$

$$A = 39.0cm$$

Η σταθερά Β θα είναι $\mathrm{B} = \dfrac{N\sum ml - \sum m \sum l}{\Delta} = \dfrac{5 \times 1622 - 30 \times 256.6}{200} \Rightarrow B = 2.06cm / kgr$

Γραφική παράσταση

Το γράφημα των προηγούμενων μετρήσεων θα είναι:



Επομένως αν κάποια μάζα επιμηκύνει το ελατήριο κατά 53.2cm τότε σύμφωνα με την εξίσωση της χ^2 ευθείας:

$$m = \frac{53.2 - 39.0}{2.06} = 6.9kg$$

Χρειάζεται να υπολογίσουμε τις αβεβαιότητες των A και B.

Αρχικά όμως ποια είναι η αβεβαιότητα των μετρήσεων y?

Έχουμε 5 μετρήσεις αλλά η διασπορά τους δεν μας δίνει την αβεβαιότητα τους. Ωστόσο κάθε μέτρηση περιμένουμε να κατανέμεται σύμφωνα με την Gaussian κατανομή γύρω από την αληθινή τιμή $y = A + Bx_i$ με εύρος σ_i

Επομένως όλες οι αποκλίσεις $y_i - A - Bx_i$ θα είναι κατανεμημένες Gaussian με κεντρική τιμή 0 και το ίδιο εύρος σ_y.

Το εύρος σ_y δίνεται από την σχέση $\boxed{\sigma_y = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}\left(y_i - A - Bx\right)^2}}$

Εύρεση αβεβαιότητας παραμέτρων ευθείας

Οι αβεβαιότητες των a και b βρίσκονται από εφαρμογή διάδοσης σφαλμάτων και …πολύ άλγεβρα

$$a = \sum_{i=1}^{N} \left[\frac{\left(x_i - \overline{x}\right) y_i}{\left(\overline{x^2} - \overline{x}^2\right)} \right] \Rightarrow \sigma_a^2 = \sum_{i=1}^{N} \sigma_{y_i}^2 \left(\frac{\partial a}{\partial y_i} \right)^2 \Rightarrow \sigma_a^2 = \sigma_y^2 \sum_{i=1}^{N} \left[\frac{x_i - \overline{x}}{N\left(\overline{x^2} - \overline{x}^2\right)} \right]^2 \Rightarrow \sigma_a^2 = \sigma_y^2 \sum_{i=1}^{N} \left[\frac{x_i^2 + \overline{x}^2 - 2\overline{x}x_i}{N^2\left(\overline{x^2} - \overline{x}^2\right)^2} \right]$$

Αλλά $$\sum_{i}^{N} \left(x_i^2 + \overline{x}^2 - 2\overline{x}x_i\right) = \sum_{i=1}^{N} x_i^2 + \sum_{i=1}^{N} \overline{x}^2 - 2\overline{x} \boxed{\sum_{i=1}^{N} x_i} = \sum_{i=1}^{N} x_i^2 + N\overline{x}^2 - 2\overline{x}\, \boxed{N\overline{x}} = \sum_{i=1}^{N} x_i^2 - N\overline{x}^2$$

Αντικαταστούμε την τελευταία σχέση στον αριθμητή της εξίσωσης για σ_a^2 οπότε

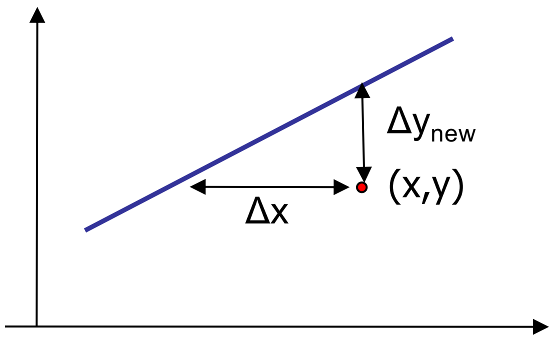
$$\sigma_a^2 = \sigma_y^2 \frac{\sum_{i=1}^{N} x_i^2 - N\overline{x}}{N^2\left(\overline{x^2} - \overline{x}^2\right)^2} = \sigma_y^2 \frac{\left(N\overline{x^2} - N\overline{x}\right)}{N^2\left(\overline{x^2} - \overline{x}^2\right)^2} = \sigma_y^2 \frac{N\left(\overline{x^2} - \overline{x}^2\right)}{N^2\left(\overline{x^2} - \overline{x}^2\right)^2} \Rightarrow \sigma_a^2 = \sigma_y^2 \frac{1}{N\left(\overline{x^2} - \overline{x}^2\right)}$$

Ανάλογα βρίσκουμε και την αβεβαιότητα του b $$\sigma_b^2 = \sigma_y^2 \frac{\overline{x^2}}{N\left(\overline{x^2} - \overline{x}^2\right)}$$

$$\sigma_y = \sqrt{\frac{1}{N-2} \sum_{i=1}^{N} \left(y_i - A - Bx\right)^2}$$

Περίπτωση με σφάλμα σε x και y

Αν οι μετρήσεις έχουν σφάλμα σε x, σ_x και y, σ_y, τότε μπορούμε να βρούμε ένα ισοδύναμο σφάλμα στην διεύθυνση y που προκαλεί το ίδιο αποτέλεσμα



Υποθέτοντας αρχικά ότι $\sigma_y = 0$

$$\Delta y_{new} = \frac{dy}{dx}\Delta x \quad \Rightarrow \sigma y_{new} = \frac{dy}{dx}\sigma_x$$

Αν σ_y δεν είναι 0 τότε το νέο σφάλμα μπορεί να γραφεί

$$\sigma y_{new} = \sqrt{\sigma_y^2 + \left(\frac{dy}{dx}\sigma_x\right)^2}$$

αφού οι δυό αβεβαιότητες είναι ανεξάρτητες μεταξύ τους

Περίληψη – Μέθοδος χ²

Οπότε συνοψίζοντας τη μέθοδο των ελαχίστων τετραγώνων για την περίπτωση της ευθείας και υποθέτοντας ότι τα σφάλματα των N επιμέρους μετρήσεων, y_i, είναι ίσα μεταξύ τους ότι :

κλίση

$$a = \frac{N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\Delta}$$

$$\sigma_a^2 = \sigma_y^2 \frac{N}{\Delta}$$

όπου:

$$\Delta = N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2$$

τεταγμένη

$$b = \frac{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i y_i}{\Delta}$$

$$\sigma_b^2 = \sigma_y^2 \frac{\sum_{i=1}^{N} x_i^2}{\Delta}$$

αβεβαιότητα μετρήσεων:

$$\sigma_y = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}\left(y_i - A - Bx\right)^2}$$

Η συσχέτιση μεταξύ των x και y:

$$r^2 = \frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2 \sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} = \frac{\left(\sum_{i=1}^{N} x_i y_i - N\bar{x}\,\bar{y}\right)^2}{\left(\sum_{i=1}^{N} x_i^2 - N\bar{x}^2\right)\left(\sum_{i=1}^{N} y_i^2 - N\bar{y}^2\right)}$$

Μέθοδος ελαχίστων τετραγώνων – σ_{y_i} άνισα

Αν τα σφάλματα των μετρήσεων είναι διαφορετικά μεταξύ τους τότε η συνάρτηση που θα πρέπει να ελαχιστοποιήσουμε είναι

$$\boxed{\chi^2 = \sum_{i=1}^{N}\left[\frac{y_i - ax_i^2 - b}{\sigma_i}\right]^2}$$

Η ελαχιστοποιήση δίνει τις ίδιες εξισώσεις μόνο που στη περιπτώση αυτή οι μέσες τιμές των μεγεθών αντιστοιχούν σε αυτές που προκύπτουν με το να ζυγίσουμε τις τιμές με τα ανάλογα βάρη ($1/\sigma_i^2$) και η κανονικοποίηση δεν γίνεται ως προς N αλλά ως προς το συνολικό βάρος $\sum_{i=1}^{N} 1/\sigma_i^2$

Οι σχέσεις που δίνουν τις παραμέτρους α και b γίνονται (y = αx + b):

$$b = \frac{\sum wx^2 \sum wy - \sum wx \sum wxy}{\Delta} \qquad \sigma_b = \sqrt{\frac{\sum wx^2}{\Delta}}$$

$$a = \frac{\sum w \sum wxy - \sum wx \sum wy}{\Delta} \qquad \sigma_a = \sqrt{\frac{\sum w}{\Delta}}$$

$$\Delta = \sum w \sum wx^2 - \left(\sum wx\right)^2$$