Έλεγχος t (t-test)
Σύγκριση δύο μέσων

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

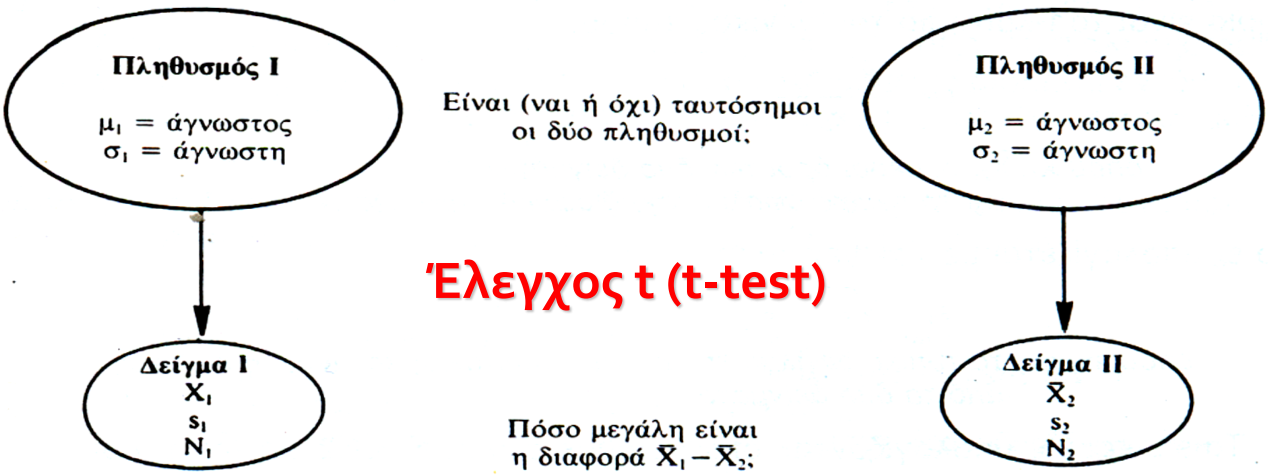
1

Βασικές στατιστικές παραδοχές της παραμετρικής ανάλυσης

- **Η κλίμακα μέτρησης των αρχικών δεδομένων πρέπει να είναι διαστήματος και κλίμακας λόγου**.
- **Παραδοχή** της **κανονικής κατανομής των μεταβλητών (normality).**
  - Οι μεταβλητές θεωρούνται ότι ανήκουν σε πληθυσμιακές κατανομές με κανονική μορφή, που έχουν δηλαδή καμπύλες συχνότητας συμμετρικές και μεσόκυρτες.
- **Παραδοχή** της **ομοιογένειας των διασπορών (homogeneity).**
  - Οι μεταβλητές θεωρούνται ότι ανήκουν σε πληθυσμιακές κατανομές με ίσες διασπορές ($\sigma_1^2 \approx \sigma_2^2$).
  - Όταν αυτή η παραδοχή δεν ικανοποιείται, τότε το λιγότερο πρέπει να είναι γνωστός ο λόγος των δύο διασπορών (σ_1^2 / σ_2^2).
  - Σε αντίθετη περίπτωση η ανάλυση πρέπει να γίνει με κάποια απαραμετρική ή άλλη ειδική παραμετρική μέθοδο.
- **Παραδοχή** της **τυχαίας δειγματοληψίας** ή **ανεξαρτησία δεδομένων (independence).**
  - Οι μεταβλητές θεωρούνται ότι προήλθαν από δειγματοληψία στην οποία κάθε μέλος του αντίστοιχου πληθυσμού είχε ισότιμη πιθανότητα συμπερίληψης στο δείγμα.

2

Έλεγχος t (t-test)



Η διαφορά μεταξύ των μέσων όρων $\bar{X}_1$ και $\bar{X}_2$ των δύο δειγμάτων είναι τόσο μεγάλη, ώστε να μπορούμε να υποστηρίζουμε ότι και οι πληθυσμοί από τους οποίους προέρχονται τα δύο δείγματα έχουν διαφορετικούς μέσους όρους μ_1 και μ_2; Ή, η διαφορά αυτή είναι μικρή. οπότε μπορούμε να αποφανθούμε ότι οι δύο αυτοί πληθυσμοί είναι ταυτόσημοι;

3

Τι σημαίνουν οι παραπάνω παραδοχές?

- Όταν οι παραδοχές αυτές ισχύουν επαρκώς, τότε μεγιστοποιείται η ισχύς του ελέγχου t και τα αποτελέσματα της ανάλυσης είναι επαρκώς βάσιμα.
- Η **ισχύς του στατιστικού ελέγχου** (power of the statistical test) αφορά τη μεγιστοποίηση της πιθανότητας απόρριψης μιας μηδενικής υπόθεσης (H_0) που δεν ισχύει στον πληθυσμό, δηλαδή της πιθανότητας επιλογής της σωστής απόφασης ή απλά της πιθανότητας αποφυγής του σφάλματος τύπου II.

4

4

Σύμφωνα με τις επισημάνσεις του Cohen (1988), η πιθανότητα του σφάλματος τύπου ΙΙ είναι **β**, οπότε η ισχύς του του στατιστικού ελέγχου είναι **1-β**.

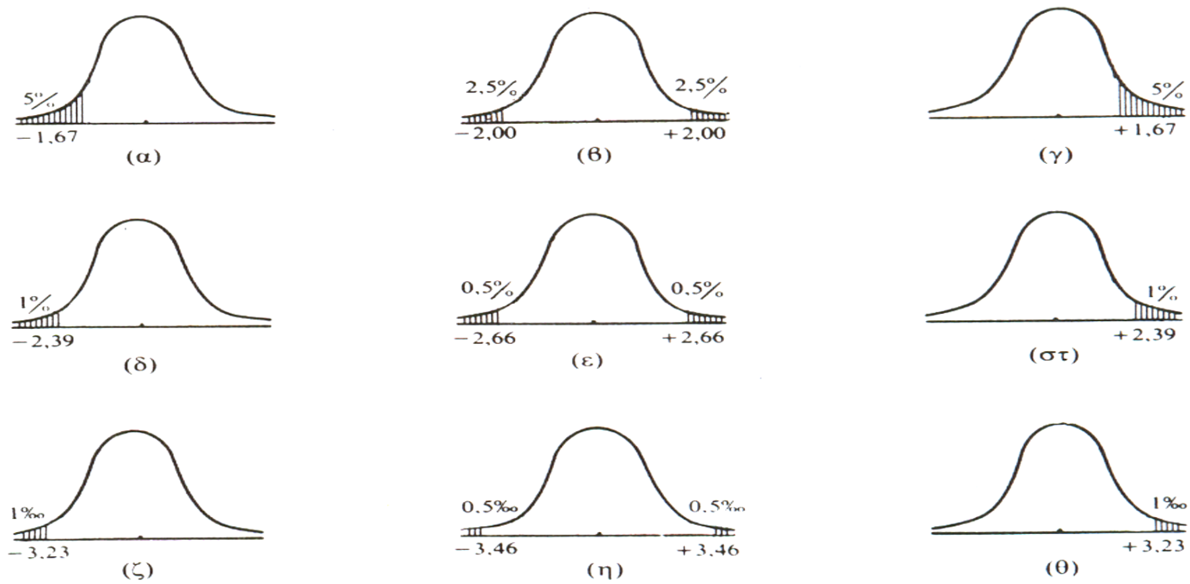
Possible Hypothesis Test Outcomes		
Decision	Accept H_0	Reject H_0
H_0 is true	Correct Decision (No error)	Type I Error
	Probability = 1 - α	Probability = α
H_0 is false	Type II Error	Correct Decision (No error)
	Probability = β	Probability = 1 - β

5

Παραμετρικοί Έλεγχοι Στατιστικής Σημαντικότητας

- Μια μεγάλη κατηγορία παραμετρικών ελέγχων στατιστικής σημαντικότητας (parametric tests of statistical significance) αφορά τις **συγκρίσεις** μεταξύ εξαρτημένων (dependent) ή ανεξάρτητων (independent) **δειγμάτων**.
- Η σύγκριση k δειγμάτων γίνεται με την ανάλυση διασποράς (ANalysis Of VAriance) ή πιο απλά ANOVA, που συγκρίνει τους αντίστοιχους k μέσους (means).
- Στη βασική της μορφή η διαδικασία αυτή περιλαμβάνει τη σύγκριση 2 μέσων (comparing two means) και τη στατιστική σημαντικότητα της διαφοράς τους (μ_1 -μ_2) είτε με τον **έλεγχο z** (z-test) είτε με τον **έλεγχο t** (t-test).

6



. Διευθέτηση κρίσιμων t-τιμών στατιστικής σημαντικότητας 5%, 1% και 1‰ μονής και διπλής κατεύθυνσης με βαθμούς ελευθερίας df=59.

7

Στατιστικές υποθέσεις
(Παραδείγματα)

- **Μηδενική Υπόθεση (H_0)**. Οι μέσοι όροι των βαθμών στο μάθημα της στατιστικής του συνόλου των ανδρών και του συνόλου των γυναικών είναι ίσοι.

- **Εναλλακτική Υπόθεση (H_α)**. Οι μέσοι όροι των βαθμών στο μάθημα της στατιστικής του συνόλου των ανδρών και του συνόλου των γυναικών είναι διαφορικοί (υπόθεση διπλής κατεύθυνσης)

ή

- **Εναλλακτική Υπόθεση (H_α)**. Ο μέσος όρος των βαθμών στο μάθημα της στατιστικής του συνόλου των ανδρών είναι μεγαλύτερος από τον μέσο όρο των βαθμών του συνόλου των γυναικών (υπόθεση μονής κατεύθυνσης).

ή

- **Εναλλακτική Υπόθεση (H_α)**. Ο μέσος όρος των βαθμών στο μάθημα της στατιστικής του συνόλου των ανδρών είναι μικρότερος από τον μέσο όρο των βαθμών του συνόλου των γυναικών (υπόθεση μονής κατεύθυνσης)

8

Έλεγχο z (z-test)

- Ο **έλεγχος z** χρησιμοποιείται σε εφαρμογές υπολογισμού πιθανότητας επιλογής μιας τιμής (μέσης ή απλής) με βάση την τυπική κανονική κατανομή.
- Ειδικά για τη δειγματική κατανομή των διαφορών δύο μέσων, το z-test χρησιμοποιείται σε δύο γενικές περιπτώσεις.

9

9

Έλεγχο z (z-test): ***Περίπτωση 1η***

- Για σύγκριση του μέσου ($\bar{X}$) ενός δείγματος με **τον μέσο (μ) του γεννήτορα πληθυσμού** μιας μεταβλητής (π.χ., Χ), όταν είναι γνωστή **η πληθυσμιακή τυπική απόκλιση σ_X**:
- τυπικό σφάλμα του μέσου: $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{N}}$
- έλεγχος z $z = (\bar{X} - \mu)/\sigma_{\bar{X}}$
- Πιθανότητα p: χώρος z → ∞ (βλ Πίνακα)
- σύγκριση p με α: (π.χ., α = 0.05 ή α = 0.01)
  - αν $p \leq \alpha$, η διαφορά είναι στατιστικώς σημαντική: άλλος πληθυσμός.
  - Αν $p > \alpha$, η διαφορά είναι στατιστικώς μη σημαντική: ίδιος πληθυσμός.

10

10

Έλεγχο z (z-test): *Περίπτωση 2η*

- Για σύγκριση 2 πληθυσμιακών μέσων (μ_1, μ_2) μίας μεταβλητής (π.χ., X_1, X_2), όπου είναι γνωστές οι τυπικές αποκλίσεις σ_1 και σ_2:
- τυπικό σφάλμα της διαφοράς των 2 μέσων:
  - για εξαρτημένους πληθυσμούς: $\sigma_d = \sqrt{\frac{\sigma_1^2}{N} + \frac{\sigma_2^2}{N} - 2\rho\frac{\sigma_1}{\sqrt{N}}\frac{\sigma_2}{\sqrt{N}}}$
  - για ανεξάρτητους πληθυσμούς: $\sigma_d = \sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}$
- έλεγχος z : $z = \frac{\mu_1 - \mu_2}{\sigma_d}$
- Πιθανότητα p: χώρος z → ∞ (βλ. Πίνακα)
- σύγκριση p με α: (π.χ. $\alpha = 0.05$ ή $\alpha = 0.01$)
  - αν $p \leq \alpha$, η διαφορά είναι στατιστικώς σημαντική: διαφορετικοί πληθυσμοί.
  - Αν $p > \alpha$, η διαφορά είναι στατιστικώς μη σημαντική: ίδιοι πληθυσμοί.

11

11

Το στατιστικό κριτήριο που είναι κατάλληλο για την ανάλυση των δειγμάτων είναι ο *έλεγχος t (t-test)*

$$t = \frac{\overline{X}_1 - \overline{X}_2}{SE_d}$$

όπου SE_D το **τυπικό σφάλμα της διαφοράς** των δύο μέσων.

Οι δύο **βασικοί τύποι** εκτίμησης του **τυπικού σφάλματος της διαφοράς δύο μέσων**

σε **εξαρτημένα δείγματα X_1 & X_2** της ίδιας μεταβλητής (X)

$$SE_D = \sqrt{(SE_1)^2 + (SE_2)^2 - 2r(SE_1)(SE_2)}$$

& σε **ανεξάρτητα δείγματα X_1 & X_2** της ίδιας μεταβλητής (X)

$$SE_D = \sqrt{(SE_1)^2 + (SE_2)^2} = \sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}$$

όπου SE_1 & SE_2 τα τυπικά σφάλματα των μέσων των δειγμάτων X_1 και X_2 και s_1 & s_2 οι αμερόληπτες τυπικές αποκλίσεις τους, αντίστοιχα.

12

12

Έλεγχος t: Σύγκριση 2 εξαρτημένων δειγμάτων

- Τα εξαρτημένα δείγματα (dependent samples) προέρχονται από εξαρτημένους, δηλαδή σχετιζόμενους σε κάποιο βαθμό πληθυσμούς.
- Στην επιστημονική έρευνα μπορούν να υπάρξουν 2, 3, …, k εξαρτημένα δείγματα σε **2 περιπτώσεις**:
- **1η περίπτωση** (η συνήθης): Επαναλαμβανόμενες μετρήσεις.
- **2η περίπτωση** (πιο σπάνια): δύο διαφορετικές ομάδες ατόμων (subjects) αντιπαραβάλλονται (matched groups) με βάση κάποια κοινά χαρακτηριστικά που έχουν σχέση με τη μεταβλητή ιδιότητα (εξαρτημένη) X της έρευνας.
- **Μηδενική υπόθεση**: $H_0: \mu_1 - \mu_2 = 0$
- **Εναλλακτική υπόθεση**: $H_\alpha: \mu_1 - \mu_2 \neq 0$

13

13

Έλεγχος t-test – Εξαρτημένα Δείγματα (r ≠ 0)

Μέθοδος συσχέτισης

- **Βήμα 1ο**: Μέσοι $\bar{X}_1 = \frac{\sum X_1}{N}$ $\quad \bar{X}_2 = \frac{\sum X_2}{N}$
- **Βήμα 2ο**: Αμερόληπτες Τυπικές αποκλίσεις

$$s_1 = \sqrt{\frac{\sum x_1^2}{N-1}} = \sqrt{\frac{N\sum X_1^2-(\sum X_1)^2}{N(N-1)}} \qquad s_2 = \sqrt{\frac{\sum x_2^2}{N-1}} = \sqrt{\frac{N\sum X_2^2-(\sum X_2)^2}{N(N-1)}}$$

- **Βήμα 3ο**: Τυπικά σφάλματα των μέσων $SE_1 = \frac{s_1}{\sqrt{N}}$ $\quad SE_2 = \frac{s_2}{\sqrt{N}}$
- **Βήμα 4ο**: Συντελεστής συσχέτισης

$$r = \frac{\sum (x_1 x_2)}{(N-1)s_1 s_2} \quad ή \quad = \frac{\sum (x_1 x_2)}{\sqrt{(\sum x_1^2)(\sum x_2^2)}}$$

- **Βήμα 5ο**: Τυπικό σφάλμα της διαφοράς (των δύο μέσων)

$$SE_D = \sqrt{(SE_1)^2 + (SE_2)^2 - 2r(SE_1)(SE_2)}$$

- **Βήμα 6ο**: Κριτήριο $t = \frac{\bar{X}_1 - \bar{X}_2}{SE_D}$

14

14

Έλεγχος t-test – Εξαρτημένα Δείγματα (r ≠ 0)

Μέθοδος συσχέτισης

- **Βήμα 7ο**: Έλεγχος της στατιστικής σημαντικότητας του t (Πίνακας): με πιθανότητα α & βαθμούς ελευθερίας **df = N – 1**, η κρίσιμη τιμή t_c είναι $t = t_{a,\,df}$ οπότε
- αν $|t| \geq t_c$ απορρίπτουμε την H_o: η διαφορά είναι στατιστικώς σημαντική,
- αν $|t| < t_c$ αποδεχόμαστε την H_o: η διαφορά δεν είναι στατιστικώς σημαντική.
- **Βήμα 8ο**: Μέγεθος Επίδρασης (Effect Size), Διαφοράς:

$d = (\bar{X}_1 - \bar{X}_2)/s$, όπου s η τυπική απόκλιση (i) της $1^{ης}$ μέτρησης (X_1) αν ο σχεδιασμός είναι Πριν-Μετά (Pre-Post design) ή (ii) και από τα 2 δείγματα $[(s_1+s_2)/2]$, αν οι 2 συνθήκες είναι ερευνητικά ισοδύναμες.

15

15

Μαθητής	Ευστοχία		Απόκλιση		Γινόμενα, Τετράγωνα		
i	Πρίν X_1	Μετά X_2	$X_1 - \bar{X}_1 = x_1$	$X_2 - \bar{X}_2 = x_2$	x_1x_2	x_1^2	x_2^2
1	4	5	0	0	0	0	0
2	5	6	1	1	1	1	1
3	2	4	-2	-1	2	4	1
4	4	5	0	0	0	0	0
5	3	4	-1	-1	1	1	1
6	4	5	0	0	0	0	0
7	2	3	-2	-2	4	4	4
8	6	7	2	2	4	4	4
9	6	6	2	1	2	4	1
Σ =	36	45	0	0	**14**	**18**	**12**

X1 = χωρίς προθέρμανση, X2 = μετά από προθέρμανση, απόκλιση x = X - μέσος

Δεδομένα δύο Εξαρτημένων δειγμάτων

- Οι στατιστικές υποθέσεις της έρευνας αυτής διατυπώνονται ως εξής:
- **Μηδενική υπόθεση**: H_0: $\mu_1 = \mu_2$
- **Εναλλακτική υπόθεση**: H_α: $\mu_1 < \mu_2$ (μονόπλευρος έλεγχος)

16

16

Βήμα 1ο - Μέσοι $\bar{X}_1 = \frac{36}{9} = 4$ $\quad \bar{X}_2 = \frac{45}{9} = 5$

Βήμα 2ο - Τυπικές αποκλίσεις (αμερόληπτες, unbiased)

$$s_1 = \sqrt{\frac{18}{9\text{-}1}} = 1.5 \qquad s_2 = \sqrt{\frac{12}{9\text{-}1}} = 1.2247$$

Βήμα 3ο - Τυπικά σφάλματα των μέσων

$$SE_1 = \frac{1.5}{\sqrt{9}} = \frac{1.5}{3} = 0.5 \qquad SE_2 = \frac{1.2247}{\sqrt{9}} = \frac{1.2247}{3} = 0.4082$$

Βήμα 4ο - Συντελεστής συσχέτισης (των 2 μετρήσεων)

$$r = \frac{14}{\sqrt{(18)(12)}} = \frac{14}{14.697} = 0.9526$$

Βήμα 5ο - Τυπικό σφάλμα της διαφοράς (των μέσων)

$$SE_D = \sqrt{(0.5)^2 + (0.4082)^2 - 2(0.9526)(0.5)(0.4082)} = 0.1667$$

Βήμα 6ο - Κριτήριο $t = \frac{5-4}{0.1667} = \frac{1}{0.1667} = 5.999 \approx 6$

Μαθητής	Ευστοχία		Απόκλιση		Γινόμενα, Τετράγωνα		
i	Πρίν X_1	Μετά X_2	X_1- $\bar{X}_1$ = x_1	X_2- $\bar{X}_2$ = x_2	x_1x_2	x_1^2	x_2^2
1	4	5	0	0	0	0	0
2	5	6	1	1	1	1	1
3	2	4	-2	-1	2	4	1
4	4	5	0	0	0	0	0
5	3	4	-1	-1	1	1	1
6	4	5	0	0	0	0	0
7	2	3	-2	-2	4	4	4
8	6	7	2	2	4	4	4
9	6	6	2	1	2	4	1
Σ =	36	45	0	0	**14**	**18**	**12**

X1 = χωρίς προθέρμανση, X2 = μετά από προθέρμανση, απόκλιση x = X - μέσος

Βήμα 7ο - Έλεγχος της στατιστικής σημαντικότητας του t με επίπεδο πιθανότητας α = 0.05 (μονόπλευρος έλεγχος) και βαθμούς ελευθερίας df = N -1=9 -1 = 8 η κρίσιμη τιμή t είναι

$t_c = t_{a,\,df} = t_{0.05,\,8} = 1.860$ και επειδή $|t| = 6 > t_c = 1.860$

με πιθανότητα σφάλματος τύπου Ι ίση με 5% απορρίπτουμε τη μηδενική υπόθεση ($H_0 : \mu_1 = \mu_2$) και αποδεχόμαστε την εναλλακτική ($H_A : \mu_1 < \mu_2$).

Στατιστικό συμπέρασμα: η διαφορά των 2 μέσων πληθυσμιακών συνθηκών ευστοχίας είναι στατιστικά σημαντική ($p < 0.05$).

17

Critical Values for Student's *t* Distribution

one-tail area	0.250	0.125	0.100	0.075	0.050	0.025	0.010	0.005	0.0005
two-tail area	0.500	0.250	0.200	0.150	0.100	0.050	0.020	0.010	0.0010
d.f. \ c	0.500	0.750	0.800	0.850	0.900	0.950	0.980	0.990	0.999
1	1.000	2.414	3.078	4.165	6.314	12.706	31.821	63.657	636.619
2	0.816	1.604	1.886	2.282	2.920	4.303	6.965	9.925	31.599
3	0.765	1.423	1.638	1.924	2.353	3.182	4.541	5.841	12.924
4	0.741	1.344	1.533	1.778	2.132	2.776	3.747	4.604	8.610
5	0.727	1.301	1.476	1.699	2.015	2.571	3.365	4.032	6.869
6	0.718	1.273	1.440	1.650	1.943	2.447	3.143	3.707	5.959
7	0.711	1.254	1.415	1.617	1.895	2.365	2.998	3.499	5.408
8	0.706	1.240	1.397	1.592	1.860	2.306	2.896	3.355	5.041
9	0.703	1.230	1.383	1.574	1.833	2.262	2.821	3.250	4.781
10	0.700	1.221	1.372	1.559	1.812	2.228	2.764	3.169	4.587
11	0.697	1.214	1.363	1.548	1.796	2.201	2.718	3.106	4.437
12	0.695	1.209	1.356	1.538	1.782	2.179	2.681	3.055	4.318
13	0.694	1.204	1.350	1.530	1.771	2.160	2.650	3.012	4.221
14	0.692	1.200	1.345	1.523	1.761	2.145	2.624	2.977	4.140
15	0.691	1.197	1.341	1.517	1.753	2.131	2.602	2.947	4.073
16	0.690	1.194	1.337	1.512	1.746	2.120	2.583	2.921	4.015
17	0.689	1.191	1.333	1.508	1.740	2.110	2.567	2.898	3.965
18	0.688	1.189	1.330	1.504	1.734	2.101	2.552	2.878	3.922
19	0.688	1.187	1.328	1.500	1.729	2.093	2.539	2.861	3.883
20	0.687	1.185	1.325	1.497	1.725	2.086	2.528	2.845	3.850
21	0.686	1.183	1.323	1.494	1.721	2.080	2.518	2.831	3.819
22	0.686	1.182	1.321	1.492	1.717	2.074	2.508	2.819	3.792
23	0.685	1.180	1.319	1.489	1.714	2.069	2.500	2.807	3.768
24	0.685	1.179	1.318	1.487	1.711	2.064	2.492	2.797	3.745
25	0.684	1.198	1.316	1.485	1.708	2.060	2.485	2.787	3.725
26	0.684	1.177	1.315	1.483	1.706	2.056	2.479	2.779	3.707
27	0.684	1.176	1.314	1.482	1.703	2.052	2.473	2.771	3.690
28	0.683	1.175	1.313	1.480	1.701	2.048	2.467	2.763	3.674
29	0.683	1.174	1.311	1.479	1.699	2.045	2.462	2.756	3.659
30	0.683	1.173	1.310	1.477	1.697	2.042	2.457	2.750	3.646
35	0.682	1.170	1.306	1.472	1.690	2.030	2.438	2.724	3.591
40	0.681	1.167	1.303	1.468	1.684	2.021	2.423	2.704	3.551
45	0.680	1.165	1.301	1.465	1.679	2.014	2.412	2.690	3.520
50	0.679	1.164	1.299	1.462	1.676	2.009	2.403	2.678	3.496
60	0.679	1.162	1.296	1.458	1.671	2.000	2.390	2.660	3.460
70	0.678	1.160	1.294	1.456	1.667	1.994	2.381	2.648	3.435
80	0.678	1.159	1.292	1.453	1.664	1.990	2.374	2.639	3.416
100	0.677	1.157	1.290	1.451	1.660	1.984	2.364	2.626	3.390
500	0.675	1.152	1.283	1.442	1.648	1.965	2.334	2.586	3.310
1000	0.675	1.151	1.282	1.441	1.646	1.962	2.330	2.581	3.300
∞	0.674	1.150	1.282	1.440	1.645	1.960	2.326	2.576	3.291

For degrees of freedom *d.f.* not in the table, use the closest *d.f.* that is *smaller*.

επίπεδο πιθανότητας α = 0.05 (μονόπλευρος έλεγχος)

Βαθμοί ελευθερίας df = N -1=9 -1 = 8

Κρίσιμη τιμή t - $t_c = t_{a,\,df} = t_{0.05,\,8} = 1.860$ και επειδή $|t| = 6 > t_c = 1.860$

Στατιστικό συμπέρασμα: η διαφορά των 2 μέσων πληθυσμιακών συνθηκών ευστοχίας είναι στατιστικά σημαντική ($p < 0.05$).

Βήμα 8ο - **Μέγεθος επίδρασης (effect size), Διαφοράς**:

$d = (\bar{X}_1 - \bar{X}_2)/S_1 = (5 - 4)/1.5 =$ -1/1·5= 0.67.

Ερευνητικό συμπέρασμα: η προθέρμανση βελτιώνει σημαντικά την ευστοχία στις ελεύθερες βολές στο μπάσκετ σε μαθητές ηλικίας 12 ετών, με το μέγεθος της βελτίωσης (0.67) να είναι μέτριο προς μεγάλο (Cohen, 1988, σελ. 25-27):

0.2 = μικρό,

0.5 = μέτριο,

0.8 = μεγάλο).

18

Έλεγχος t-test – Εξαρτημένα Δείγματα (r ≠ 0): *Μέθοδος Διαφορών (D)*

Βήμα 1ο - Μέση διαφορά $\bar{D} = \bar{X}_1 - \bar{X}_2 = \frac{\sum D}{N}$

Βήμα 2ο - Αμερόληπτη τυπική απόκλιση της διαφοράς $S_D = \sqrt{\frac{N\sum D^2 - (\sum D)^2}{N(N-1)}}$

Βήμα 3ο - Τυπικό σφάλμα της μέσης διαφοράς $SE_D = \frac{s_D}{\sqrt{N}}$

Βήμα 4ο - Κριτήριο $t = \frac{\bar{X}_1 - \bar{X}_2}{SE_D} = \frac{\bar{D}}{SE_D}$

Βήμα 5ο - Έλεγχος της στατιστικής σημαντικότητας του t (Πίν. 21.Δ)
με πιθανότητα α & βαθμούς ελευθερίας $df = N - 1$
η κρίσιμη τιμή t είναι $t_c = t_{\alpha,\, df}$, οπότε

- αν $|t| \geq t_c$ απορρίπτουμε την H_0 : η διαφορά είναι στατιστικώς σημαντική,
- αν $|t| < t_c$ αποδεχόμαστε την H_0 : η διαφορά δεν είναι στατιστικώς σημαντική.

Βήμα 6ο - Μέγεθος επίδρασης (effect size): όπως πιο πάνω (Βήμα 8ο).

19

Μαθητής	Ευστοχία		Διαφορά	
i	Πρίν X_1	Μετά X_2	D	D^2
1	4	5	-1	1
2	5	6	-1	1
3	2	4	-2	4
4	4	5	-1	1
5	3	4	-1	1
6	4	5	-1	1
7	2	3	-1	1
8	6	7	-1	1
9	6	6	0	0
Σ =	36	45	-9	11

X_1 = χωρίς προθέρμανση, X_2 = μετά από προθέρμανση, $D = X_1 - X_2$ =διαφορά.

Βήμα 1ο - Μέση διαφορά

$$\bar{D} = \bar{X}_1 - \bar{X}_2 = 4 - 5 = -1 \quad ή \quad \frac{\sum D}{N} = \frac{-9}{9} = -1$$

Βήμα 2ο - Αμερόληπτη τυπική απόκλιση της διαφοράς

$$S_D = \sqrt{\frac{9(11) - (-9)^2}{9(9\text{-}1)}} = \sqrt{\frac{18}{72}} = 0.5$$

Βήμα 3ο - Τυπικό σφάλμα της μέσης διαφοράς

$$SE_D = \frac{0.5}{\sqrt{9}} = \frac{0.5}{3} = 0.1667$$

Βήμα 4ο - Κριτήριο $t = \frac{4-5}{0.1667} = \frac{-1}{0.1667} = -5.999 \approx -6$

Βήμα 5ο - Έλεγχος της στατιστικής σημαντικότητας του t (Πίνακας)

- με επίπεδο πιθανότητας α = 0.05 (μονόπλευρος έλεγχος) και
- βαθμούς ελευθερίας df=N -1 = 9 -1 = 8
- η κρίσιμη τιμή t είναι $t_c = t_a > df = t_{0.05,\,8} = 1.860$ και επειδή $|t| = 6 > t_c = 1.860$

με πιθανότητα σφάλματος τύπου Ι ίση με 5%, καταλήγουμε στα ίδια συμπεράσματα, όπως με τη μέθοδο συσχέτισης.

- Βήμα 6ο - Μέγεθος Επίδρασης (effect size) Διαφοράς

20

Έλεγχος t: Σύγκριση 2 ανεξάρτητων δειγμάτων (1)

Οι στατιστικές διερευνήσεις αυτού του είδους βασίζονται στις **στατιστικές παραδοχές** ***της κανονικότητας των κατανομών*** και ***της ομοιογένειας των διασπορών*** και στοχεύουν στον πιθανολογικό έλεγχο της

μηδενικής υπόθεσης H_0: $\mu_1 = \mu_2$

με την **εναλλακτική υπόθεση** H_A: $\mu_1 \neq \mu_2$

21

Έλεγχος t: Σύγκριση 2 ανεξάρτητων δειγμάτων (2)

Οι 2 ομάδες υποκειμένων σχεδιάζονται να είναι **ισάριθμες** ($N_1 = N_2$).

Συχνά, όμως, για διάφορους λόγους (απόσυρση, ελλιπείς μετρήσεις κ.λπ.) οι 2 ομάδες γίνονται **άνισες** ($N_1 \neq N_2$).

Ακόμα, ενίοτε οι **διασπορές των 2 πληθυσμών** είτε είναι **άνισες** ($\sigma_1^2 \neq \sigma_2^2$), είτε εκτιμώνται ως άνισες στα αντίστοιχα δείγματα, οπότε δεν ισχύει η παραδοχή της ομοιογένειας των διασπορών και αυτό συνυπολογίζεται.

22

Έλεγχος t: Σύγκριση 2 ανεξάρτητων δειγμάτων (3)

Ο κλασικός **έλεγχος της ομοιογένειας** των 2 **διασπορών** γίνεται με τη **μέθοδο Levene**, που προηγείται του ελέγχου t, σύμφωνα με τον οποίο:

- Αν δειχθεί ότι οι 2 διασπορές της εξαρτημένης μεταβλητής (X) **δεν** **διαφέρουν** σημαντικά, τότε οι 2 διασπορές συνδυάζονται (pooled variance) και ο έλεγχος t γίνεται με βαθμούς ελευθερίας df = $N_1 + N_2 - 2$.
- Αν δειχθεί ότι οι 2 διασπορές της εξαρτημένης μεταβλητής (X) **διαφέρουν** σημαντικά, τότε ο έλεγχος t γίνεται με χωριστές διασπορές (separate variances) και με διόρθωση Welch (1938) στους βαθμούς ελευθερίας.

23

Βήμα 1ο – Μέσοι $\bar{X}_1 = \frac{\sum X_1}{N_1}$ $\bar{X}_2 = \frac{\sum X_2}{N_2}$

Βήμα 2ο - Αμερόληπτες διασπορές (unbiased variances)

$$s_1^2 = \frac{\sum x_1^2}{N_1 - 1} = \frac{N_1 \sum X_1^2 - (\sum X_1)^2}{N_1(N_1-1)} \qquad s_2^2 = \frac{\sum x_2^2}{N_2 - 1} = \frac{N_2 \sum X_2^2 - (\sum X_2)^2}{N_2(N_2-1)}$$

Ίσες Διασπορές
$(\sigma_1^2 = \sigma_2^2)$
Άνισα Ανεξάρτητα δείγματα $(N_1 \neq N_2)$

Βήμα 3ο - **Συνδυασμένη** διασπορά (pooled variance)

$$s^2 = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{(N_1 - 1) + (N_2 - 1)}$$

Βήμα 4ο - Τυπικό σφάλμα της διαφοράς (των 2 μέσων)

$$SE_D = \sqrt{\frac{s^2}{N_1} + \frac{s^2}{N_2}} = s\sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

Βήμα 5ο - Κριτήριο $t = \frac{\bar{X}_1 - \bar{X}_2}{SE_D}$

Βήμα 6ο - Έλεγχος της στατιστικής σημαντικότητας του t (Πίν 21.Δ):

με πιθανότητα α & βαθμούς ελευθερίας df = $(N_1 - 1) + (N_2 - 1) = N_1 + N_2 - 2$

η κρίσιμη τιμή t είναι $t_c = t_{\alpha, df}$ οπότε

- αν $|t| \geq t_c$ απορρίπτουμε την H_0 : η διαφορά είναι στατιστικώς σημαντική,
- αν $|t| < t_c$ αποδεχόμαστε την H_0 : η διαφορά είναι στατιστικώς μη σημαντική

Βήμα 7ο - **Μέγεθος επίδρασης (effect size)**, Διαφοράς:

$$d = (\bar{X}_1 - \bar{X}_2)/S, \text{ όπου S η συνδυασμένη τυπική απόκλιση } = \sqrt{s^2}.$$

24

24

Μη Αθλητές	X_1	X_1^2
1	3	9
2	4	16
3	5	25
4	5	25
5	5	25
6	6	36
7	7	49
$\Sigma_1 =$	35	185
Αθλητές	X_2	X_2^2
1	2	4
2	3	9
3	3	9
4	5	25
5	2	4
6	3	9
$\Sigma_2 =$	18	60

Π.χ., Ίσες Διασπορές ($\sigma_1^2 = \sigma_2^2$)

Άνισα Ανεξάρτητα δείγματα ($N_1 \neq N_2$)

Βήμα 1ο - Μέσοι $\bar{X}_1 = \frac{35}{7} = 5 \qquad \bar{X}_2 = \frac{18}{6} = 3$

Βήμα 2ο - Αμερόληπτες διασπορές

$$s_1^2 = \frac{7(185) - (35)^2}{7(7\text{-}1)} = 1.6667 \qquad s_2^2 = \frac{6(60) - (18)^2}{6(6\text{-}1)} = 1.2$$

Βήμα 3ο - **Συνδυασμένη διασπορά**

$$s^2 = \frac{(7-1)(1.6667) + (6-1)(1.2)}{(7\text{-}1) + (6\text{-}1)} = 1.4546$$

Βήμα 4ο - Τυπικό σφάλμα της διαφοράς (των 2 μέσων)

$$SE_D = \sqrt{\frac{1.4546}{7} + \frac{1.4546}{6}} = 0.671$$

Βήμα 5ο - κριτήριο t: $t = \frac{5-3}{0.671} = \frac{2}{0.671} = 2.981$

25

Critical Values for Student's *t* Distribution

one-tail area	0.250	0.125	0.100	0.075	0.050	0.025	0.010	0.005	0.0005
two-tail area	0.500	0.250	0.200	0.150	0.100	0.050	0.020	0.010	0.0010
d.f. \ *c*	0.500	0.750	0.800	0.850	0.900	0.950	0.980	0.990	0.999
1	1.000	2.414	3.078	4.165	6.314	12.706	31.821	63.657	636.619
2	0.816	1.604	1.886	2.282	2.920	4.303	6.965	9.925	31.599
3	0.765	1.423	1.638	1.924	2.353	3.182	4.541	5.841	12.924
4	0.741	1.344	1.533	1.778	2.132	2.776	3.747	4.604	8.610
5	0.727	1.301	1.476	1.699	2.015	2.571	3.365	4.032	6.869
6	0.718	1.273	1.440	1.650	1.943	2.447	3.143	3.707	5.959
7	0.711	1.254	1.415	1.617	1.895	2.365	2.998	3.499	5.408
8	0.706	1.240	1.397	1.592	1.860	2.306	2.896	3.355	5.041
9	0.703	1.230	1.383	1.574	1.833	2.262	2.821	3.250	4.781
10	0.700	1.221	1.372	1.559	1.812	2.228	2.764	3.169	4.587
11	0.697	1.214	1.363	1.548	1.796	2.201	2.718	3.106	4.437
12	0.695	1.209	1.356	1.538	1.782	2.179	2.681	3.055	4.318
13	0.694	1.204	1.350	1.530	1.771	2.160	2.650	3.012	4.221
14	0.692	1.200	1.345	1.523	1.761	2.145	2.624	2.977	4.140
15	0.691	1.197	1.341	1.517	1.753	2.131	2.602	2.947	4.073
16	0.690	1.194	1.337	1.512	1.746	2.120	2.583	2.921	4.015
17	0.689	1.191	1.333	1.508	1.740	2.110	2.567	2.898	3.965
18	0.688	1.189	1.330	1.504	1.734	2.101	2.552	2.878	3.922
19	0.688	1.187	1.328	1.500	1.729	2.093	2.539	2.861	3.883
20	0.687	1.185	1.325	1.497	1.725	2.086	2.528	2.845	3.850
21	0.686	1.183	1.323	1.494	1.721	2.080	2.518	2.831	3.819
22	0.686	1.182	1.321	1.492	1.717	2.074	2.508	2.819	3.792
23	0.685	1.180	1.319	1.489	1.714	2.069	2.500	2.807	3.768
24	0.685	1.179	1.318	1.487	1.711	2.064	2.492	2.797	3.745
25	0.684	1.198	1.316	1.485	1.708	2.060	2.485	2.787	3.725
26	0.684	1.177	1.315	1.483	1.706	2.056	2.479	2.779	3.707
27	0.684	1.176	1.314	1.482	1.703	2.052	2.473	2.771	3.690
28	0.683	1.175	1.313	1.480	1.701	2.048	2.467	2.763	3.674
29	0.683	1.174	1.311	1.479	1.699	2.045	2.462	2.756	3.659
30	0.683	1.173	1.310	1.477	1.697	2.042	2.457	2.750	3.646
35	0.682	1.170	1.306	1.472	1.690	2.030	2.438	2.724	3.591
40	0.681	1.167	1.303	1.468	1.684	2.021	2.423	2.704	3.551
45	0.680	1.165	1.301	1.465	1.679	2.014	2.412	2.690	3.520
50	0.679	1.164	1.299	1.462	1.676	2.009	2.403	2.678	3.496
60	0.679	1.162	1.296	1.458	1.671	2.000	2.390	2.660	3.460
70	0.678	1.160	1.294	1.456	1.667	1.994	2.381	2.648	3.435
80	0.678	1.159	1.292	1.453	1.664	1.990	2.374	2.639	3.416
100	0.677	1.157	1.290	1.451	1.660	1.984	2.364	2.626	3.390
500	0.675	1.152	1.283	1.442	1.648	1.965	2.334	2.586	3.310
1000	0.675	1.151	1.282	1.441	1.646	1.962	2.330	2.581	3.300
∞	0.674	1.150	1.282	1.440	1.645	1.960	2.326	2.576	3.291

For degrees of freedom *d.f.* not in the table, use the closest *d.f.* that is *smaller.*

Βήμα 6ο - Έλεγχος της στατιστικής σημαντικότητας του t

- με επίπεδο πιθανότητας $\alpha = 0.05$ (δίπλευρος έλεγχος) και
- βαθμούς ελευθερίας $df = N_1 + N_2 - 2 = 7 + 6 - 2 = 11$
- η κρίσιμη τιμή t είναι $t_c = t_{a,\, df} = t_{0.05,\, 11} = 2.201$

οπότε, επειδή $t = 2.981 > t_c = 2.201$, με πιθανότητα σφάλματος τύπου Ι ίση με 5% απορρίπτουμε τη μηδενική υπόθεση (H_0: $\mu_1 = \mu_2$) και αποδεχόμαστε την εναλλακτική (H_A: $\mu_1 < \mu_2$).

Στατιστικό συμπέρασμα: η διαφορά των 2 πληθυσμιακών μέσων σφαλμάτων χρόνου αντίδρασης είναι στατιστικώς σημαντική ($p < 0.05$).

Βήμα 7ο - Μέγεθος επίδρασης (effect size):
$d = (\bar{X}_1 - \bar{X}_2)/S = (5 - 3) / 1.2061 = 1.66$

Ερευνητικό συμπέρασμα: οι αθλητές κάνουν στατιστικώς σημαντικά λιγότερα σφάλματα χρόνου αντίδρασης από τους μη αθλητές, με το μέγεθος της διαφοράς (1.66) να είναι πολύ μεγάλο (Cohen, 1988, σελ. 25-27: 0.2 = μικρό, 0.5 = μέτριο, 0.8 = μεγάλο).

26

Άνισες Διασπορές ($\sigma_1^2 \neq \sigma_2^2$)

Βήμα 1ο – Μέσοι $\bar{X}_1 = \frac{\sum X_1}{N_1}$ $\quad \bar{X}_2 = \frac{\sum X_2}{N_2}$

Βήμα 2ο - Αμερόληπτες διασπορές

$$s_1^2 = \frac{\sum x_1^2}{N_1 - 1} = \frac{N_1 \sum X_1^2 - (\sum X_1)^2}{N_1(N_1 - 1)} \qquad s_2^2 = \frac{\sum x_2^2}{N_2 - 1} = \frac{N_2 \sum X_2^2 - (\sum X_2)^2}{N_2(N_2 - 1)}$$

Βήμα 3ο - Τυπικό σφάλμα της διαφοράς (των 2 μέσων)

$$SE_D = \sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}$$

Βήμα 4ο - Κριτήριο $\quad t = \frac{\bar{X}_1 - \bar{X}_2}{SE_D}$

Βήμα 5ο - Βαθμοί ελευθερίας (**df**$_c$) με τη μέθοδο **Welch** (1938)

$$\text{Welch } df = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{(s_1^2)^2}{N_1^2(N_1 - 1)} + \frac{(s_2^2)^2}{N_2^2(N_2 - 1)}}$$

όπου $(s_1^2/N_1) = (SE_1)^2$ η διασπορά σφάλματος του μέσου της 1ης ομάδας
$(s_1^2/N_2) = (SE_2)^2$ η διασπορά σφάλματος του μέσου της 2ης ομάδας.

Βήμα 6ο - Έλεγχος της στατιστικής σημαντικότητας του t (Πίν. 21.Δ):

με πιθανότητα α & βαθμούς ελευθερίας κατά Welch df$_c$
η κρίσιμη τιμή t είναι $t_c = t_{\alpha, dfc}$ οπότε

- αν $|t| \geq t_c$ απορρίπτουμε την H_0 : η διαφορά είναι στατιστικώς σημαντική,
- αν $|t| < t_c$ αποδεχόμαστε την H_0 : η διαφορά είναι στατιστικώς μη σημαντική.

Βήμα 7ο - **Μέγεθος επίδρασης (effect size)**, Διαφοράς:

$$d = (\bar{X}_1 - \bar{X}_2)/S, \text{ όπου S η συνδυασμένη τυπική απόκλιση } = \sqrt{S^2}$$

$$\sqrt{[(N_1-1)S_1^2 + (N_2-1)S_2^2]/(N_1 + N_2 - 2)}$$

27

27

Π.χ., Άνισες Διασπορές ($\sigma_1^2 \neq \sigma_2^2$)

Ομάδα 1	X_1	X_1^2
1	0	0
2	1	1
3	1	1
4	2	4
Σ_1 =	4	6
Ομάδα 2	X_2	X_2^2
1	3	9
2	3	9
3	4	16
4	4	16
5	7	49
6	9	81
Σ_2 =	30	180

Η παραδοχή των ίσων διασπορών (**$\sigma_1^2 = \sigma_2^2$**) εξετάστηκε με τον έλεγχο Levene και δεν επιβεβαιώθηκε: η διαφορά μεταξύ των 2 διασπορών ήταν στατιστικώς σημαντική ($F = 6.171$, $p = 0.038$).

Έτσι, ο έλεγχος t έγινε μετά από διόρθωση των βαθμών ελευθερίας με τη μέθοδο Welch, για να προσαρμοστεί η ανάλυση στη διαφορά των 2 διασπορών.

Βήμα 1ο - Μέσοι $\bar{X}_1 = \frac{4}{4} = 1 \qquad \bar{X}_2 = \frac{30}{6} = 5$

Βήμα 2ο - Αμερόληπτες διασπορές

$$s_1^2 = \frac{4(6) - 4^2}{4(4\text{-}1)} = 0.6667 \qquad s_2^2 = \frac{6(180) - 30^2}{6(6\text{-}1)} = 6$$

Βήμα 3ο - Τυπικό σφάλμα της διαφοράς (των 2 μέσων)

$$SE_D = \sqrt{\frac{0.6667}{4} + \frac{6}{6}} = 1.0801$$

Βήμα 4ο - Κριτήριο $\quad t = \frac{5 - 1}{1.0801} = 3.7034$

28

Critical Values for Student's *t* Distribution

one-tail area	0.250	0.125	0.100	0.075	0.050	0.025	0.010	0.005	0.0005
two-tail area	0.500	0.250	0.200	0.150	0.100	0.050	0.020	0.010	0.0010
d.f. \ *c*	0.500	0.750	0.800	0.850	0.900	0.950	0.980	0.990	0.999
1	1.000	2.414	3.078	4.165	6.314	12.706	31.821	63.657	636.619
2	0.816	1.604	1.886	2.282	2.920	4.303	6.965	9.925	31.599
3	0.765	1.423	1.638	1.924	2.353	3.182	4.541	5.841	12.924
4	0.741	1.344	1.533	1.778	2.132	2.776	3.747	4.604	8.610
5	0.727	1.301	1.476	1.699	2.015	2.571	3.365	4.032	6.869
6	0.718	1.273	1.440	1.650	1.943	2.447	3.143	3.707	5.959
7	0.711	1.254	1.415	1.617	1.895	2.365	2.998	3.499	5.408
8	0.706	1.240	1.397	1.592	1.860	2.306	2.896	3.355	5.041
9	0.703	1.230	1.383	1.574	1.833	2.262	2.821	3.250	4.781
10	0.700	1.221	1.372	1.559	1.812	2.228	2.764	3.169	4.587
11	0.697	1.214	1.363	1.548	1.796	2.201	2.718	3.106	4.437
12	0.695	1.209	1.356	1.538	1.782	2.179	2.681	3.055	4.318
13	0.694	1.204	1.350	1.530	1.771	2.160	2.650	3.012	4.221
14	0.692	1.200	1.345	1.523	1.761	2.145	2.624	2.977	4.140
15	0.691	1.197	1.341	1.517	1.753	2.131	2.602	2.947	4.073
16	0.690	1.194	1.337	1.512	1.746	2.120	2.583	2.921	4.015
17	0.689	1.191	1.333	1.508	1.740	2.110	2.567	2.898	3.965
18	0.688	1.189	1.330	1.504	1.734	2.101	2.552	2.878	3.922
19	0.688	1.187	1.328	1.500	1.729	2.093	2.539	2.861	3.883
20	0.687	1.185	1.325	1.497	1.725	2.086	2.528	2.845	3.850
21	0.686	1.183	1.323	1.494	1.721	2.080	2.518	2.831	3.819
22	0.686	1.182	1.321	1.492	1.717	2.074	2.508	2.819	3.792
23	0.685	1.180	1.319	1.489	1.714	2.069	2.500	2.807	3.768
24	0.685	1.179	1.318	1.487	1.711	2.064	2.492	2.797	3.745
25	0.684	1.198	1.316	1.485	1.708	2.060	2.485	2.787	3.725
26	0.684	1.177	1.315	1.483	1.706	2.056	2.479	2.779	3.707
27	0.684	1.176	1.314	1.482	1.703	2.052	2.473	2.771	3.690
28	0.683	1.175	1.313	1.480	1.701	2.048	2.467	2.763	3.674
29	0.683	1.174	1.311	1.479	1.699	2.045	2.462	2.756	3.659
30	0.683	1.173	1.310	1.477	1.697	2.042	2.457	2.750	3.646
35	0.682	1.170	1.306	1.472	1.690	2.030	2.438	2.724	3.591
40	0.681	1.167	1.303	1.468	1.684	2.021	2.423	2.704	3.551
45	0.680	1.165	1.301	1.465	1.679	2.014	2.412	2.690	3.520
50	0.679	1.164	1.299	1.462	1.676	2.009	2.403	2.678	3.496
60	0.679	1.162	1.296	1.458	1.671	2.000	2.390	2.660	3.460
70	0.678	1.160	1.294	1.456	1.667	1.994	2.381	2.648	3.435
80	0.678	1.159	1.292	1.453	1.664	1.990	2.374	2.639	3.416
100	0.677	1.157	1.290	1.451	1.660	1.984	2.364	2.626	3.390
500	0.675	1.152	1.283	1.442	1.648	1.965	2.334	2.586	3.310
1000	0.675	1.151	1.282	1.441	1.646	1.962	2.330	2.581	3.300
∞	0.674	1.150	1.282	1.440	1.645	1.960	2.326	2.576	3.291

For degrees of freedom *d.f.* not in the table, use the closest *d.f.* that is *smaller*.

Π.χ., Άνισες Διασπορές ($\sigma_1^2 \neq \sigma_2^2$)

Βήμα 5ο - Βαθμοί ελευθερίας (**df_c**) με τη μέθοδο Welch (1938)

$$\text{Welch } df = \frac{\left(\frac{0.6667}{4}+\frac{6}{6}\right)^2}{\frac{(0.6667)^2}{4^2(4-1)}+\frac{6^2}{6^2(6-1)}} = \frac{1.3611}{0.20926} = 6.5045$$

Βήμα 6ο - **Έλεγχος της στατιστικής σημαντικότητας** του t (Πίν.)

με επίπεδο πιθανότητας $\alpha = 0.05$ (δίπλευρος έλεγχος)

η κρίσιμη τιμή t για $df = 7$ είναι $t_c = t_{a,\, df} = t_{0.05,\, 7} = 2.365$ και

για $df = 6$ είναι $t_{0.05,\, 6} = 2.447$

οπότε η κρίσιμη τιμή t για $df_c = 6.5045$ είναι $\approx t_c =$

$t_{0.05,\, 6{,}5045} \approx 2.365 + (2.447 - 2.365)*(0.5045) \approx 2.406$

και επειδή $t = 3.70 > t_c = 2.406$, με πιθανότητα σφάλματος τύπου Ι ίση με 5% απορρίπτουμε τη μηδενική υπόθεση (H_0: $\mu_1 = \mu_2$) και αποδεχόμαστε την εναλλακτική (H_A: $\mu_1 < \mu_2$).

Στατιστικό συμπέρασμα: η διαφορά των 2 πληθυσμιακών μέσων της μαθησιακής ικανότητας X είναι στατιστικώς σημαντική ($p < 0.05$).

29

Ομάδα 1	X_1	X_1^2
1	0	0
2	1	1
3	1	1
4	2	4
Σ_1 =	4	6
Ομάδα 2	X_2	X_2^2
1	3	9
2	3	9
3	4	16
4	4	16
5	7	49
6	9	81
Σ_2 =	30	180

Π.χ., Άνισες Διασπορές ($\sigma_1^2 \neq \sigma_2^2$)

Στατιστικό συμπέρασμα: η διαφορά των 2 πληθυσμιακών μέσων της μαθησιακής ικανότητας X είναι στατιστικώς σημαντική ($p < 0.05$).

Βήμα 7ο - Μέγεθος επίδρασης (effect size), Διαφοράς:

με συνδυασμένη τυπική απόκλιση $S = \sqrt{\frac{3*0.6667+5*6}{3+5}} = \sqrt{\frac{32}{8}} = \sqrt{4} = 2$

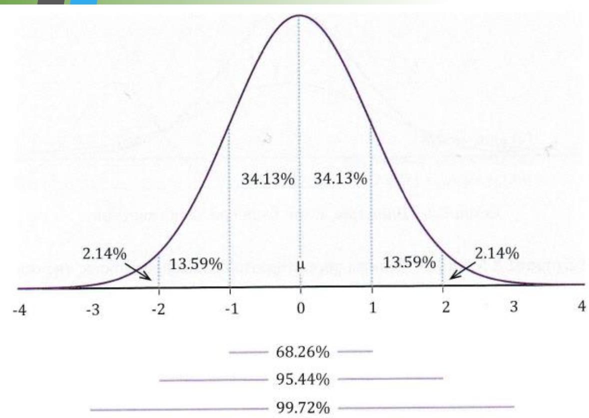
το μέγεθος της διαφοράς είναι: $d = (\bar{X}_2 - \bar{X}_1)/S = (5-1)/2 = 4/2 = 2$.

Ερευνητικό συμπέρασμα: οι μαθητές που εκπαιδεύτηκαν είχαν στατιστικώς σημαντικά υψηλότερη μάθηση από τους μαθητές που δεν εκπαιδεύτηκαν, με το μέγεθος της διαφοράς (2) να είναι πολύ μεγάλο. (Cohen, 1988, σελ. 26, μεγάλη διαφορά = 0.80).

30

30

Διάστημα Εμπιστοσύνης Διαφοράς 2 Μέσων



Σχήμα 8.3 - Κανονική κατανομή συχνοτήτων (%f) ανά τυπική απόκλιση (σ): χώροι (regions) πιθανοτήτων (αναλογιών) στο σύνολο του πληθυσμού (n)

- Αρχική αναφορά στην εκτίμηση διαστήματος (interval estimate) γίνεται στις προηγούμενες διαλέξεις, όπου είδαμε πως ορίζονται συγκεκριμένα διαστήματα (συμμετρικοί χώροι κάτω από την καμπύλη) τυπικών αποκλίσεων (σ), όπως, για παράδειγμα, το κεντρικό 68%, 95% και 99.7% του πληθυσμού (n).
- Τα διαστήματα αυτά ερμηνεύτηκαν ως πιθανότητες εμφάνισης μιας τυχαίας τιμής X.

31

31

Διάστημα Εμπιστοσύνης Διαφοράς 2 Μέσων

Το **διάστημα εμπιστοσύνης 95% του μέσου** (95% confidence interval for mean) εξηγήθηκε σε προηγούμενα κεφάλαια.

(95% CI): Mean $\pm$ SE_{mean} * $t_{a,df}$

Διάστημα εμπιστοσύνης 95% της μέσης διαφοράς (95% confidence interval for mean difference)

Διαφορά των 2 μέσων ($\bar{D}$) $\pm$ (**t** με δίπλευρο έλεγχο στο **α = 0.05** και με **N -1** βαθμούς ελευθερίας) * (Τυπικό Σφάλμα (**SE**) της Διαφοράς (D) των 2 μέσων)

$\bar{D} \pm t_{0.05,\,df} * (SE_D)$

32

32

33