1

Βασικοί ορισμοί

- **Πληθυσμός**: Οποιαδήποτε συλλογή οντοτήτων, έμψυχων ή άψυχων, τις οποίες εξετάζουμε ως προς ένα ή περισσότερα χαρακτηριστικά. Αδύνατη συνήθως η μελέτη του στο σύνολό του. Προσεγγίζεται με την επιλογή του κατάλληλου δείγματος.

- **Δείγμα**: Τμήμα του πληθυσμού το οποίο επιλέγουμε για μελέτη. Πρέπει να είναι αντιπροσωπευτικό του πληθυσμού και επαρκές ώστε να επιτρέπει τη γενίκευση συμπερασμάτων στον πληθυσμό.

- Η στατιστική **μεταβλητή** (variable) αποτελεί ένα χαρακτηριστικό του πληθυσμού που μελετάμε, ενώ οι τιμές που λαμβάνει ονομάζονται **δεδομένα**.

2

Αρχές Δειγματοληψίας

- Η δειγματοληψία απαιτεί ακριβή ορισμό του **πληθυσμού αναφοράς** και του **δειγματοληπτικού πλαισίου.**

- Το δείγμα πρέπει να είναι **αντιπροσωπευτικό** και **επαρκές** ώστε να επιτρέπει την <u>γενίκευση</u> των συμπερασμάτων της έρευνας στο σύνολο του πληθυσμού.

- Η **γενίκευση** γίνεται μόνο στον πληθυσμό που καθορίζεται από το δειγματοληπτικό πλαίσιο που έχει χρησιμοποιηθεί.

- Υπάρχει το ενδεχόμενο **λάθους δειγματοληψίας**: Μόνο ένα από τα πολλά πιθανά δείγματα μελετάται.

3

Μέγεθος δείγματος – Γενικές αρχές

Το **επαρκές** μέγεθος του δείγματος είναι <u>μεγίστης</u> σημασίας για την <u>αξιοπιστία</u> της έρευνας:

Μικρό δείγμα → Η μελέτη μπορεί να αποτύχει να απαντήσει το ερώτημα λόγω μειωμένης ισχύος
Μεγάλο δείγμα → Ασύμμετρη και ασύμφορη χρήση πόρων (κόστος, χρόνος κ.λπ.)

Γενικά, <u>χρειαζόμαστε πιο μεγάλο δείγμα</u>:

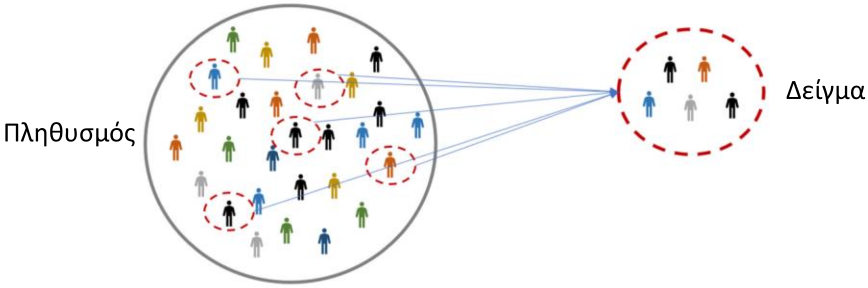
- Εάν υπάρχει ανομοιογένεια στον πληθυσμό.
- Εάν επιθυμούμε πιο μεγάλη ακρίβεια.
- Εάν η δειγματοληψία μας είναι κατά συστάδες ή στρωματοποιημένη.
- Εάν η στατιστική ανάλυση είναι πιο σύνθετη.

4

Συμβολισμός (πληθυσμός & δείγμα)

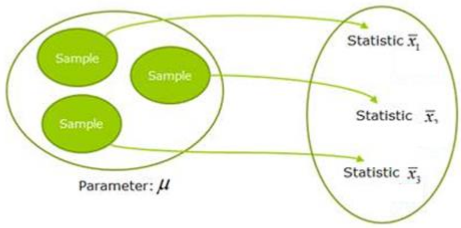
- **Πληθυσμός**:
- Ελληνικοί χαρακτήρες
  - μέση τιμή → μ
  - σταθερή απόκλιση → σ

- **Δείγμα**:
- Λατινικοί χαρακτήρες
  - μέση τιμή → $\bar{X}$
  - σταθερή απόκλιση → S



Πληθυσμός & δείγμα

- Επιλέγουμε ένα δείγμα και μετράμε το χαρακτηριστικό.
- Είναι απίθανο η μέση τιμή που προκύπτει από το δείγμα που έχουμε επιλέξει να είναι ίση με τη μέση τιμή του πληθυσμού.
- Η επιλογή ενός άλλου δείγματος θα οδηγούσε σε άλλη μέση τιμή – Αν υποθέσουμε ότι παίρνουμε πολλά δείγματα.



Τυπικό Σφάλμα (Standard Error) Μέσης Τιμής

Μέτρο της απόστασης της μέσης τιμής του δείγματος από την πραγματική στον πληθυσμό.

$$SE = \frac{SD}{\sqrt{n}}$$

7

Προσοχή! Το Standard Error (SE) διαφέρει από τη Standard Deviation (SD)

ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ (s ή SD)	ΤΥΠΙΚΟ ΣΦΑΛΜΑ (SE)
$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n-1}}$	$SE = \frac{SD}{\sqrt{n}}$
Εκφράζει τη διασπορά των παρατηρήσεων	Εκφράζει την απόσταση του δειγματικού μέσου από τον πραγματικό μέσο
Δεν επηρεάζεται ιδιαίτερα από το μέγεθος του δείγματος	Επηρεάζεται από το μέγεθος του δείγματος

8

Διάστημα εμπιστοσύνης (Confidence intervals, CI) (I)

- Δεδομένου ότι δεν μπορούμε να μετρήσουμε όλον τον πληθυσμό, πρέπει να βγάλουμε συμπέρασμα από το ένα δείγμα που μελετάμε και την εκτίμηση της μέσης τιμής και της τυπικής απόκλισης από το δείγμα αυτό.

- Εκτιμούμε τη μέση τιμή του δείγματος και συνοδεύουμε την εκτίμηση αυτή με ένα διάστημα εντός του οποίου είμαστε σχεδόν σίγουροι ότι βρίσκεται η πραγματική μέση τιμή.

- 95% ή 99% σίγουροι → 95% ή 99% διάστημα εμπιστοσύνης

9

Διάστημα εμπιστοσύνης (Confidence intervals, CI) (II)

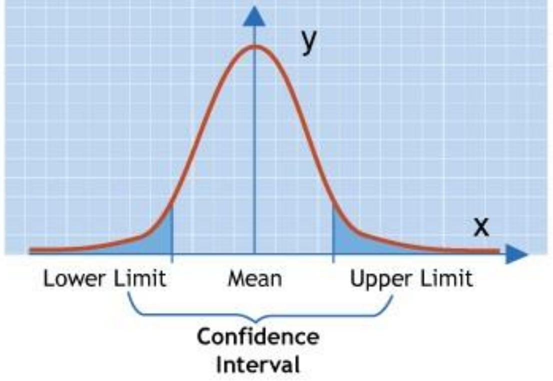
- Τι εκφράζει;

  Αν παίρναμε 100 διαφορετικά δείγματα και υπολογίζαμε το 95% διάστημα εμπιστοσύνης σε κάθε ένα από αυτά τότε τα 95 από αυτά θα περιλάμβαναν τη μέση τιμή του πληθυσμού.

10

95% Διάστημα Εμπιστοσύνης Μέσης Τιμής (I)

Το 95% Διάστημα Εμπιστοσύνης της μέσης τιμής είναι ένα διάστημα τιμών το οποίο περιλαμβάνει την πραγματική μέση τιμή του μεγέθους στον πληθυσμό με πιθανότητα 95%.



11

95% Διάστημα Εμπιστοσύνης Μέσης Τιμής (II)

Πώς υπολογίζεται :

Η εκτίμησή μας για τη μέση τιμή $\pm$ περιθώριο σφάλματος

↓ $\bar{X}$

↓ Καθορίζεται από το πόσο σίγουροι θέλουμε να είμαστε (95%), το μέγεθος του δείγματος και την ετερογένεια του χαρακτηριστικού στον πληθυσμό

12

95% Διάστημα Εμπιστοσύνης Μέσης Τιμής (III)

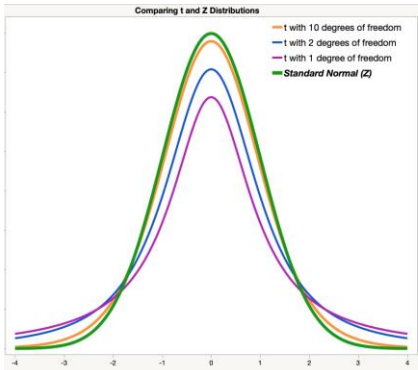
- Αν δείγμα **<u>μεγάλο (n>200)</u>** το 95% ΔΕ δίνεται από τον τύπο:

(μέση τιμή ± 1,96*SE)

- Αν δείγμα **<u>μικρό (n<200)</u>** τότε το 95% ΔΕ δίνεται από τον τύπο:

(μέση τιμή ± $t_{n-1,\,0.05}$*SE)

όπου t_{n-1} η οριακή τιμή της κατανομής t στο 5% για n-1 βαθμούς ελευθερίας, η οποία βρίσκεται από τους πίνακες της κατανομής t



13

Διάστημα Εμπιστοσύνης Μέσης Τιμής Μικρό Δείγμα (n<200)

Όταν έχω μικρό δείγμα, η κατανομή της μέσης τιμής αποκεντρώνεται, άρα για να έχω το 95% ΔΕ θα πρέπει να πάρω μεγαλύτερο-ευρύτερο διάστημα

95% Διάστημα Εμπιστοσύνης

M.T. ± t_{n-1} * SE

όπου n μέγεθος δείγματος

Οριακές τιμές αυξάνουν όσο n α ↓

	Πιθανότητα λάθους (α)		
ΒΕ	**10%**	**5%**	**1%**
1	6.31375	12.7062	63.6567
2	2.91999	4.30265	9.92484
3	2.35336	3.18245	5.84091
4	2.13185	2.77645	4.60409
5	2.01505	2.57058	4.03214
...	...	...	...
30	1.69726	2.04227	2.75000
100	1.66023	1.98397	2.62589
>500	1.64487	1.95999	2.57588

95% ΔΕ

14

Παράδειγμα

Έστω ότι η μέση τιμή **600 μετρήσεων** αναστημάτων ήταν 164 εκ. και το SE ήταν 0,5 εκ.

Τότε, το 95% διάστημα εμπιστοσύνης για τη μέση τιμή είναι το διάστημα (164-1.96*0.5 , 164+1.96*0.5)

Αυτό σημαίνει ότι είμαστε 95% σίγουροι ότι η πραγματική μέση τιμή του πληθυσμού βρίσκεται στο διάστημα (163.02 , 164.98)

Η πιθανότητα να κάνουμε λάθος είναι 5% = 0,05.

15

Κλάδοι στατιστικής

- Η **περιγραφική στατιστική** ανάλυση (descriptive statistics) ασχολείται με τη συνοπτική και <u>περιγραφική παρουσίαση</u> των δεδομένων μίας στατιστικής έρευνας.

- Η <u>ανάλυση</u> και <u>εξαγωγή</u> <u>συμπερασμάτων</u> ονομάζεται επαγωγική στατιστική ή **στατιστική συμπερασματολογία** (statistical inference, inferential statistics).

16

Στατιστική συμπερασματολογία

17

Έλεγχος υποθέσεων

Εφαρμόζεται μια **στατιστική φόρμουλα** για να υπολογιστεί η πιθανότητα ότι τα ευρήματα του μεγέθους που παρατηρούνται προέκυψαν λόγω τύχης.

18

Έλεγχος υπόθεσης (I)

Ο έλεγχος υπόθεσης είναι μια διαδικασία βάσει της οποίας συνάγουμε συμπεράσματα για μια παράμετρο του πληθυσμού (π.χ. τη μέση τιμή), χρησιμοποιώντας πληροφορίες που προέρχονται από το δείγμα μας.

Διατυπώνουμε το ερώτημα και καθορίζουμε τη μηδενική και την εναλλακτική υπόθεση.

Μηδενική υπόθεση (H_0):
Δεν υπάρχει διαφορά **ΣΤΟΝ ΠΛΗΘΥΣΜΟ**

Εναλλακτική υπόθεση (H_A ή H_1):
Υπάρχει διαφορά **ΣΤΟΝ ΠΛΗΘΥΣΜΟ**

Πάντα ξεκινάμε από τη μηδενική υπόθεση (H_0)

Σκοπός: Να απορρίψουμε ή όχι τη μηδενική υπόθεση

19

Έλεγχος υπόθεσης (II)

Όταν θέλουμε να συγκρίνουμε τη μέση τιμή μιας παραμέτρου σε 2 ανεξάρτητους πληθυσμούς:

Πάντα ξεκινάμε από τη μηδενική υπόθεση (H_0): Υποθέτουμε ότι η πραγματική μέση τιμή ενός πληθυσμού είναι ίση με ενός άλλου

$$H_0: \mu_1=\mu_2 \rightarrow \Delta= \mu_1-\mu_2 =0$$

Η εναλλακτική υπόθεση (HA) είναι:

$$H_A: \mu_1 \neq \mu_2 \rightarrow \Delta= \mu_1-\mu_2 \neq 0$$

Σκοπός: Να απορρίψουμε ή όχι την μηδενική υπόθεση.

20

Ορισμός της p-value στον έλεγχο υποθέσεων:

p-value είναι η πιθανότητα να παρατηρήσουμε ένα αποτέλεσμα τόσο ή περισσότερο ακραίο όσο το αποτέλεσμα ενός συγκεκριμένου δείγματος δεδομένου ότι ισχύει η μηδενική υπόθεση

p-value πιθανότητα ($0 < p < 1$)

Μικρό p-value: Τα αποτελέσματα από το δείγμα είναι μη πιθανά δεδομένης της H_0

21

Ερμηνεία p-value

- Αν για παράδειγμα: *p* = 0,046
  - Αυτό ερμηνεύεται ως ότι η πιθανότητα να προέκυψε το παρατηρούμενο αποτέλεσμα από τύχη είναι 0.046 ή ως ότι η πιθανότητα η μηδενική υπόθεση να είναι σωστή είναι 0.046.

22

Επίπεδο σημαντικότητας (α) (I)

Επίπεδο σημαντικότητας: Για να αποφασίσουμε ποιες τιμές της p-value θεωρούνται ασυνήθιστες, πρέπει να υπάρχει κάποιο όριο ώστε όλες οι p-values που είναι **μικρότερες** αυτού να οδηγούν σε **απόρριψη** της μηδενικής.

Αυτό το όριο είναι το επίπεδο σημαντικότητας α.

Επίπεδο σημαντικότητας (α) II

- ✓ Αν p-value < α, τότε απορρίπτουμε την H_0 και αποδεχόμαστε την H_A
- ✓ Αν p-value > α, τότε αποτυγχάνουμε να απορρίψουμε την H_0

Χρησιμοποιούμε το 0,05 (**5%**) (αυθαίρετα)!

Αν **p-value < 0.05** το αποτέλεσμα ονομάζεται στατιστικά σημαντικό

Άρα, στο επίπεδο του 5%, έχουμε στατιστικά σημαντική διαφορά μεταξύ δύο μέσων τιμών όταν p-value < 0.05

Εισαγωγή στον έλεγχο υποθέσεων

Σε μία έρευνα συλλέγονται, μεταξύ άλλων, τα εξής δεδομένα για τους συμμετέχοντες:

- Ηλικία (έτη)
- Φύλο (άνδρας/γυναίκα)
- Συστολική πίεση (mm Hg)
- Κάπνισμα (ναι/οχι)

Πιθανά ερωτήματα

- Διαφέρουν τα επίπεδα διαστολικής πίεσης μεταξύ ανδρών και γυναικών;
  - **Συστολική πίεση**: ποσοτική μεταβλητή
  - **Φύλο**: ποιοτική μεταβλητή με 2 επίπεδα

- Υπάρχει σχέση μεταξύ ηλικίας και διαστολικής πίεσης;
  - **Ηλικία**: ποσοτική μεταβλητή
  - **Συστολική πίεση**: ποσοτική μεταβλητή

Διερεύνηση σχέσεων (Ι)

	Έκβαση (ανταπόκριση)	
Έκθεση / Παρέμβαση	Ποσοτική	Ποιοτική
Ποιοτική	• t-test	•X2
Ποσοτική	•Συσχέτιση	

27

Διερεύνηση σχέσεων (ΙΙ)

	Έκβαση (ανταπόκριση)	
Έκθεση / Παρέμβαση	**Ποσοτική**	Ποιοτική
Ποιοτική	• **t-test**	•X^2
Ποσοτική	•Συσχέτιση	

28

t-test

29

Είδη t-test

Έλεγχος ενός δείγματος (one-sample t-test):

- Συγκρίνουμε τη μέση τιμή ενός δείγματος με μία συγκεκριμένη τιμή. Παράδειγμα: Υπάρχει διαφορά στο δείκτη μάζας σώματος μεταξύ ασθενών με σακχαρώδη διαβήτη και γενικού πληθυσμού;

Πολλαπλά δείγματα:

Εξαρτημένα δείγματα (**paired t-test**):

- Συγκρίνουμε τη μέση τιμή δύο κατανομών στις οποίες οι παρατηρήσεις αφορούν τον ίδιο πληθυσμό. Παράδειγμα: Υπάρχει διαφορά στο δείκτη μάζας σώματος των γυναικών πριν και μετά την εγκυμοσύνη;

Ανεξάρτητα (μη σχετιζόμενα) δείγματα (**two-sample t-test**):

- Συγκρίνουμε τη μέση τιμή δύο ανεξάρτητων κατανομών. Παράδειγμα: Υπάρχει διαφορά στο δείκτη μάζας σώματος μεταξύ ανδρών και γυναικών;

30

One-sample t-test

31

Βήματα ανάλυσης

1. Προσδιορίζουμε τους πληθυσμούς
2. Καθορίζουμε τη μηδενική και την εναλλακτική υπόθεση
3. Καθορίζουμε τα χαρακτηριστικά της κατανομής προς σύγκριση
4. Προσδιορίζουμε την κρίσιμη τιμή
5. Υπολογίζουμε το αποτέλεσμα της στατιστικής δοκιμασίας
6. Συγκρίνουμε και βγάζουμε συμπεράσματα

32

Έλεγχος ενός δείγματος (one-sample t-test)

Χρησιμοποιείται για να συγκρίνει τη μέση τιμή ενός δείγματος με μία συγκεκριμένη τιμή (εάν και εφόσον το δείγμα ακολουθεί κανονική κατανομή!!).

Παράδειγμα:

1. Διερεύνηση αν οι φοιτητές μίας Σχολής διαφέρουν από το γενικό πληθυσμό αναφορικά με κάποιο χαρακτηριστικό (π.χ. BMI) που είναι γνωστή η μέση τιμή του για τον πληθυσμό αυτό.
2. Διερεύνηση αν η διαφορά στις τιμές παρατηρείται κατά τύχη.

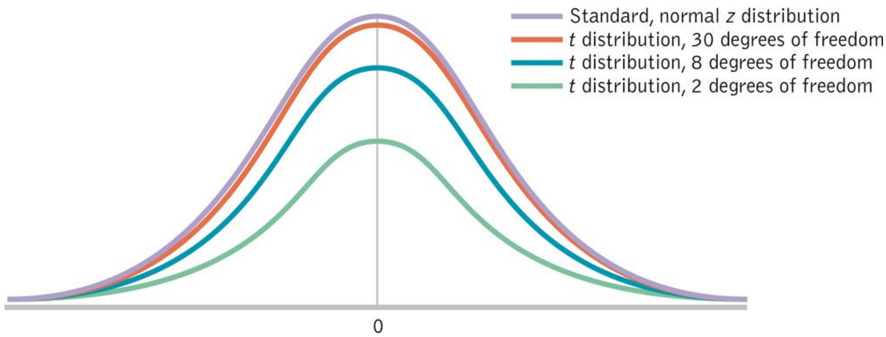
Εκτίμηση του t:

$$t = \frac{\textit{mean - comparison value}}{\textit{Standard Error}}$$

Υπολογίζει τη διαφορά της μέσης τιμής ενός δείγματος από έναν πληθυσμό αναφορικά με το τυπικό σφάλμα.

33

Student's t κατανομή



Ονομάζεται **"student's" t distribution** με *n* –1 βαθμούς ελευθερίας.

34

Βαθμοί ελευθερίας student's t κατανομής

- *Ο αριθμός των τιμών που είναι **ελεύθερος** να μεταβάλλεται όταν εκτιμάται μία παράμετρος του πληθυσμού από το δείγμα:*
  - *df* = *N* – 1 (έλεγχος ενός δείγματος)

Παράδειγμα: Εκτίμηση πόσο συχνά χρησιμοποιούμε οδοντικό νήμα [μέση τιμή = 2 (φορές την εβδομάδα)]

3	Free	2
5	Free	1
1	Free	0
0	Free	0
2	Free	0
1	LOCKED	9
Average = 2		**Average = 2**

Οι βαθμοί ελευθερίας δηλώνουν τις ελεύθερες (τυχαίες) τιμές που υπάρχουν στο πρόβλημα που μελετάμε. Στο παράδειγμά μας ενώ αρχικά οι βαθμοί ελευθερίας είναι n, όσες και οι παρατηρήσεις στο δείγμα, στη συνέχεια επειδή στον ορισμό της διασποράς περιλαμβάνεται η μέση τιμή του δείγματος, δεσμεύονται οι ελεύθερες τιμές με τη συνθήκη να ικανοποιούν την εξίσωση που δίνει τη μέση τιμή. Έτσι χάνεται ένας βαθμός ελευθερίας και οι βαθμοί ελευθερίας είναι n − 1.

35

Student's t statistic

- Η κατανομή του δείγματος:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

36

Δείγμα

- Έστω x_1, x_2, x_3 , … , x_n παρατηρήσεις ενός δείγματος πληθυσμού που ακολουθεί κανονική κατανομή.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \text{the sample mean}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} = \text{the sample standard deviation}$$

37

Έλεγχος υπόθεσης

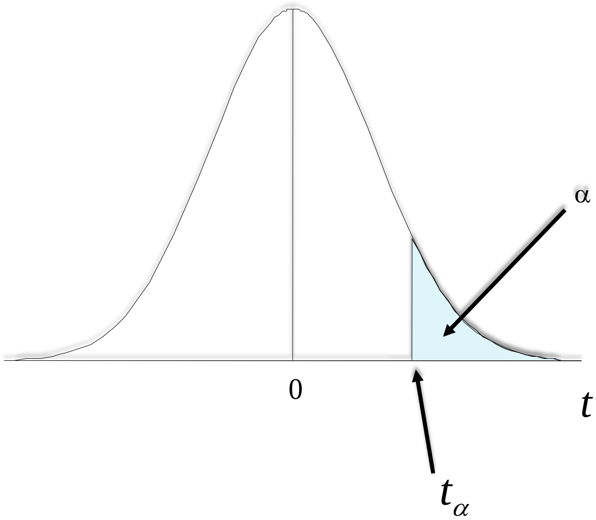
Εναλλακτική Υπόθεση H_A	Κρίσιμη περιοχή
$H_A : \mu \neq \mu_0$	$t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$H_A : \mu > \mu_0$	$t > t_\alpha$
$H_A : \mu < \mu_0$	$t < -t_\alpha$

t_a και $t_{a/2}$ είναι κρίσιμες τιμές της t κατανομής με $n – 1$ βαθμούς ελευθερίας

Θέλουμε να ελέγξουμε εάν η μέση τιμή του δείγματος (μ) είναι ίση με μία συγκεκριμένη τιμή (μ0).

38

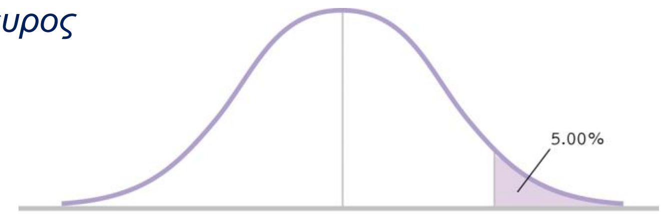
Κρίσιμες τιμές student’s t κατανομής



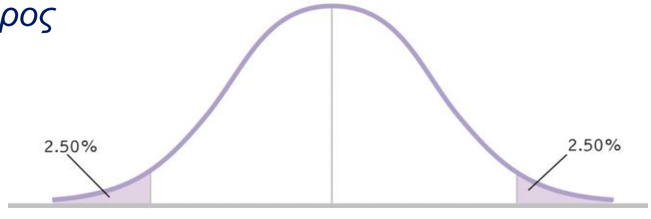
39

Στατιστικοί έλεγχοι

Μονόπλευρος έλεγχος



Αμφίπλευρος έλεγχος



40

Έλεγχος στη δοκιμασία t

- Την υπολογισθείσα τιμή του t-test την συγκρίνουμε με τις οριακές τιμές στον αντίστοιχο πίνακα:
  - Σε επίπεδο σημαντικότητας 5%
  - Σε βαθμούς ελευθερίας = n-1

α=0.05

Για n=7,

df=n-1=6

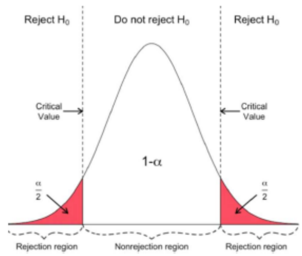
	α					
df	0.250	0.100	0.050	0.025	0.010	0.005
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106

41

Έλεγχος στη δοκιμασία t (II)

•**Αν** $|\bar{X} - \mu_0|/SE > t_{n-1,0.05}$
→ η διαφορά <u>είναι</u> στατιστικά σημαντική
(απορρίπτω τη μηδενική υπόθεση)

•**Αν** $|\bar{X} - \mu_0|/SE < t_{n-1,0.05}$
→ η διαφορά <u>δεν είναι</u> στατιστικά σημαντική
→μπορεί δηλαδή να οφείλεται στην τύχη
(δεν απορρίπτω τη μηδενική υπόθεση)



42

Παράδειγμα

- Έστω ότι οι τιμές x_1, x_2, x_3, x_4, x_5, x_6 αναπαριστούν τον αριθμό κιλών που έχασαν 6 άτομα σε δίαιτα.
- Θεωρούμε ότι οι x_1, x_2, x_3, x_4, x_5, x_6 ακολουθούν κανονική κατανομή με μέση τιμή μ και τυπική απόκλιση s.
- Θέλουμε να ελέγξουμε:

$H_0 : \mu \leq 0$ Η δίαιτα δεν είναι αποτελεσματική

Σε σχέση $H_A : \mu > 0$ Η δίαιτα είναι αποτελεσματική

43

Δεδομένα

1	2	3	4	5	6
2.0	1.0	1.4	-1.8	0.9	2.3

$$\bar{x} = 0.96667 \text{ and } s = 1.462418$$

44

Έλεγχος στη δοκιμασία t

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{0.96667 - 0}{(1.462418) / \sqrt{6}} = 1.619$$

Κρίσιμη περιοχή (*a* = 0.05). Βαθμοί ελευθερίας (n-1=6-1=5).

$$t_{0.05} = 2.015$$

$$t < ta$$

<u>Συμπέρασμα</u>: Δεν απορρίπτεται η H_0.
Η δίαιτα δεν είναι αποτελεσματική.

45

Two-sample t-test

46

46

Ερώτημα

– Υπάρχει διαφορά της αρτηριακής πίεσης μεταξύ των δύο φύλων;

47

Ερώτημα (απάντηση)

– Υπάρχει διαφορά της αρτηριακής πίεσης μεταξύ των δύο φύλων;
  - Αρτηριακή πίεση: ποσοτική μεταβλητή (κανονική κατανομή)
  - Φύλο: ποιοτική μεταβλητή (2 κατηγορίες)

  → t-test (two-sample t-test)

48

Two-sample t-test

Έστω ότι θέλουμε να συγκρίνουμε αν διαφέρει μία **ποσοτική μεταβλητή** στα **2 επίπεδα** μίας **ποιοτικής**.

Δηλαδή, έστω ότι θέλουμε να συγκρίνουμε δύο διαφορετικές μέσες τιμές, που προέρχονται από δυο **ανεξάρτητους** πληθυσμούς.

Προϋποθέσεις:

✓ Η μεταβλητή που μας ενδιαφέρει ακολουθεί την **κανονική κατανομή** και στους 2 ανεξάρτητους πληθυσμούς.

✓ Οι **τυπικές αποκλίσεις** δεν διαφέρουν (F-test) (εναλλακτικά ελέγχουμε εάν η μια δεν είναι διπλάσια της άλλης)

49

Παράδειγμα

Επίπεδα άλφα-φετοπρωτεΐνης σε έγκυες γυναίκες ανάλογα με τη φυλή.

Ερευνητική υπόθεση: Διαφέρουν τα επίπεδα σε Ασιάτισσες και Καυκάσιες έγκυες γυναίκες;

Θέλουμε να ελέγξουμε αν οι 2 μέσες τιμές είναι ίσες

H_0: $\mu_1=\mu_2$ (Μηδενική υπόθεση)

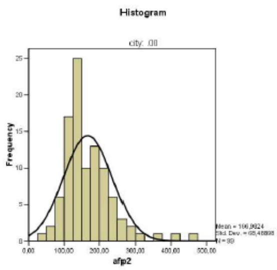
H_A: $\mu_1 \neq \mu_2$ (Εναλλακτική υπόθεση)

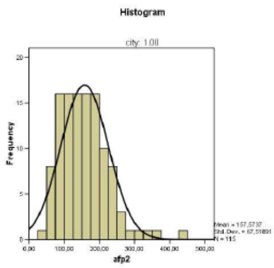
- Δίνονται οι τιμές της αλφα-φετοπρωτεΐνης σε n_1=99 Ασιάτισσες και n_2=115 Καυκάσιες έγκυες γυναίκες

Δείγμα 1 (Ασιάτισσες)	Δείγμα 2 (Καυκάσιες)
– n_1: Αριθμός ατόμων στο δείγμα 1	– n_2: Αριθμός ατόμων στο δείγμα 2
– $\bar{X}_1$: μέση τιμή	– $\bar{X}_2$: μέση τιμή
– SD_1 : σταθ. απόκλιση	– SD_2 : σταθ. απόκλιση

50

Έλεγχος προϋποθέσεων για εφαρμογή two sample t-test





Ομάδα 1:
n1=99 έγκυες γυναίκες στη Σαγκάη
$\bar{X}_1$=167 μmol/L, SD_1=68.5 μmol/L

Ομάδα 2:
n2=115 έγκυες γυναίκες στη Βοστώνη
$\bar{X}_2$=157.6 μmol/L, SD_2=67.5 μmol/L

Το μέγεθος κανονικά κατανεμημένο στις 2 ομάδες & παρόμοιες διασπορές (προϋποθέσεις για την εφαρμογή του t-test)

51

Εφαρμογή two sample t-test

- $\bar{X}_1$-$\bar{X}_2$= 9.4 → είναι μεγάλη αυτή η διαφορά;
- $t = \dfrac{|\bar{X}_1 - \bar{X}_2|}{\mathbf{SE}_{(\bar{X}_1 - \bar{X}_2)}} = \dfrac{9.4}{9.3}$=1.01
- Συγκρίνω με τιμή της t κατανομής στο α=5% και n_1+ n_2-2 ΒΕ (99+115-2=212) (→ κρίσιμη τιμή)
- t=1.01 <1.97 → **η διαφορά δεν είναι στατιστικά σημαντική στο 5%** (ούτε και στο 10%) – τα δεδομένα δεν επιτρέπουν να απορρίψουμε τη μηδενική υπόθεση (άρα δεν διαφέρουν τα επίπεδα ανάλογα με τη φυλή)
- Από στατιστικό πακέτο: p=0.314

Degrees of freedom	Significance level 10%	5%	1%
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.894	2.365	3.499
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
60	1.671	2.000	2.660
70	1.667	1.994	2.648
80	1.664	1.990	2.639
90	1.662	1.987	2.632
100	1.660	1.984	2.626
120	1.658	1.980	2.617
150	1.655	1.976	2.609
200	1.653	1.972	2.601
300	1.650	1.968	2.592
400	1.649	1.966	2.588
500	1.648	1.965	2.586
600	1.647	1.964	2.584
∞	1.645	1.960	2.576

52

Παράδειγμα αλφα-φετοπρωτεΐνης (ΙΙΙ)

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means (P-value)				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
afp2	Equal variances assumed	.035	.853	1.010	212	.314	9.40877	9.31862
	Equal variances not assumed			1.009	206.389	.314	9.40877	9.32861

Τα στατιστικά πακέτα εκτιμάνε το **p-value**, δηλαδή την πιθανότητα να παρατηρήσουμε αυτή τη διαφορά (9,4) στα μέσα επίπεδα αλφα-φετοπρωτεΐνης μεταξύ των 2 ομάδων γυναικών ή μια πιο ακραία διαφορά αν υποθέσουμε ότι ισχύει η μηδενική υπόθεση (ότι τα επίπεδα άλφα-φετοπρωτεΐνης δεν διαφέρουν ανάλογα με τη φυλή).

Συμβατικά, επιλέγουμε το 5% ως το επίπεδο σημαντικότητας. Άρα στο παράδειγμα, σε επίπεδο σημαντικότητας 5%, δεν έχουμε στατιστικά σημαντική διαφορά στα επίπεδα της πρωτεΐνης μεταξύ των γυναικών των 2 διαφορετικών φυλών, **αφού το p-value > 0.05.**

53

95% διάστημα εμπιστοσύνης για τη διαφορά των μέσων τιμών

Το 95% διάστημα εμπιστοσύνης για τη διαφορά δύο μέσων τιμών από δείγματα n_1 και n_2 είναι:

$$\boldsymbol{\delta} \pm t_{(n_1+n_2-2,\, 0.05)} * \mathbf{SE}_{\boldsymbol{\delta}}$$

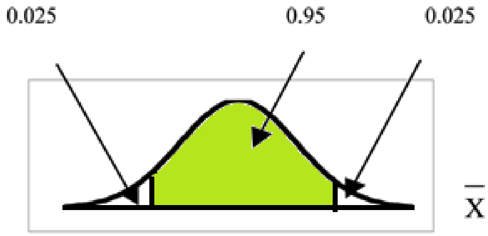
όπου t η τιμή της t κατανομής στο επίπεδο του 5% και στους n_1+n_2-2 βαθμούς ελευθερίας

54

Έλεγχος υπόθεσης και διάστημα εμπιστοσύνης

Το διάστημα εμπιστοσύνης παρέχει εύρος πιθανών τιμών για την τιμή της παραμέτρου στον πληθυσμό.

95% Δ.Ε. στη σύγκριση μέσων τιμών:
Αν το 0 **<u>δεν</u>** ανήκει στο 95% ΔΕ, τότε **<u>απορρίπτουμε την H_0</u>** ότι **Δ** = 0 σε ε.σ. α= 0.05 (pvalue < 0.05)



55

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
afp2	Equal variances assumed	.035	.853	1.010	212	.314	9.40877	9.31862	-8.96025	27.77780
	Equal variances not assumed			1.009	206.389	.314	9.40877	9.32861	-8.98281	27.80036

Τα στατιστικά πακέτα εκτιμάνε το 95% ΔΕ για τη διαφορά των δύο μέσων τιμών.

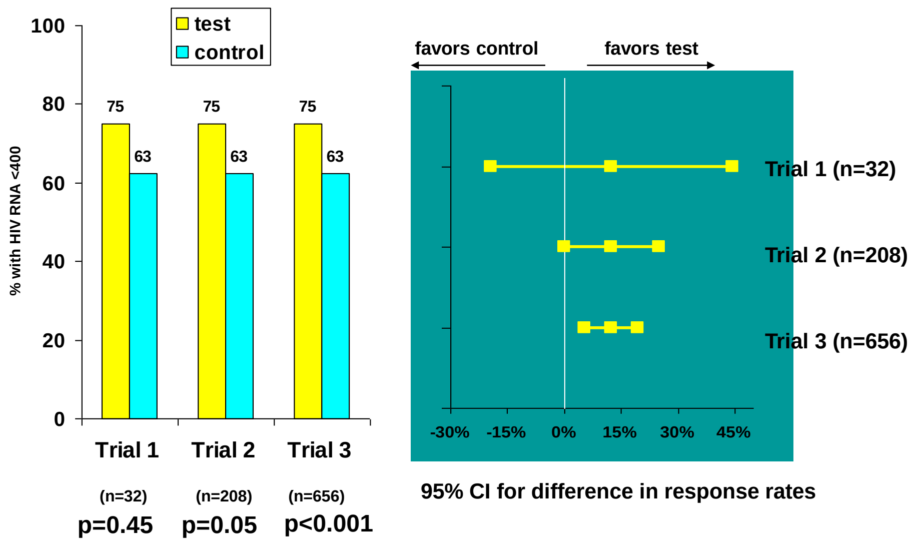
95% Δ.Ε: (-8.9 , 27.8).

Είμαστε 95% σίγουροι ότι το διάστημα αυτό συμπεριλαμβάνει την πραγματική διαφορά στις μέσες τιμές στους δύο διαφορετικούς πληθυσμούς.

Το παραπάνω διάστημα περιλαμβάνει το 0 ΑΡΑ **<u>δεν</u>** <u>απορρίπτουμε την H_0 σε επίπεδο 5%</u> (συμβαδίζει με τον έλεγχο υπόθεσης στο επίπεδο 5%). Οι μέσες τιμές πρωτεΐνης ΔΕΝ διαφέρουν στατιστικά σημαντικά σε επίπεδο 5% στους 2 διαφορετικούς πληθυσμούς.

56

p-value και όρια αξιοπιστίας ανάλογα με το μέγεθος του δείγματος



57

Παράδειγμα: Βάρος γέννησης νεογνού και φυλή (I)

Group Statistics

	city	N	Mean	Std. Deviation	Std. Error Mean	
bweight	.00	99	3431.5455	490.91128	49.33844	← Γυναίκες Σαγκάης
	1.00	115	3577.9130	481.35818	44.88688	← Γυναίκες Βοστόνης

<u>Ερώτημα</u>: Διαφέρει το βάρος γέννησης ανάλογα τη φυλή που ανήκει η μητέρα;

<u>**Έλεγχος Υπόθεσης:**</u>

H_0: Το βάρος γέννησης <u>δεν</u> διαφέρει ανάλογα τη φυλή:

$\mu_1=\mu_2 \rightarrow \Delta= \mu_1-\mu_2 =0$

H_A: Το βάρος γέννησης διαφέρει ανάλογα τη φυλή:

$\mu_1 \neq \mu_2 \rightarrow \Delta= \mu_1-\mu_2 \neq 0$

58

Παράδειγμα: Βάρος γέννησης νεογνού και φυλή (II)

Group Statistics

	city	N	Mean	Std. Deviation	Std. Error Mean	
bweight	.00	99	3431.5455	490.91128	49.33844	← Γυναίκες Σαγκάης
	1.00	115	3577.9130	481.35818	44.88688	← Γυναίκες Βοστόνης

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
bweight	Equal variances assumed	.038	.846	-2.198	212	.029	-146.36759	66.60330	-277.657	-15.07804
	Equal variances not assumed			-2.194	206.029	.029	-146.36759	66.70168	-277.873	-14.86223

Αφού p-value <0,05 και το 0 δεν περιλαμβάνεται στο 95% ΔΕ της διαφοράς: Απορρίπτουμε την H_0 (H_0: το βάρος γέννησης δεν διαφέρει μεταξύ των 2 φυλών). **Άρα το βάρος γέννησης διαφέρει σε βαθμό στατιστικά σημαντικό σε επίπεδο 5% μεταξύ των 2 φυλών.**

Παράδειγμα: Βάρος γέννησης νεογνού και φυλή (III)

Group Statistics

	city	N	Mean	Std. Deviation	Std. Error Mean	
bweight	.00	99	3431.5455	490.91128	49.33844	← Γυναίκες Σαγκάης
	1.00	115	3577.9130	481.35818	44.88688	← Γυναίκες Βοστόνης

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
bweight	Equal variances assumed	.038	.846	-2.198	212	.029	-146.36759	66.60330	-277.657	-15.07804
	Equal variances not assumed			-2.194	206.029	.029	-146.36759	66.70168	-277.873	-14.86223

Η μέση διαφορά βάρους γέννησης μεταξύ των 2 φυλών είναι -146.4 γραμμάρια και το 95% ΔΕ της διαφοράς είναι -277.7 έως -15.1 γραμμάρια, με το βάρος γέννησης να είναι μικρότερο στα μωρά των γυναικών από την Σαγκάη κατά μέσο όρο σε σχέση με τα μωρά των γυναικών από τη Βοστόνη.

Συνοπτικά t-test (I)

Αν θέλουμε να συγκρίνουμε 2 μέσες τιμών ενός ποσοτικού μεγέθους σε 2 ανεξάρτητους πληθυσμούς: **Independent t-test**

- Καθορίζουμε H_0 και H_A:

H_0: Δεν υπάρχει πραγματική διαφορά μεταξύ των μ_1 και μ_2 των 2 ανεξάρτητων πληθυσμών → $\mu_1 = \mu_2$ ή $\mu_1 - \mu_2 = \Delta = 0$

H_A : Υπάρχει πραγματική διαφορά μεταξύ των μ_1 και μ_2 των 2 ανεξάρτητων πληθυσμών → $\mu_1 \neq \mu_2$ ή $\mu_1 - \mu_2 = \Delta \neq 0$

61

Συνοπτικά t-test (II)

- Μετά τη στατιστική δοκιμασία, **απορρίπτουμε την H_0** αν:
  - ✓ Καταλήξουμε σε πολύ μικρή p-value (< 0,05)
  - ✓ Το μηδέν δεν περιλαμβάνεται στο 95% ΔΕ της διαφοράς των μέσων τιμών

Οι μέσες τιμές των 2 ανεξάρτητων πληθυσμών διαφέρουν στατιστικά σημαντικά σε επίπεδο 5%

- Μετά τη στατιστική δοκιμασία, **αποτυγχάνουμε να απορρίψουμε την H_0** αν:
  - ✓ Καταλήξουμε σε μεγάλη p-value (> 0,05)
  - ✓ Το μηδέν περιλαμβάνεται στο 95% ΔΕ της διαφοράς των μέσων τιμών

Οι μέσες τιμές των 2 ανεξάρτητων πληθυσμών δεν διαφέρουν στατιστικά σημαντικά σε επίπεδο 5%

62

Παράδειγμα

Table II. Patients' clinicopathologic characteristics.

Variable	No complications	Complications	*p*-Value
Age	61.6 ± 11.2	65.9 ± 12.0	< 0.001
Gender			0.187
Male	1737	328	
Female	629	101	
BMI (kg/m^2)	23.6 ± 3.4	24.1 ± 3.8	0.059
Hemoglobin level (g/dl	12.4 ± 2.3	11.9 ± 2.3	< 0.001
Serum prealbum (g/dl)	23.3 ± 5.6	22.3 ± 5.7	0.001

Zhou et al. 2015

63