t – test for dependent data
t - test για παρατηρήσεις κατά ζεύγη (paired t - test)

Ζιντζαράς Ηλίας, M.Sc., Ph.D.

Καθηγητής Βιομαθηματικών-Βιομετρίας
Εργαστήριο Βιομαθηματικών
Τμήμα Ιατρικής
Πανεπιστήμιο Θεσσαλίας

Institute for Clinical Research and Health Policy Studies
Tufts University School of Medicine
Boston, MA, USA

Θεόδωρος Μπρότσης, MSc, PhD
Εντεταλμένος Διδάσκων
(http://biomath.med.uth.gr)
Πανεπιστήμιο Θεσσαλίας
Email: tmprotsis@uth.gr

Μεταβολή αιμοσφαιρίνης

Έστω ότι θέλουμε να συγκρίνουμε την μεταβολή της αιμοσφαιρίνης μετά την χορήγηση ερυθροποιητίνης (EPO) σε 9 ασθενείς

Στο τέλος της μελέτης θέλουμε ελέγξουμε αν υπήρξε ή όχι σημαντική μεταβολή της αιμοσφαιρίνης, συγκρίνοντας τις τιμές πριν και μετά την χορήγηση ερυθροποιητίνης

Subject	Before	After
1	135	160
2	126	157
3	165	153
4	122	165
5	162	155
6	122	160
7	116	165
8	136	170
9	168	157

Πριν τη χορήγηση ερυθροποιητίνης	**Υπόθεση;**	Μετά τη χορήγηση ερυθροποιητίνης

Η μέση τιμή των διαφορών είναι μηδέν

$\mu_d = 0$

Subject 1	hgb_{before}	hgb_{after}
Subject 2	hgb_{before}	hgb_{after}
Subject 3	hgb_{before}	hgb_{after}
	⋮	⋮

Ανεξαρτησία;

Αυτές οι δύο μετρήσεις ΔΕΝ είναι ανεξάρτητες μεταξύ τους.

Είναι κατά ΖΕΥΓΗ.

Υπόθεση

Η **μηδενική υπόθεση** είναι ότι δεν υπάρχει μεταβολή στην αιμοσφαιρίνη πριν και μετά τη χορήγηση ερυθροποιητίνης. Η μέση τιμή των διαφορών είναι $= 0$

$$H_0: \mu_d = 0 \qquad\qquad H_a: \mu_d \neq 0$$

Επίπεδο σημαντικότητας: $a = 0.05$

Καθώς πρόκειται για παρατηρήσεις κατά ζεύγη με $\boldsymbol{n = 9}$, θα κάνουμε χρήση της $\boldsymbol{t - κατανομης}$ με $\boldsymbol{βαθμους\ ελευθεριας\ (degrees\ of\ freedom) = 8\ (n-1)}$

Κανόνας απόφασης:

Επειδή το $a = 0.05$ και επειδή κάνουμε χρήση της $t -$ κατανομής, η H_0 απορρίπτεται αν το στατιστικό τεστ είναι > 2.31 (5% σημείο της $t -$ κατανομής για $\boldsymbol{n - 1 = 8}$ βαθμούς ελευθερίας)

Περιγραφική Στατιστική

Κατά **μέσο όρο**, η μεταβολή στην **αιμοσφαιρίνη** είναι $\mathbf{21.11}$ μονάδες, η οποία είναι σημαντικά μεγαλύτερη από το **μηδέν** και συνεπώς υπάρχει ένδειξη ότι η αιμοσφαιρίνη (hgb) **αυξάνεται** λόγω της ερυθροποιητίνης (EPO)

Ωστόσο, η μεταβλητότητα είναι αρκετά μεγάλη ($\boldsymbol{CV = 115\%}$)**,** το μέγεθος του δείγματος είναι μικρό ($\boldsymbol{n = 9}$) και έτσι το σφάλμα της μέσης τιμής θεωρείται αρκετά μεγάλο ($\boldsymbol{SE = 8.11}$)

Subject	hgb_{before}	hgb_{after}	$hgb_{before} - hgb_{after}$
1	135	160	-25
2	126	157	-31
3	165	153	12
4	122	165	-43
5	162	155	7
6	122	160	-38
7	116	165	-49
8	136	170	-34
9	168	157	11
		N	9
		Average	-21.11
		Standard Deviation (SD)	24.34
		Coefficient of variation (CV)	-115
		Standard error ($SE = \frac{s}{\sqrt{n}}$)	8.11

Αξιολόγηση αποτελεσματικότητας φαρμάκου

Για να αξιολογήσουμε την αποτελεσματικότητα του φαρμάκου, θα πρέπει να απαντήσουμε στην παρακάτω ερώτηση:

Πόσο **βέβαιοι** είμαστε ότι η **μέση βελτίωση** (21.11) είναι **σημαντική**, δηλ. διαφορετική από το μηδέν;

Εναλλακτικά, πόσο **σίγουροι** είμαστε ότι η μέση βελτίωση δεν οφείλεται στην **τύχη**;

Subject	hgb_{before}	hgb_{after}	$hgb_{before} - hgb_{after}$
1	135	160	-25
2	126	157	-31
3	165	153	12
4	122	165	-43
5	162	155	7
6	122	160	-38
7	116	165	-49
8	136	170	-34
9	168	157	11
		N	9
		Average	-21.11
		Standard Deviation (SD)	24.34
		Coefficient of variation (CV)	-115
		Standard error ($SE = \frac{s}{\sqrt{n}}$)	8.11

Αξιολόγηση αποτελεσματικότητας φαρμάκου

Για να ελέγξουμε εάν η μέση βελτίωση (21.11) είναι σημαντική, πρέπει να εξετάσουμε την **μέση τιμή** σε συνδυασμό με τη **μεταβλητότητα** (τυπική απόκλιση) και το **μέγεθος** της δοκιμής, δηλαδή το **τυπικό σφάλμα**

Subject	hgb_{before}	hgb_{after}	$hgb_{before} - hgb_{after}$
1	135	160	-25
2	126	157	-31
3	165	153	12
4	122	165	-43
5	162	155	7
6	122	160	-38
7	116	165	-49
8	136	170	-34
9	168	157	11
		N	9
		Average	-21.11
		Standard Deviation (SD)	24.34
		Coefficient of variation (CV)	-115
		Standard error ($SE = \frac{s}{\sqrt{n}}$)	8.11

T-test

Στη συνέχεια, μπορούμε να εξετάσουμε αν η μέση διαφορά είναι σημαντική με τη χρήση του **t-test**:

$$t = \frac{average\ difference}{SE}$$

$$t = \frac{\bar{d}}{SE}$$

$$t = \frac{21.11}{8.11}$$

$$t = 2.60$$

Subject	hgb_{before}	hgb_{after}	$hgb_{before} - hgb_{after}$
1	135	160	-25
2	126	157	-31
3	165	153	12
4	122	165	-43
5	162	155	7
6	122	160	-38
7	116	165	-49
8	136	170	-34
9	168	157	11
		N	9
		Average	-21.11
		Standard Deviation (SD)	24.34
		Coefficient of variation (CV)	-115
		Standard error (SE $= \frac{s}{\sqrt{n}}$)	8.11

Αν πάρουμε την διαφορά (after - before) το πρόσημο αγνοείται

Σημαντικότητα

$$t = \frac{average\ difference}{SE} \qquad t = \frac{\bar{d}}{SE}$$

$$t = \frac{21.11}{8.11} \qquad t = 2.60$$

Τώρα, θα πρέπει να απαντήσουμε στην ακόλουθη ερώτηση:

Πόσο βέβαιοι είμαστε ότι το $\boldsymbol{t = 2.60}$ είναι σημαντικό;

Αν το $\boldsymbol{t = 2.60}$ είναι **σημαντικό** τότε καταλήγουμε στο συμπέρασμα ότι η μέση διαφορά είναι **σημαντική**

Subject	hgb_{before}	hgb_{after}	$hgb_{before} - hgb_{after}$
1	135	160	-25
2	126	157	-31
3	165	153	12
4	122	165	-43
5	162	155	7
6	122	160	-38
7	116	165	-49
8	136	170	-34
9	168	157	11
		N	**9**
		Average	**-21.11**
		Standard Deviation (SD)	**24.34**
		Coefficient of variation (CV)	**-115**
		Standard error ($\mathrm{SE} = \frac{s}{\sqrt{n}}$)	**8.11**

Σημαντικότητα

Στον πίνακα της $t-$κατανομής βρίσκουμε το $t_{\text{κρίσιμο σημείο}}$, για επίπεδο σημαντικότητας $a = 0.05$, για $n - 1 = 9 - 1 = 8$ βαθμούς ελευθερίας

Η τιμή αυτή είναι $\mathbf{2.31}$

Η τιμή $\mathbf{2.31}$ είναι το **κατώφλι/όριο (threshold)** για να αποφασίσουμε αν η τιμή του $t - test$ $(t = 2.60)$ είναι **σημαντική (significant)** για μία μελέτη με 8 βαθμούς ελευθερίας/df (9 subjects)

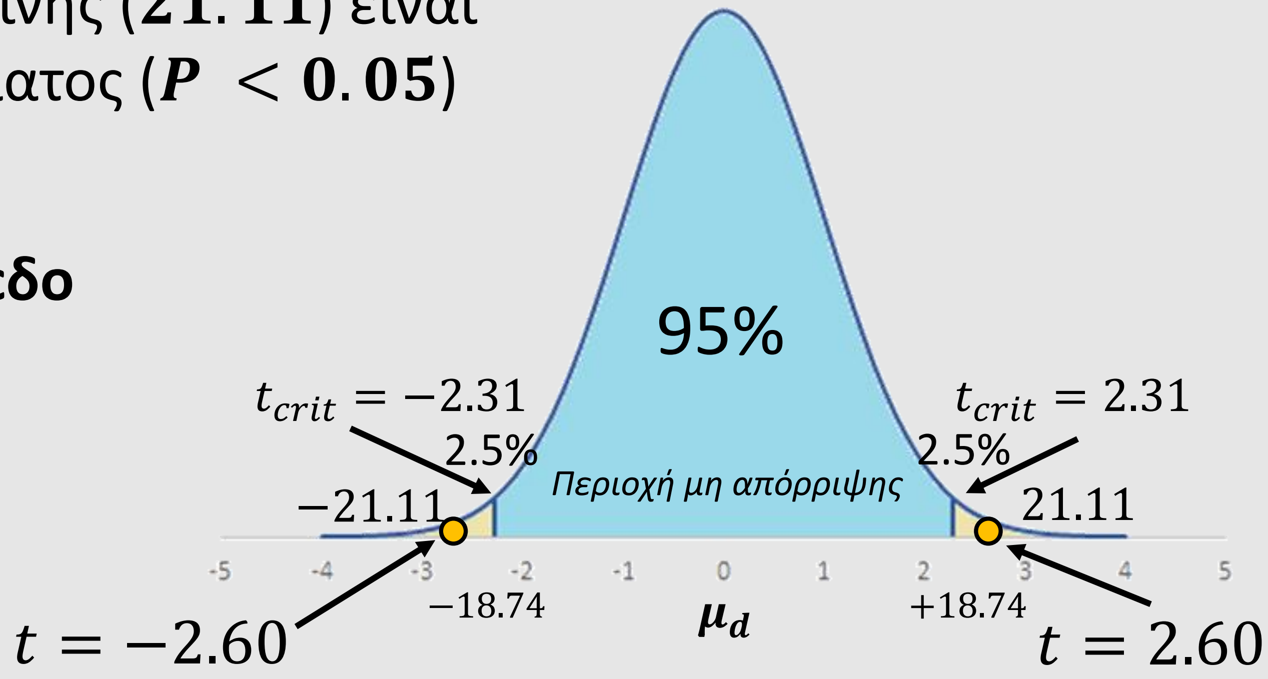
	Percentage points of the t distribution		
	p-value		
df (=n-1)	0.05	0.01	0.001
1	12.71	63.66	636.62
2	4.3	9.92	31.6
3	3.18	5.84	12.92
4	2.78	4.6	8.61
5	2.57	4.03	6.87
6	2.45	3.71	5.96
7	2.36	3.5	5.41
8	2.31	3.36	5.04
9	2.26	3.25	4.78
10	2.23	3.17	4.59
20	2.09	2.85	3.85
30	2.04	2.75	3.65
40	2.02	2.7	3.55
120	1.98	2.62	3.37
∞	1.96	2.58	3.29

P-Value

Το t-test $\boldsymbol{t} = \mathbf{2.60}$ είναι **μεγαλύτερο** από $\mathbf{2.31}$ (ή $t = -2.60$ είναι μικρότερο από -2.31)

Επομένως, η **μέση διαφορά** της αιμοσφαιρίνης ($\mathbf{21.11}$) είναι **σημαντική** με μία μικρή πιθανότητα σφάλματος ($\boldsymbol{P} < \mathbf{0.05}$)

Η τιμή p (**p-value**) ονομάζεται αλλιώς **επίπεδο σημαντικότητας (significant level)**



95 % διάστημα εμπιστοσύνης (Confidence Interval)

Το 95% CI (διάστημα εμπιστοσύνης) για τη μέση διαφορά μας δίνει το διάστημα που η **πραγματική μέση διαφορά** μπορεί να **βρίσκεται με 95% εμπιστοσύνη** (πιθανότητα)

$$(\text{μέση διαφορά} - \mathrm{t} \cdot \mathrm{SE}, \text{μέση διαφορά} + \ \mathrm{t} \cdot \mathrm{SE})$$

Το $\boldsymbol{t}$ είναι το 5% σημείο της $t -$ κατανομής για $n - 1 = 9 - 1 = 8$ βαθμούς ελευθερίας, η τιμή της οποίας είναι $\boldsymbol{t = 2.31}$

95 % διάστημα εμπιστοσύνης (Confidence Interval)

(μέση διαφορά - t · SE, μέση διαφορά + t · SE)

Έτσι, το 95% CI (διάστημα εμπιστοσύνης) για τη μέση τιμή των διαφορών είναι:

$$(21.11 - 2.31 \cdot 8.11, 21.11 + 2.31 \cdot 8.11)$$

ή

$$(2.40, 39.8)$$

Συμπέρασμα

Σύμφωνα με τα δεδομένα του δείγματος το 95% διάστημα εμπιστοσύνης για τη διαφορά μεταξύ των ζευγών είναι

$$2.40 - 39.8 \text{ g/L}$$

Έτσι, με 95% εμπιστοσύνη η πραγματική μέση τιμή των διαφορών βρίσκεται εντός αυτού του διαστήματος

Περιέχει αυτό το διάστημα την υποθετική μέση διαφορά που είναι ΜΗΔΕΝ;
ΟΧΙ. Το μηδέν βρίσκεται έξω από αυτό το διάστημα εμπιστοσύνης και έτσι η πραγματική μέση τιμή των διαφορών είναι πολύ διαφορετική (μακριά) από το μηδέν.

ΣΥΜΠΕΡΑΣΜΑ:
Απορρίπτουμε τη μηδενική υπόθεση H_0 ότι η μέση τιμή των πριν/μετά διαφορών της αιμοσφαιρίνης είναι **ΜΗΔΕΝ**. Το στατιστικό test και η $\bar{d}$ βρίσκονται πέρα από την t_{crit} για $df = 8$. Επομένως το φάρμακο (EPO) είναι αποτελεσματικό και η μεταβολή στην αιμοσφαιρίνη κυμαίνεται από 2.40 μέχρι 39.8 g/L