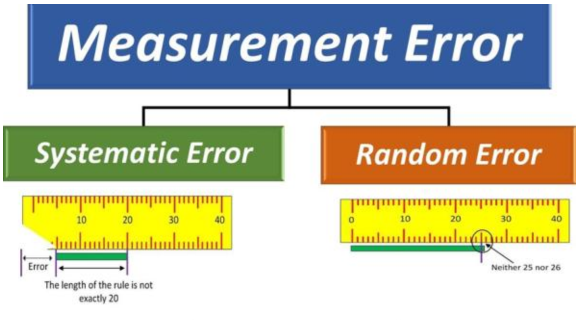


Ανάλυση Σφάλματος Μέτρησης

Αντώνης Κ. Τραυλός (B.A., M.A., Ph.D.)
Καθηγητής
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
Σχολή Επιστημών Ανθρώπινης Κίνησης και Ποιότητας Ζωής
Τμήμα Οργάνωσης και Διαχείρισης Αθλητισμού

Η διάλεξη αυτή βασίζεται σε ενότητες του Κεφαλαίου 1 του Βαγενάς (2019).

1

Το δυσκολότερο μέρος της στατιστικής μεθοδολογίας

- Κατά γενική παραδοχή το δυσκολότερο μέρος της στατιστικής μεθοδολογίας είναι αυτό της συλλογής έγκυρων δεδομένων.
- Η **συλλογή των αρχικών δεδομένων** πρέπει να γίνεται με τον πιο αντικειμενικό και έγκυρο τρόπο που προσφέρεται ανάλογα με το είδος του ερευνητικού προβλήματος.

2

Εγκυρότητα και αξιοπιστία δεδομένων

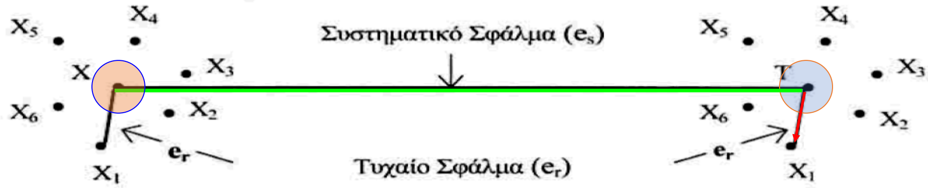
- Αν τα αρχικά δεδομένα δεν έχουν αποδεδειγμένη **εγκυρότητα (validity)** και **αξιοπιστία (reliability)**, τότε καμία στατιστική ανάλυση δεν μπορεί να τα αναβαθμίσει.
- **Ακρίβεια (accuracy)**: βαθμός εγγύτητας ενός μετρητικού αποτελέσματος (X) προς την πραγματική τιμή (T), ενώ **επακρίβεια (precision)** είναι ο βαθμός επανάληψης, αναπαραγωγής του μετρητικού αποτελέσματος (X_1, X_2,..., X_N).
- Οι ιδιότητες αυτές αφορούν δύο διαφορετικές πηγές του ολικού σφάλματος μέτρησης:
- (α) την **ανακρίβεια (inaccuracy)** ή **μεροληψία** (**bias**), που προκαλεί το μεροληπτικό ή **συστηματικό σφάλμα (systematic error, e_s),** και
- (β) την **ανεπακρίβεια (imprecision)** ή **μη επαναληπτικότητα** (**non-repeatability**), που προκαλεί το αμερόληπτο ή **τυχαίο σφάλμα (random error, e_r).**

Ολικό Σφάλμα Μέτρησης : $e = e_s + e_r$

$$X_1 = T + e = T + e_s + e_r = T + TX + XX_1$$

- Έστω **T** η πραγματική τιμή (true value) μιας υπό μέτρηση ποσότητας X.
- Η **τιμή T** υπάρχει, είναι συγκεκριμένη, αλλά δεν είναι ποτέ άμεσα γνωστή.
- **Κάθε προσπάθεια για τη μέτρηση της T** δίνει μια τιμή X_l, X_2, X_3, X_4, X_5, X_6 (τα 6 σημεία κοντά στην T).
- Οι **αποστάσεις των σημείων αυτών από την T** (π.χ., TX) συμβολίζουν τη διασπορά του τυχαίου σφάλματος ($T – X_1 = e_r$), που προκαλείται από αποκλίσεις μεταξύ των 6 μετρήσεων.

Ολικό Σφάλμα Μέτρησης : $e = e_s + e_r$



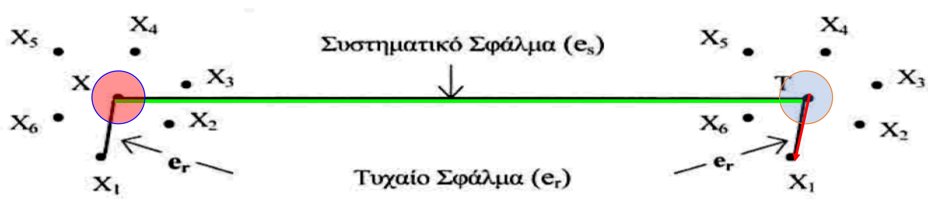
$$X_1 = T + e = T + e_s + e_r = T + TX + XX_1$$

- Η **απόσταση TX συμβολίζει το συστηματικό σφάλμα (e_s)**, που υπεισέρχεται σε κάθε μέτρηση ως σταθερή απόκλιση από την πραγματική τιμή T και είναι αποτέλεσμα βλάβης ή κακής διαμέτρησης του οργάνου μέτρησης, που διορθώνεται με βαθμονόμηση (calibration).

5

5

Ολικό Σφάλμα Μέτρησης : $e = e_s + e_r$



$$X_1 = T + e = T + e_s + e_r = T + TX + XX_1$$

- Έτσι, αν σε μια σειρά μετρήσεων της ίδιας ποσότητας T υπεισέλθουν και οι δύο πηγές σφάλματος, τότε το αποτέλεσμα (X) είναι ίσο με $T + e_s \pm e_r$.
- Έτσι, σε κάθε μέτρηση αντί της πραγματικής τιμής T βρίσκουμε μια τιμή $X_i = T \pm TX \pm XX_i$.
- Αν εξαλειφθεί το συστηματικό σφάλμα ($e_s = \pm TX = 0$), τότε η μέτρηση περιέχει μόνο τυχαίο σφάλμα $X_i = T \pm XX_i$ και η πραγματική τιμή της μετρούμενης ποσότητας $T = (X_1 + X_2 + . . . + X_N) / N$.

6

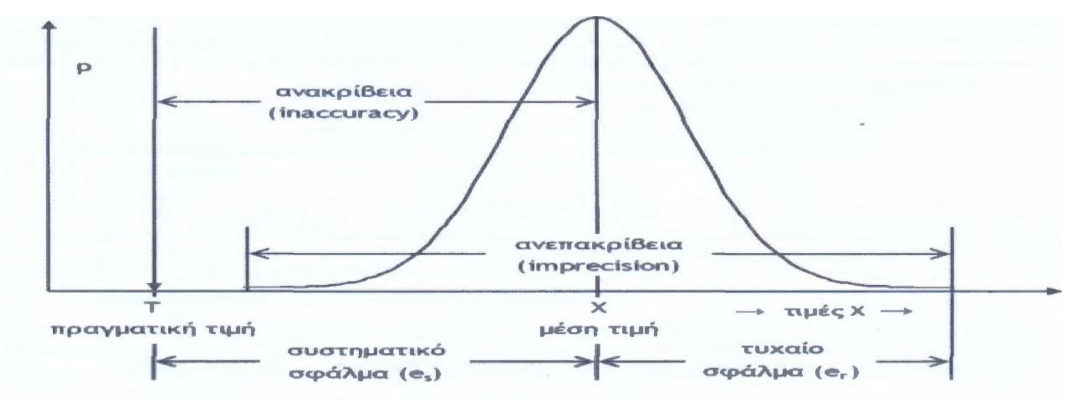
6

$$X_1 = T + e = T + e_s + e_r = T + TX + XX_1$$

- Ο λόγος $T / (T + e_s + e_r)$ αφορά την **εγκυρότητα (validity)**, δηλαδή την καταλληλόλητα του οργάνου μέτρησης. **Όσο πιο μικρό το σφάλμα** (συστηματικό + τυχαίο) **τόσο μεγαλύτερη η εγκυρότητα** (π.χ., 0.7, 0.8).
- Ο λόγος $(T + e_s) / (T + e_s + e_r)$ αφορά την **αξιοπιστία (reliability)**, δηλαδή τη δυνατότητα της οργάνου να αναπαράγει την ίδια τιμή. **Όσο πιο μικρό το τυχαίο σφάλμα** **τόσο μεγαλύτερη η αξιοπιστία** (π.χ., 0.8, 0.9).

7

7



- Αν ένα μέγεθος με πραγματική τιμή T υποβληθεί σε ένα μεγάλο αριθμό επαναληπτικών μετρήσεων (n), τότε το **συστηματικό σφάλμα** (αν υπάρχει) παραμένει σταθερό (e_s), ενώ το **τυχαίο σφάλμα** (e_r) σχηματίζει μια συμμετρική κατανομή (probability density, p) με κεντρική τη μέση τιμή (X) των n μετρήσεων.
- Έτσι, το τυχαίο σφάλμα "εκμηδενίζεται” στατιστικά μέσω της παραδοχής (assumption) της τυχαίας διασποράς των τιμών X ως προς την T με αθροιστικό αποτέλεσμα $\Sigma e_r \rightarrow 0$, όπου e_r η τυχαία απόκλιση (σφάλμα) κάθε επαναληπτικής τιμής X από την πραγματική τιμή T.

8

8

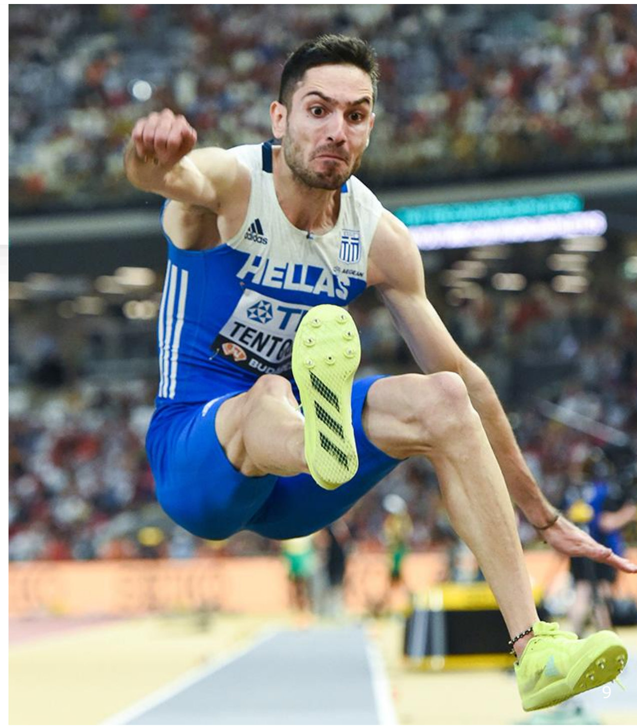
Ανάλυση σφάλματος μέτρησης: *άλμα σε μήκος*

Υποθέστε λοιπόν ότι ένας άλτης κάνει 4 διαδοχικά άλματα και ότι οι **πραγματικές τιμές** των επιδόσεων του είναι:

6.825 m, 7.045 m, 6.585 m και 7.065 m,

ενώ οι **αντίστοιχες τιμές που δίνει ο κριτής** μετρώντας τις επιδόσεις με μια μετροταινία είναι:

6.85 m, 7.06 m, 6.61 m και 7.08 m.



9

- Η μικρότερη μονάδα της μετροταινίας είναι 1 cm και κατά συνέπεια το μέγιστο τυχαίο σφάλμα που υπεισέρχεται σε κάθε μέτρηση **είναι 0.5 cm ή 0.005 m**. Στις μετρήσεις υπεισήλθε ένα **συστηματικό σφάλμα +2 cm** και οι 4 επιδόσεις πρέπει να παρασταθούν με τις αριθμητικές εκφράσεις:

6.825 m + 0.02 m + 0.005 m = 6.85 m

7.045 m + 0.02 m - 0.005 m = 7.06 m

6.585 m + 0.02 m + 0.005 m = 6.61 m

7.065 m + 0.02 m - 0,005 m = 7.08 m

10

10

ΤΙΜΗ		ΣΦΑΛΜΑ			
Πραγματική (T)	Μέτρησης (X)	Ολικό (e=X-T)	Συστηματικό (e_s)	Τυχαίο (e_r)	Σχετικό (%e)
6.825 m	6.85 m	+ 25 mm	+ 20 mm	+ 5 mm	0.37%
7.045 m	7.06 m	+ 15 mm	+ 20 mm	- 5 mm	0.21%
6.585 m	6.61 m	+ 25 mm	+ 20 mm	+ 5 mm	0.38%
7.065 m	7.08 m	+ 15 mm	+ 20 mm	- 5 mm	0.21%

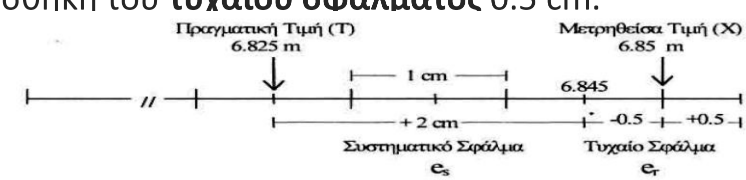
- Το **τυχαίο σφάλμα** εξηγείται με τον εξής τρόπο. Μεταξύ όλων των ανά δύο διαδοχικών ολόκληρων μονάδων (cm) δεν υπάρχει μικρότερη κλιμάκωση ένδειξης (π.χ., της τάξης του 0.5 cm ή 0.2 cm).
- Έτσι, η τιμή μέτρησης κάθε επίδοσης αναφέρεται αναγκαστικά κατά προσέγγιση προς την πλησιέστερη ένδειξη της μετροταινίας, δηλαδή το 1 cm. Οτιδήποτε συμβαίνει μεταξύ δύο ολόκληρων cm χάνεται με ένα σφάλμα, που, όπως λογαριάσαμε παραπάνω, είναι της τάξης του 0.5 cm.
- Το σφάλμα αυτό είναι πότε +0.005 m και πότε -0.005 m, ενώ μετά από ένα μεγάλο αριθμό μετρήσεων μπορεί να εκμηδενιστεί αθροιστικά.

11

11

ΤΙΜΗ		ΣΦΑΛΜΑ			
Πραγματική (T)	Μέτρησης (X)	Ολικό (e=X-T)	Συστηματικό (e_s)	Τυχαίο (e_r)	Σχετικό (%e)
6.825 m	6.85 m	+ 25 mm	+ 20 mm	+ 5 mm	0.37%
7.045 m	7.06 m	+ 15 mm	+ 20 mm	- 5 mm	0.21%
6.585 m	6.61 m	+ 25 mm	+ 20 mm	+ 5 mm	0.38%
7.065 m	7.08 m	+ 15 mm	+ 20 mm	- 5 mm	0.21%

- Το **συστηματικό σφάλμα** προκύπτει από την ανακρίβεια του οργάνου μέτρησης, που στην περίπτωση της μετροταινίας ήταν 0.02 m και φυσικά εισήγαγε ισόποσο σφάλμα +2 cm σε όλες τις μετρήσεις.
- Έτσι, η **πραγματική τιμή 6.825** m έγινε 6.845 m με την προσθήκη των 2 cm και 6.85 m με την προσθήκη του **τυχαίου σφάλματος** 0.5 cm.



Μετρηθείσα Τιμή = Πραγματική Τιμή + Συστηματικό Σφάλμα + Τυχαίο Σφάλμα :

$$X = T + e_s + e_r = 6.825 + 0.02 + 0.005 = 6.85$$

12

ΤΙΜΗ		ΣΦΑΛΜΑ			
Πραγματική (T)	Μέτρησης (X)	Ολικό (e=X-T)	Συστηματικό (e_s)	Τυχαίο (e_r)	Σχετικό (%e)
6.825 m	6.85 m	+ 25 mm	+ 20 mm	+ 5 mm	0.37%
7.045 m	7.06 m	+ 15 mm	+ 20 mm	- 5 mm	0.21%
6.585 m	6.61 m	+ 25 mm	+ 20 mm	+ 5 mm	0.38%
7.065 m	7.08 m	+ 15 mm	+ 20 mm	- 5 mm	0.21%

- Όταν **δεν γνωρίζουμε την πραγματική τιμή** της μετρούμενης ποσότητας, τότε το **σχετικό σφάλμα** είναι:

  Σχετικό Σφάλμα = (Ολικό Σφάλμα)/(Τιμή Μέτρησης) X 100

  όπου απόλυτο σφάλμα είναι το σφάλμα που μπορεί να εκτιμηθεί συνολικά μετά από σχετική ανάλυση σφάλματος.
- Όταν **γνωρίζουμε την πραγματική τιμή** της μετρούμενης ποσότητας, τότε το **σχετικό σφάλμα** είναι:

  Σχετικό Σφάλμα = (Ολικό Σφάλμα)/(Πραγματική Τιμή) X 100

13

13

Παράδειγμα 1.2 - Σχετικό σφάλμα: *πάχος μηριαίου οστού.*

- Από ακτινογραφία της κεφαλής του δεξιού μηρού ενός αθλητή μετράμε το διακονδυλικό πάχος και το βρίσκουμε 11.85 cm (**πραγματική τιμή, T**).
- Με διαστημόμετρο επακρίβειας 1 mm μετράμε πάνω στον αθλητή (εξωτερικά) το πάχος αυτό και το βρίσκουμε 12.9 cm (**τιμή μέτρησης, X**). Ας δεχτούμε ότι η μέτρηση με το διαστημόμετρο δεν περιείχε σφάλμα χειρισμού και εφαρμογής της διαδικασίας.
- Το **ολικό σφάλμα είναι e** = X - T = 12.9 -11.85 = 1.05 cm (1 cm συστηματικό και 0.05 cm τυχαίο) και δείχνει ότι η τιμή μέτρησης του διακονδυλικού πάχους "πέφτει" κάπου μεταξύ 12.85 cm και 12.95 cm, δηλαδή μέσα **στα όρια σφάλματος** που προβλέπει η κλίμακα του χιλιοστόμετρου και σε απόσταση 1 cm από την πραγματική τιμή (T).

14

14

- Το **σχετικό σφάλμα** στην περίπτωση αυτή είναι:
- **1) με (υποθετικά) άγνωστη την πραγματική τιμή Τ**

(e / X) * 100 = (1.05 cm / 12.9 cm) * 100 = (0.081 * 100) = 8.1 %

δηλαδή 8.1 % της μέτρησης (Χ) του διακονδυλικού πάχους, και

- **2) με γνωστή την πραγματική τιμή Τ**

(e / T) * 100 = (1.05 cm / 11.85 cm) * 100 = (0.089 * 100) = 8.9 %

δηλαδή 8.9% της πραγματικής τιμής (Τ) του διακονδυλικού πάχους.

Το ποσοστό αυτό δείχνει ότι το διαστημόμετρο αυτό δεν είναι ακριβές και θα πρέπει να βαθμονομηθεί, ώστε να εκμηδενιστεί **το συστηματικό σφάλμα** (1 cm) και να παραμείνει μόνο το τυχαίο σφάλμα (+ 0.05 cm), που είναι αναπόφευκτο και βελτιώνεται μόνο με άλλο καλύτερης ευαισθησίας.

15

15

Αριθμητικά Όρια

- Τα **αριθμητικά όρια (numerical limits)** ενός αριθμού π.χ. Χ είναι Χ ± e_r. Κάθε αριθμός Χ αποτελεί τη μεσαία τιμή ενός διαστήματος τιμών με μικρότερη την Χ - e_r και μεγαλύτερη τιμή την Χ + e_r, δηλαδή από μισή μονάδα μέτρησης πριν μέχρι μισή μονάδα μέτρησης μετά την Χ.
- Για παράδειγμα, η τιμή Χ = 1.2 cm έχει μικρότερη μονάδα μέτρησης το 0.1 cm, τυχαίο σφάλμα το μισό τού 0.1 cm, δηλαδή 0.05 cm, και **αριθμητικά όρια** 1.2 ± 0.05: κατώτερο όριο 1.15 cm και ανώτερο όριο 1.25 cm.
- Ομοίως, η τιμή Υ = 35.4 cm έχει μικρότερη μονάδα μέτρησης (επίσης) το 0.1 cm, τυχαίο σφάλμα το μισό τού 0.1 cm, δηλαδή 0.05 cm, και **αριθμητικά όρια** 35.4 ± 0.05: κατώτερο όριο 35.35 cm και ανώτερο όριο 35.45 cm.

16

16

- Έστω ότι θέλουμε **να προσθέσουμε** τα μήκη X = 1.2 cm και Y = 35.4 cm.
- Ποια είναι τα **αριθμητικά όρια του αθροίσματος** αυτού;

- Η τιμή X = 1.2 cm έχει κατώτερο όριο 1.15 cm και ανώτερο 1.25 cm.
- Η τιμή Y = 35.4 cm έχει κατώτερο όριο 35.35 cm και ανώτερο 35.45 cm.
- → ελάχιστο άθροισμα = 1.15 + 35.35 = 36.50 cm

  μέγιστο άθροισμα = 1.25 + 35.45 = 36.70 cm

- Έτσι, το άθροισμα X + Y = 1.2 + 35.4 = 36.6 cm έχει μέγιστο ολικό σφάλμα 0.10 cm (άθροισμα σφαλμάτων των 2 μετρήσεων: 0.05 + 0.05 cm) και αριθμητικά όρια 36.50 cm - 36.70 cm.

Για την επόμενη διάλεξη Κατανομές Συχνότητας