

Στατιστική ανάλυση δεδομένων

Γεωργία Σαλαντή

Λέκτορας

Εργαστήριο Υγιεινής και Επιδημιολογίας

Ιατρική Σχολή

Πανεπιστήμιο Ιωαννίνων

Στατιστική

1. Εκτιμήσεις

- Μεγέθη συσχέτισης και διαστήματα εμπιστοσύνης

2. Έλεγχοι υποθέσεων

- Τεστ, P-values

1. Εκτιμήσεις και αβεβαιότητα

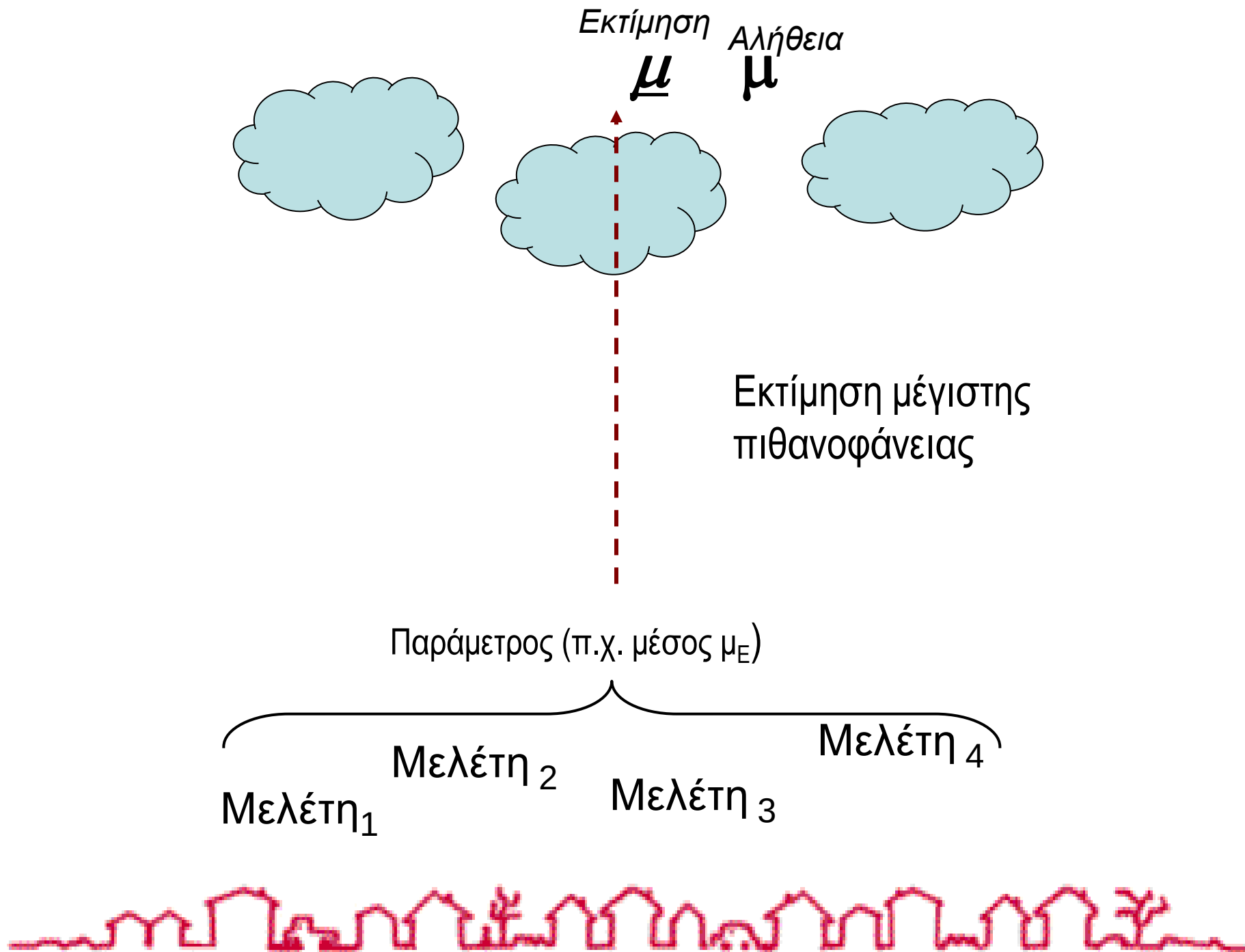
- Συνεχή δεδομένα
 - Διαφορά μέσων
- Διχότομα δεδομένα
 - Λόγοι κινδύνων, λόγοι αναλογιών

Project

- Είναι οι Έλληνες πιο ψηλοί από τους Σουηδούς;
- Υπολογίστε το μέσο ύψος για τους Έλληνες μ_E
- Υπολογίστε το μέσο ύψος για τους Σουηδούς μ_S
- Συγκρίνετε μ_S και μ_E

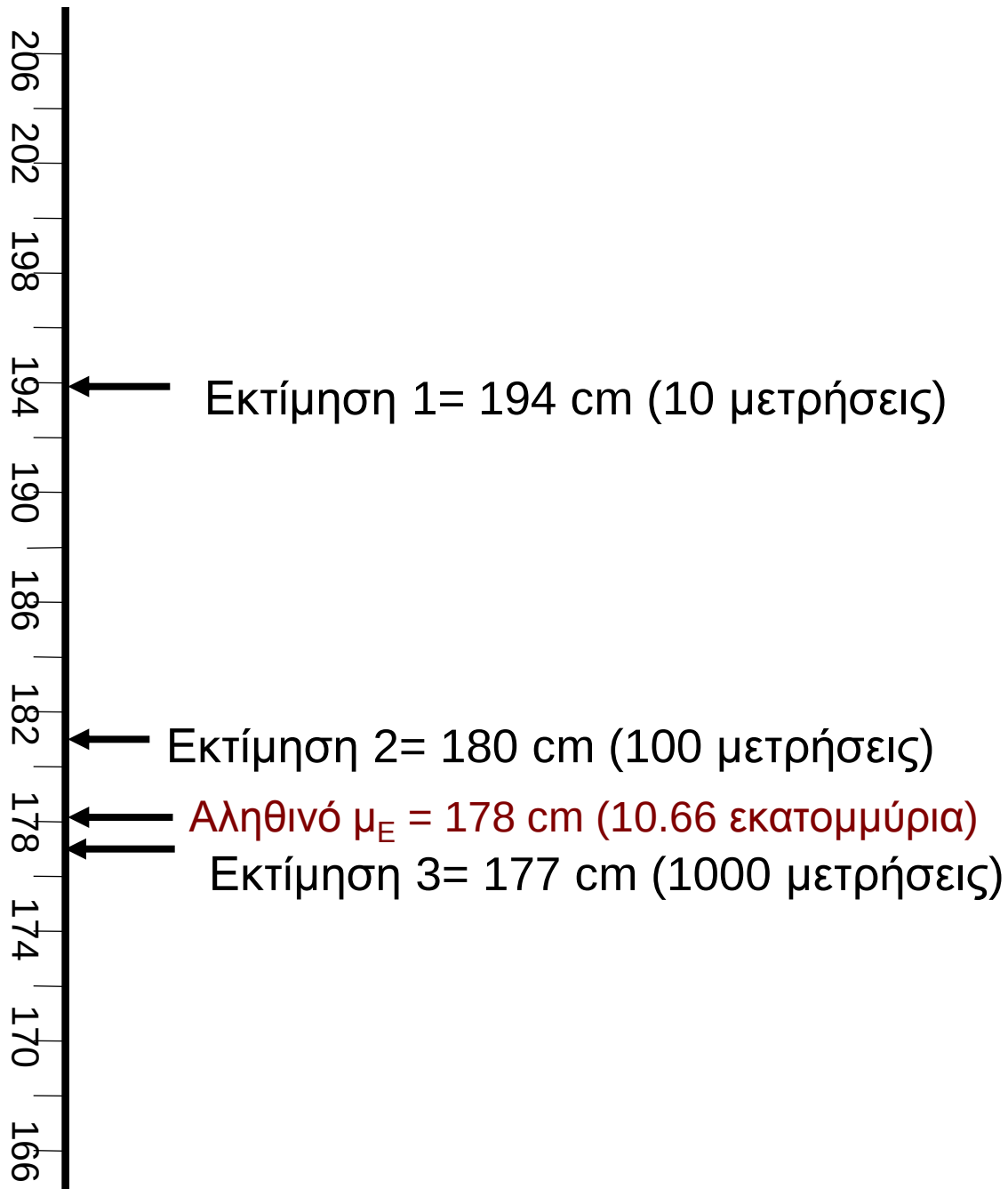
Η στατιστική βασίζεται στην εκτίμηση

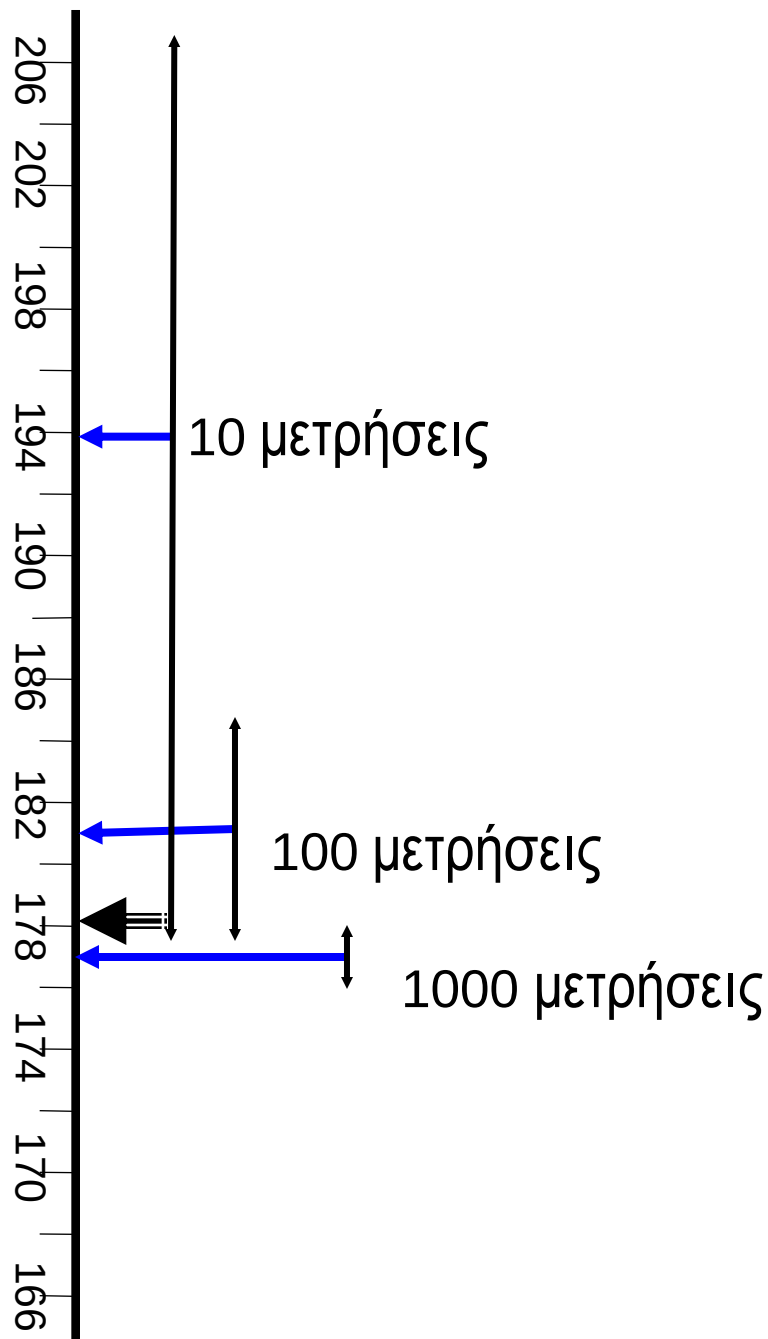
- Εκτίμηση του μέσου ύψους
 - ‘αληθινό’ μέσο ύψους του πληθυσμού
 - Εκτίμηση (από ένα δείγμα)



Εκτίμηση

- Πάρτε ένα δείγμα από 10 Έλληνες και μετρήστε το μέσο όρο τού ύψους τους
- Φυσικά, αυτό δεν είναι το μέσο ύψος όλων των Ελλήνων αλλά ελπίζουμε να είναι κάτι πολύ κοντά του
- $\underline{\mu}_E = 194 \text{ cm!!}$





Όσο πιο πολλούς
ανθρώπους μετράμε, τόσο
βεβαιότεροι είμαστε για
την εκτίμηση του
‘αληθινού’ ύψος του
πληθυσμού

Ο βαθμός της
βεβαιότητας μας
εκφράζεται στο
διάστημα εμπιστοσύνης
 $\Delta\epsilon$

Confidence Interval

Διαστήματα εμπιστοσύνης

Μέτρα διασποράς

- Τυπικό/προτυπομένο σφάλμα
 - standard error =SE
- Τυπική απόκλιση=τ.α.
 - standard deviation=sd
- Διασπορά=variance

Ο χρυσός κανόνας για διαστήματα εμπιστοσύνης

- Για μεγέθη που ακολουθούν την κανονική κατανομή, το 95% διάστημα εμπιστοσύνης (95% Confidence Interval) για το *κατι*

$$\text{κατι} \pm 1.96 \cdot \tau.σ.(\text{κατι})$$

Παράδειγμα

- Μέτρηση 100 νεογέννητων μωρών από μητέρες που κάπνιζαν. Μέσο βάρος=2.7kgr και τ.α=1.1
- Μέτρηση 90 νεογέννητων μωρών από μητέρες που δεν κάπνιζαν. Μέσο βάρος=3.1kgr και τ.α=1.2
- Σχετίζεται το κάπνισμα με χαμηλότερο βάρος;

Παράδειγμα : διαστήματα εμπιστοσύνης

- Καπνιστές: $\tau.σ.=1.1/10=0.11$
95% ΔΕ: $(2.7-1.96 \cdot 0.11, 2.7+1.96 \cdot 0.11)$
95% ΔΕ: (2.48, 2.92)
- Μη καπνιστές: $\tau.σ.=1.2/9.49=0.13$
95% ΔΕ: $(3.1-1.96 \cdot 0.13, 3.1+1.96 \cdot 0.13)$
95% ΔΕ: (2.85, 3.35)
- Δεν υπάρχει διαφορά

Διάστημα εμπιστοσύνης

Εκφράζει διακύμανση και μέγεθος δείγματος

= το πιθανό διάστημα της 'αλήθειας' (μ_E)

-Πόσο πιθανό;

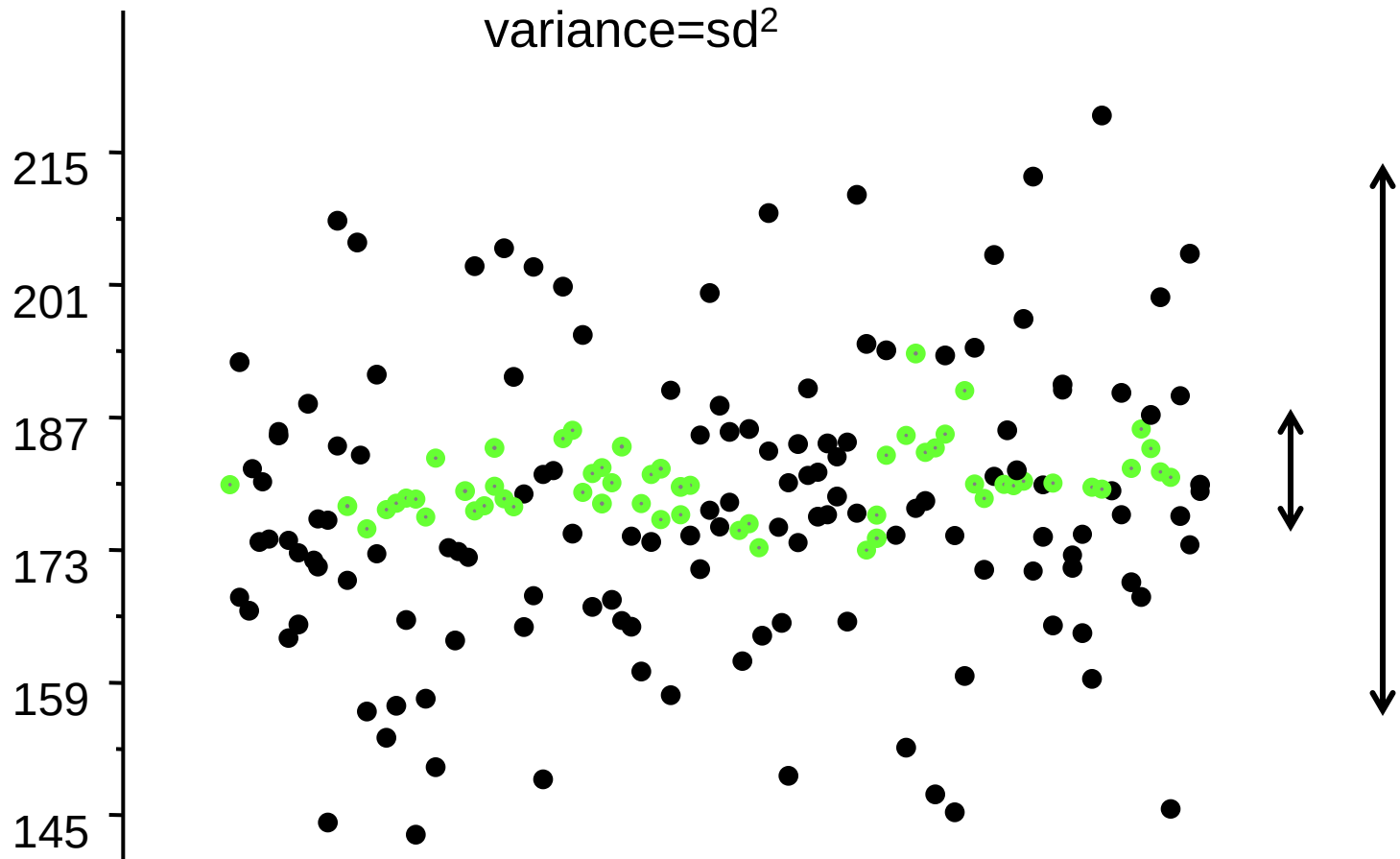
-95% πιθανό

Το δεδομένο διάστημα έχει πιθανότητα 95% να περιέχει την 'αλήθεια'

Εμπιστοσύνη και διακύμανση

Πόσο διασκορπισμένες είναι οι παρατηρήσεις?

Διακύμανση (variance), τυπική απόκλιση (standard deviation)

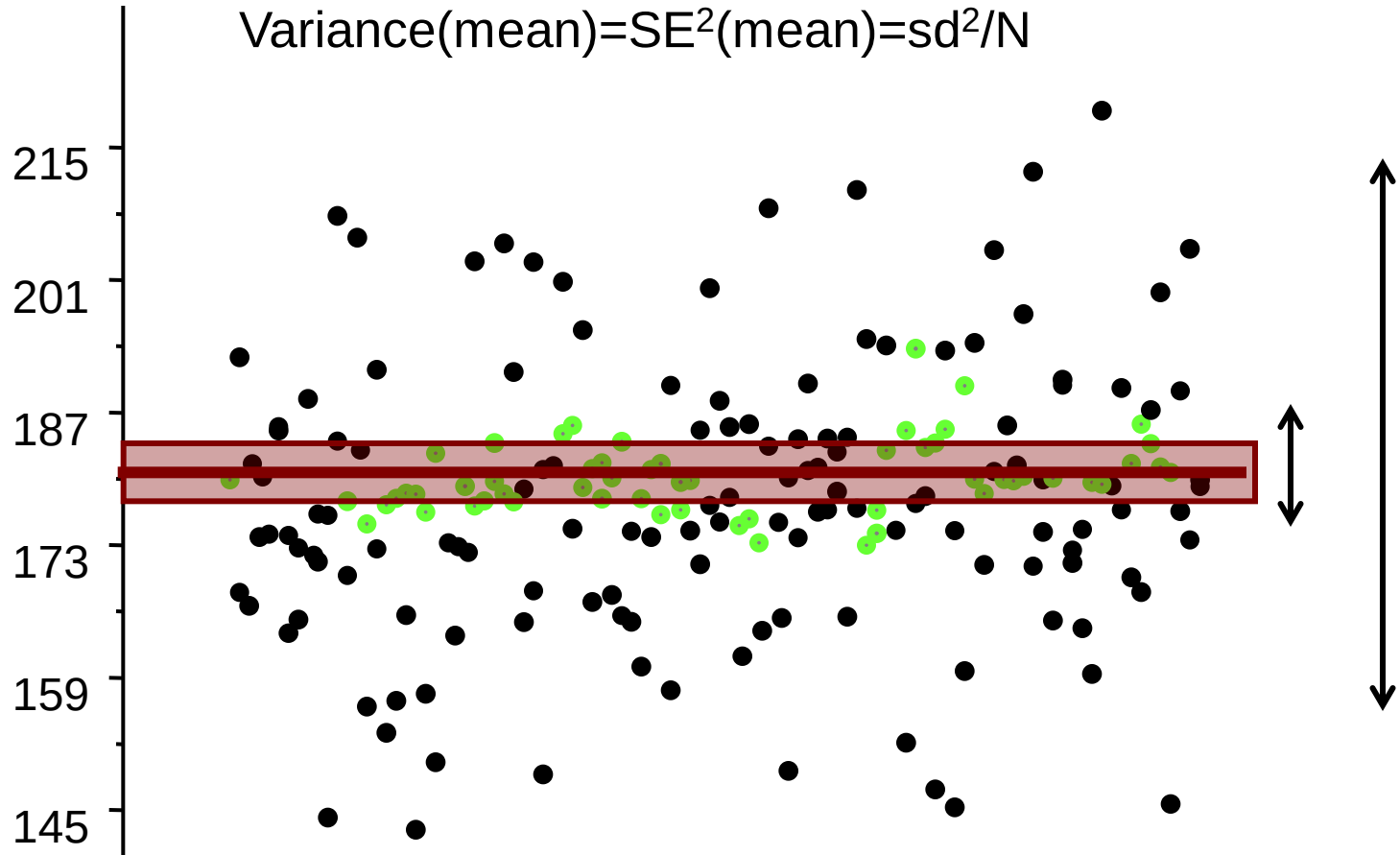


Εμπιστοσύνη και διακύμανση

Πόσο αβέβαιος είναι ο μέσος;

Τυπικό σφάλμα μέσου (Standard error of the mean SE), διασπορά του μέσου (variance of the mean)

$$\text{Variance}(\text{mean}) = \text{SE}^2(\text{mean}) = \text{sd}^2 / N$$



SOS

- Για εκτιμήσεις
 - $SE^2 = \text{variance}$
 - Για μέσους ισχύει $SE^2 = sd^2/N$, N το μέγεθος του δείγματος

Πάμε πίσω στο Project

Μετά από 1000 μετρήσεις:

- Έλληνες: $\mu_E = 177\text{cm}$, 95% $\Delta E = [176, 178]$
- Σουηδοί: $\mu_S = 183\text{cm}$, 95% $\Delta E = [182, 184]$
- Ποιοι είναι πιο ψηλοί;

Διάστημα εμπιστοσύνης (184cm, 186cm)

- ~~95%~~
(184

- Ο μ
(184

- Κατ
έχει



Two 'average' men having an 'average' meal.

ξύ

μός

Διάστημα εμπιστοσύνης και μέγεθος δείγματος

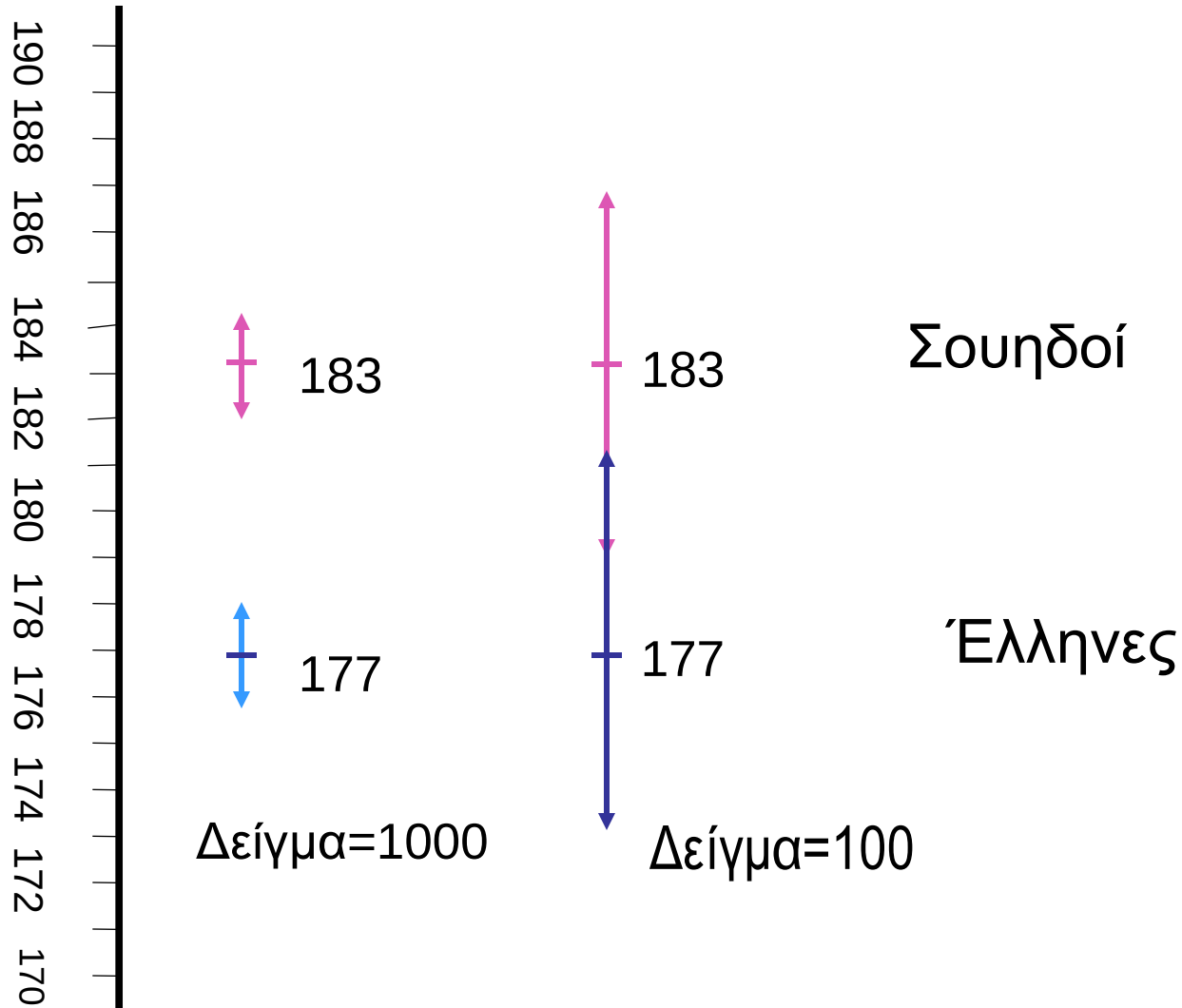
Διάστημα εμπιστοσύνης και εμπιστοσύνη					
<u>μέσος</u>	<u>σ.α.</u>	<u>δείγμα</u>	<u>εμπιστοσύνη</u>	<u>Δ.Ε.</u>	
180	20	100	99%	175	185
180	20	100	95%	176	184
180	20	100	90%	177	183
Διάστημα εμπιστοσύνης και μέγεθος δείγματος					
<u>μέσος</u>	<u>σ.α.</u>	<u>δείγμα</u>	<u>εμπιστοσύνη</u>	<u>Δ.Ε.</u>	
180	20	10	95%	168	192
180	20	100	95%	176	184
180	20	1000	95%	179	181

Πάμε πίσω στο Project

Μετά από 1000 μετρήσεις:

- Έλληνες: $\mu_E = 177\text{cm}$, 95% $\Delta E = [176, 178]$
- Σουηδοί: $\mu_S = 183\text{cm}$, 95% $\Delta E = [182, 184]$
- Ποιοι είναι πιο ψηλοί;

Εξέταση υπόθεσης με το διάστημα εμπιστοσύνης



Μέτρα συσχέτισης για διχότομα δεδομένα

- Borchgrevink 1966

Εγκεφαλικό	Ναι	Όχι	Σύνολο
Omega 3	<i>10</i>	<i>90</i>	<i>100</i>
Τίποτα	<i>14</i>	<i>86</i>	<i>100</i>
	<i>24</i>	<i>186</i>	<i>200</i>

Κίνδυνος

- 20 φοιτητές πίνουν ρούμι πριν τις εξετάσεις, 5 κόβονται
- **Κίνδυνος**
 - = 5 κόβονται/20 όλοι
 - = $5/20 = 1/4 = 0.25 = 25\%$

$$\text{Κίνδυνος} = \frac{\# \text{ γεγονότων}}{\# \text{ παρατηρήσεων}}$$

- Η πιθανότητα να κοπείς είναι 25% αν πιεις ρούμι την προηγούμενη νύχτα

Αναλογία

- 20 φοιτητές πίνουν ρούμι πριν τις εξετάσεις, 5 κόβονται
- **Αναλογία**

= 5 κόβονται/15 περνάνε

= $5/15 = 1/3 = 0.33$

$$\text{Αναλογία} = \frac{\# \text{ γεγονότων}}{\# \text{ παρατηρήσεων} - \# \text{ γεγονότων}}$$

- Η πιθανότητα να κοπείς όταν πίνεις ρούμι είναι ένα τρίτο της πιθανότητας να περάσεις
- Ένας θα κοπεί για κάθε τρεις που θα περνάνε
- Οι πιθανότητες είναι τρία προς ένα κατά

Λόγος κινδύνων και λόγος αναλογιών

$$\Lambda K = \frac{\# \text{ Κίνδυνος Ομάδα 1}}{\# \text{ Κίνδυνος Ομάδα 2}} \quad \textbf{Risk Ratio (RR)}$$

$$\Lambda A = \frac{\# \text{ Αναλογία Ομάδα 1}}{\# \text{ Αναλογία Ομάδα 2}} \quad \textbf{Odds Ratio (OR)}$$

	Γεγονός	Όχι γεγονός	Συνολο
Νέα θεραπεία	a	b	$a+b$
Παλιά θεραπεία	c	d	$c+d$
	$a+c$	$b+d$	n

$$\Lambda K = \frac{a (c+d)}{c (a+b)}$$

$$\Lambda A = \frac{a d}{c b}$$

Ερμηνεία λόγων

- $\Lambda A = 1$ ή $\Lambda K = 1$ σημαίνει ότι η έκθεση δεν σχετίζεται με τη νόσο
- Ο κίνδυνος είναι ο ίδιος στις δύο ομάδες
- $\Lambda K = 2$: Ο κίνδυνος (πιθανότητα της νόσου) στην ομάδα A είναι διπλάσιος σε σχέση με την ομάδα B
- $\Lambda A = 2$: Η αναλογία στην ομάδα A είναι διπλάσια σε σχέση με την ομάδα B

Παράδειγμα: Borchgrevink 1966

$$\Lambda_K = \frac{\frac{10}{14}}{\frac{100}{100}} = 0.71 = 71\%$$

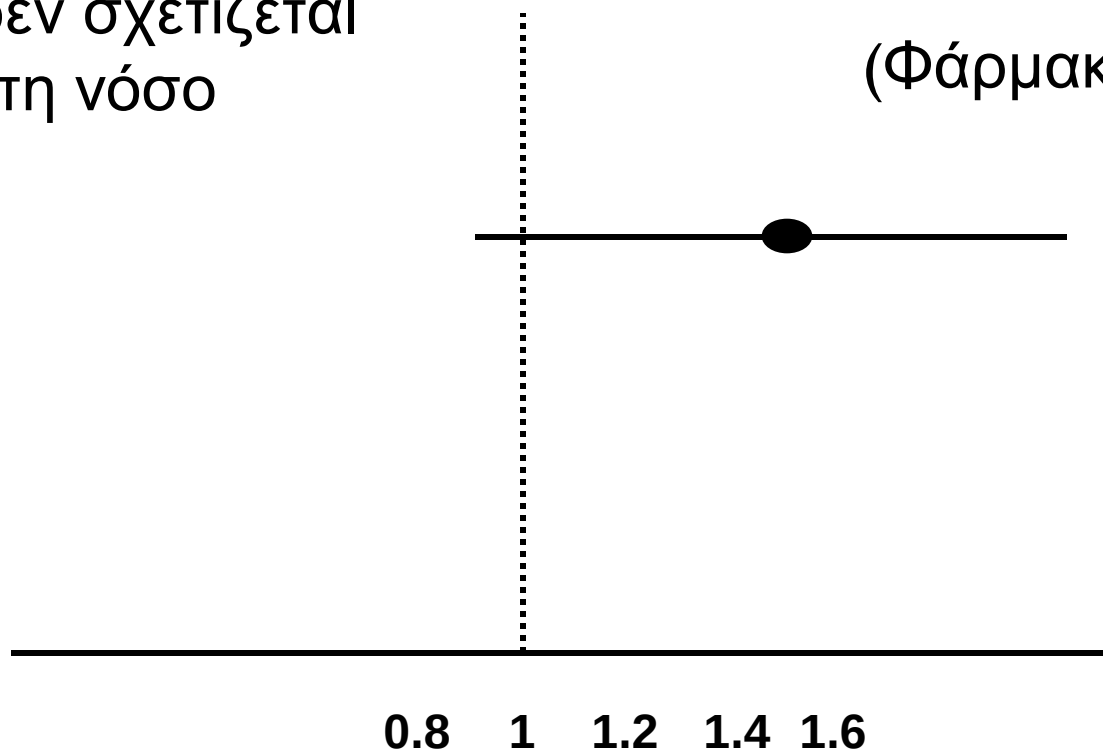
$$\Lambda_A = \frac{\frac{10}{14}}{\frac{90}{86}} = 0.68$$

Εγκεφαλικό	Ναι	Όχι	Σύνολο
Omega 3	10	90	100
Τίποτα	14	86	100
	24	186	200

Λόγος αναλογιών και το ΔΕ

$\Lambda A = 1$ σημαίνει ότι η
έκθεση δεν σχετίζεται
με τη νόσο

ΛA της νόσου
(Φάρμακο ναι/όχι)



Βελτίωση

2. Έλεγχοι υποθέσεων

- Θέλουμε να συγκρίνουμε δύο δείγματα
 - Συνεχή δεδομένα
 - πχ Μέση αρτηριακή πίεση σε δύο ομάδες
 - Διχότομα δεδομένα
 - Πιθανότητα θανάτου με δύο διαφορετικά είδη αντικαταθλιπτικών

Μηδενική υπόθεση

- Μία υπόθεση που θέλουμε να ελέγξουμε
- Αναφέρεται σε κάποια χαρακτηριστικά των δειγμάτων που εξετάζουμε εκφρασμένα σε στατιστικές ποσότητες
- $H_0: \bar{X}_1 = \bar{X}_2$ (η μέση αρτηριακή πίεση είναι ίδια σε δύο ομάδες)
- $H_1: \bar{X}_1 > \bar{X}_2$
- $H_1: \bar{X}_1 < \bar{X}_2$
- $H_1: \bar{X}_1 > \bar{X}_2$ ή $\bar{X}_1 < \bar{X}_2$

Ποιο τέστ θα χρησιμοποιήσω;

- Συνεχή ή διχότομα δεδομένα;
 - Διχότομα: X-squared test, McNeamar test
 - Αν είναι συνεχή, ακολουθούν κανονική κατανομή ή όχι;
 - Για κανονική κατανομή: z-test, t-test
 - Για μη κανονική κατανομή: Wilcoxon, Mann-Witney test
- Σε κάθε περίπτωση, το αποτέλεσμα ενός ελέγχου δίνει μια p-value

Εξέταση υπόθεσης H_0 με έλεγχο

- H_0 : 'Οι Σουηδοί δεν είναι πιο ψηλοί'
- H_0 : Όσοι παίρνουν ταμπλέτες με Omega 3 **δεν έχουν λιγότερο κίνδυνο** για καρδιαγγειακά σε σχέση με αυτούς που δεν παίρνουν'
- H_0 : 'κίνδυνος στην ομάδα 1 = κίνδυνος στην ομάδα 2'
- H_0 : $\Lambda A=1$ ή $\Lambda K=1$ ή $\Delta K=0$



Έλεγχος (test) 

P-value

Παράδειγμα

- Χοληστερίνη με το φάρμακο $\mu_{\phi}=130$,
Χοληστερίνη χωρίς το φάρμακο $\mu_{\chi\phi}=136$
- Διαφορά=6
- Διαφορά πολύ κοντά στο μηδέν, έτσι θα μπορούσαμε να πούμε ότι είναι περίπου ίσα, ε;
- H_0 : Χοληστερίνη με το φάρμακο = Χοληστερίνη χωρίς το φάρμακο
- Ο έλεγχος (t-test) δίνει p-value=1%

P-value

- Εάν επαναλάβω αυτή τη μέτρηση σε άλλα 100 άτομα θα βρω το ίδιο;
- Εάν το επαναλάβω πολλές φορές (πχ 1000 φορές) σε δείγματα του ιδίου μεγέθους, πόσες φορές θα βρω κάτι τόσο 'ακραίο' όπως διαφορά=6;
- $\pi/1000$ φορές = p-value

P-value

- Εάν η τιμή από το δείγμα δεν είναι στο **5% των πιο ακραίων τιμών**, τότε το εύρημα θεωρείται 'στατιστικά μη σημαντικό'
- Δεν απορρίπτουμε την H_0
- Εάν είναι στο 5%, τότε απορρίπτουμε την H_0
- $P\text{-value} < 0.05$: απορρίπτουμε την H_0 , το εύρημα που έχουμε (δηλ. μια μέση διαφορά στην χολιστερίνη=6) δεν μπορεί να εξηγηθεί από την τύχη

Παράδειγμα

- Χοληστερίνη με το φάρμακο $\mu_{\varphi}=130$,
Χοληστερίνη χωρίς το φάρμακο $\mu_{\chi\varphi}=136$
- **Διαφορά=6**
- Διαφορά πολύ κοντά στο μηδέν, έτσι θα μπορούσαμε να πούμε ότι είναι περίπου ίσα, ε;
- H_0 : Χοληστερίνη με το φάρμακο = Χοληστερίνη χωρίς το φάρμακο
- Ο έλεγχος (t-test) δίνει p-value=1%
 - Η πιθανότητα η διαφορά που βλέπω (**το 6**) να είναι *τυχαία* είναι 1%
 - Είναι **μικρή ή μεγάλη** αυτή η πιθανότητα;
 - Χρησιμοποιούμε **κατώφλι 5%**

Τι είναι η p-value;

- κατώφλι: $p\text{-value}=5\%$
- $p\text{-value} < 5\%$: στατιστικά σημαντική διαφορά
Η παρατηρηθείσα διαφορά μεταξύ των δύο ομάδων είναι πάρα πολύ μεγάλη για να εξηγηθεί από την τύχη!
- Απορρίπτουμε την H_0
- $p\text{-value} > 5\%$: στατιστικά μη σημαντική διαφορά

Αμφίπλευρα και μονόπλευρα ΤΕΣΤ

- $H_0: x_1 = x_2$ με $H_1: x_1 > x_2$ ή $x_1 < x_2$ (αμφίπλευρο)
p-value=0.09
- $H_0: x_1 = x_2$ με $H_1: x_1 > x_2$ (μονόπλευρο)
p-value=0.045
- Τα αμφίπλευρα τεστ είναι **πιο 'συντηρητικά'**
(= δεν απορρίπτουν τόσο εύκολα τη μηδενική υπόθεση)

Ισχύς ενός τεστ

- Η 'δύναμη' που έχει ένα τεστ να απορρίπτει τη μηδενική υπόθεση (=να δίνει $p\text{-value} < 0.05$) όταν αυτή δεν ισχύει

		Πραγματικότητα	
ΤΕΣΤ		H_0 σωστή	H_0 λάθος
	H_0 απορρίπτεται	Σφάλμα τύπου I =0.05	Ισχύς
	H_0 δεν απορρίπτεται		Σφάλμα τύπου II

Ισχύς

- Εξαρτάται από **το μέγεθος του δείγματος**
 - Όσο πιο μεγάλο, τόσο πιο μεγάλη η ισχύς
- Εξαρτάται από **τη διαφορά** ανάμεσα στις δύο ομάδες
 - Όσο πιο μεγάλη η διαφορά, τόσο πιο μεγάλη η ισχύς

Υπολογισμός μεγέθους δείγματος

- Πρέπει από πριν να καθορίσουμε
 - Το τεστ που θα χρησιμοποιήσουμε
 - Την προσδοκώμενη διαφορά ανάμεσα στις δύο ομάδες
 - Την ισχύ που θέλουμε ($X\%$, πχ 80% πιθανότητα αν η H_0 είναι λάθος το τεστ να μπορέσει να την απορρίψει)
- Έπειτα υπολογίζουμε το απαιτούμενο μέγεθος δείγματος για να έχουμε ισχύ $X\%$

Προσέξτε τις p -values...

- Ευαίσθητες στο μέγεθος δείγματος!
- Προβλήματα με υπερβολική ισχύ και έλλειψη ισχύος!
 - Μικρό δείγμα: λίγα δεδομένα για να αποφασίσουμε και να απορρίψουμε την υπόθεση
 - Πολύ μεγάλο δείγμα: κάνει ακόμα και μικρές διαφορές να φαίνονται τεράστιες!!
- Τις πιο πολλές φορές τα ΔΕ δίνουν πιο πολλές πληροφορίες από τις p -values

Μπορεί κάτι τέτοιο να συμβεί;

206 202 198 194 190 186 182 178 174 170 166

↕ 10 μετρήσεις

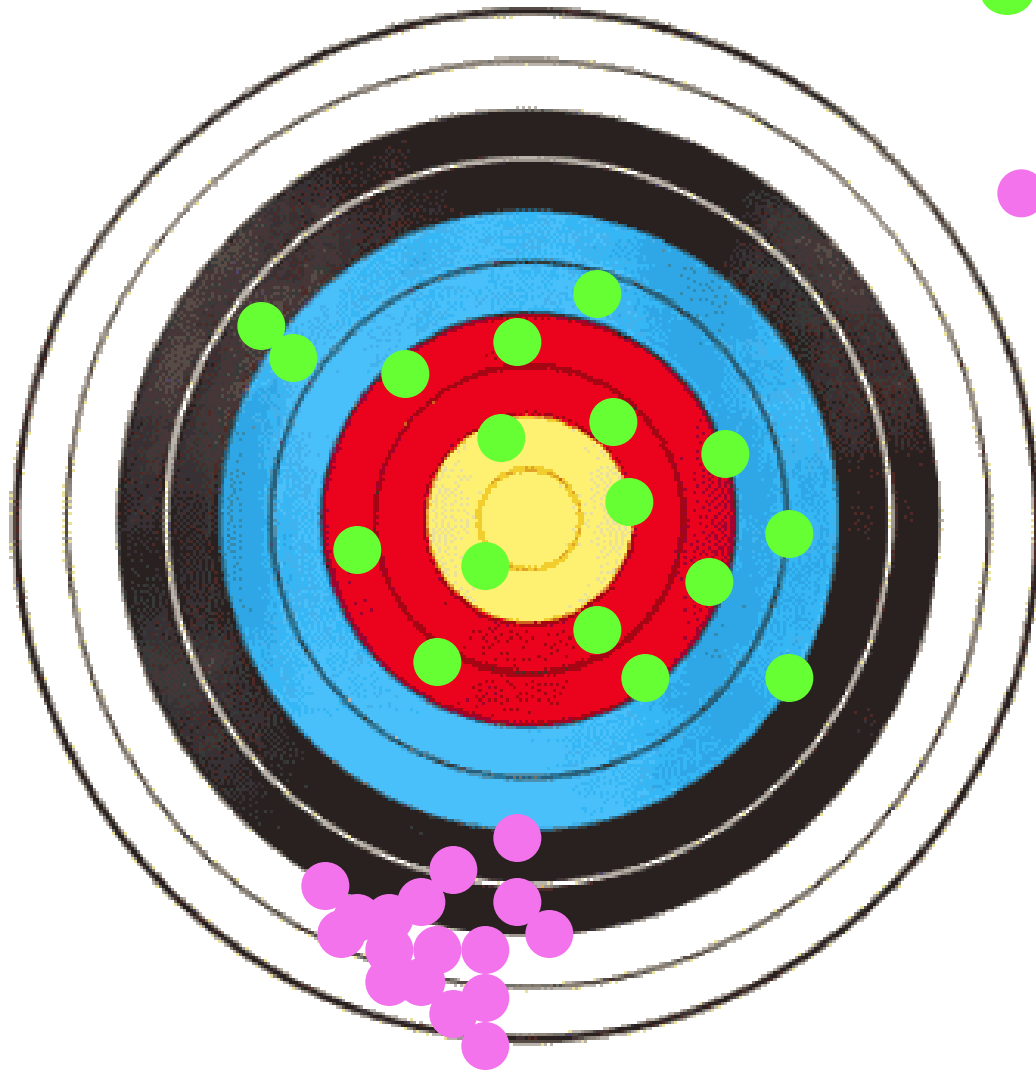
↑



Τα διάστημα εμπιστοσύνης δεν είναι η απάντηση στα πάντα!

- Τα διαστήματα εμπιστοσύνης απεικονίζουν ΜΟΝΟ την αβεβαιότητα - καμία πληροφορία για προκατάληψη και συστηματικό σφάλμα

Διασπορά και συστηματικό σφάλμα



● Διασπορά αλλά όχι
συστηματικό σφάλμα

● Συστηματικό σφάλμα
αλλά όχι διασπορά

Τι μπορεί να προσφέρει η στατιστική?

- Παρέχει μια **ωραία περίληψη** των αριθμητικών πληροφοριών – διαγράμματα
- **Απαντά υποθέσεις** – π.χ. είναι τα κινητά τηλέφωνα καρκινογόνα? - μέτρο για το πόσο σίγουρο είναι κάποιο συμπέρασμα
- **Συγκρίνει τις πληροφορίες** από τις διαφορετικές ομάδες – π.χ. την επιβίωση σε εκείνους που λαμβάνουν το φάρμακο Α και σε εκείνους που λαμβάνουν το εικονικό φάρμακο
- Δίνει μια **εκτίμηση για το πόσο ακριβής είναι η πρόβλεψη**

Τι είναι που;

(λίγο spoon feeding επειδή είναι αρχή...)

- Από το βιβλίο 'Αρχές αποδεικτικής ιατρικής' κοιτάξτε στις σελίδες
 - Για σχεδιασμούς μελετών κεφαλαία 12, 13, 14 (μόνο τις εισαγωγές!)
 - Σελ. 30-32
 - Σελ. 102-104