

Έλεγχος με την κατανομή t για ανεξάρτητα δείγματα (t test for independent samples)

1. Παράδειγμα

Θα δουλέψουμε με τα ακόλουθα 2 δείγματα μισθών καθηγητών σε δημόσια (PUBLIC) και ιδιωτικά (PRIVATE) πανεπιστήμια των ΗΠΑ ([Weiss, 2002](#)), που έχουν συλλεχθεί από το American Association of University Professors (σε χιλιάδες US\$):

INDEX	PUBLIC	PRIVATE
1	34.2	92.9
2	90	52
3	100.4	63.1
4	24.6	118.5
5	107.4	37.7
6	63.6	102.2
7	56.8	62.9
8	41.4	53.8
9	35	101
10	54.2	68.6
11	24.4	51.5
12	56	46.4
13	58.2	45.2
14	76.8	76
15	84.2	56.1
16	79.4	77.6
17	42.2	61.6
18	81.8	78.3
19	29.2	66.3
20	15.8	31.1
21	33.8	71.1
22	40.2	73.5
23	51.2	67.6
24	41.2	52.4
25	60.2	47.2
26	88.2	59.3
27	44.6	97.5
28	64.4	27.2
29	74	81.2
30	71	24.8
31	—	71
32	—	97.3
33	—	92.6
34	—	56
35	—	62.3

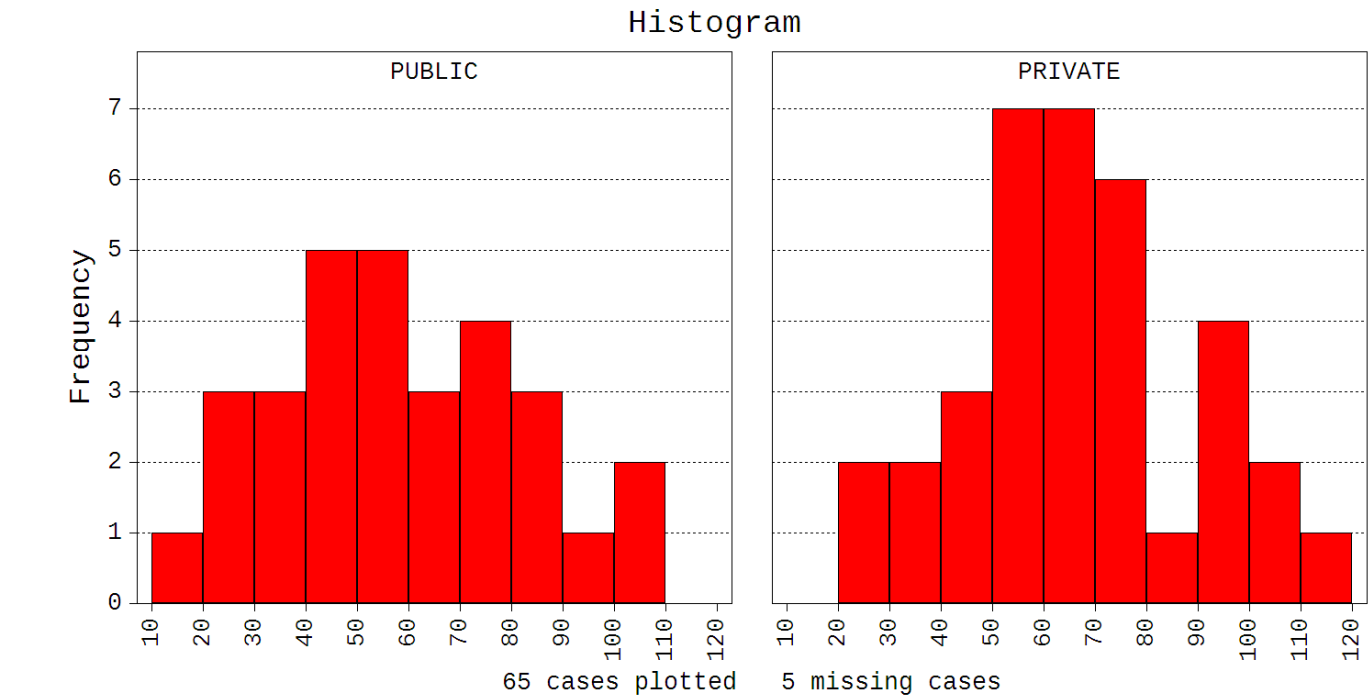
Τα περιγραφικά στατιστικά μεγέθη των ανωτέρω μεταβλητών έχουν ως ακολούθως

Descriptive Statistics

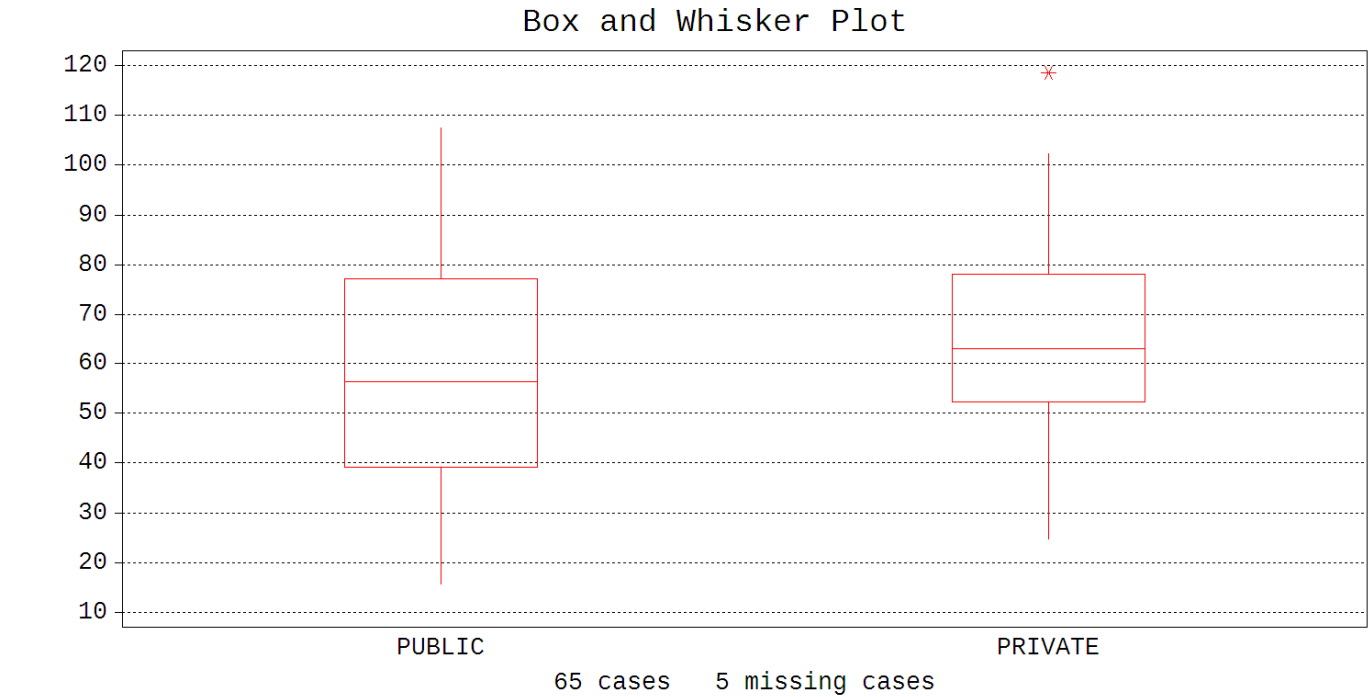
Variable	N	Mean	SD	Minimum	Maximum
PUBLIC	30	57.480	23.953	15.800	107.40
PRIVATE	35	66.394	22.261	24.800	118.50

όπου φαίνονται **μεγέθη δειγμάτων**, **μέσες τιμές** και **τυπικές αποκλίσεις**.

Τα ακόλουθα διπλανά ιστογράμματα δείχνουν την κατανομή των μεταβλητών:



Τέλος, τα ακόλουθα διπλανά box plots δείχνουν τη σχετική κατανομή των τιμών τους:



Το δεξί δείγμα φαίνεται να περιέχει μια ακραία τιμή (outlier) όμως το ιστόγραμμα του είναι OK.

Χρησιμοποιούμε τα ακόλουθα σύμβολα

μ_1 : μέσος μισθός καθηγητών σε δημόσια πανεπιστήμια

μ_2 : μέσος μισθός καθηγητών σε ιδιωτικά πανεπιστήμια

και επισημαίνουμε ότι τα δεδομένα συλλέχθηκαν προκειμένου να ελέγξουμε τις υποθέσεις

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

2. Δειγματοληπτική κατανομή

Υπενθυμίζουμε ότι η στατιστική θεωρία λέει ότι η δειγματοληπτική κατανομή των μέσων όρων δειγμάτων μεγέθους n , δηλαδή η κατανομή της μεταβλητής $\bar{x}$, έχει τα εξής χαρακτηριστικά:

- είναι κανονική,
- έχει μέσο όρο $\mu_{\bar{x}} = \mu$
- και τυπική απόκλιση ίση $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ (με το τυπικό σφάλμα).

Στο σημείο αυτό μαθαίνουμε ότι η δειγματοληπτική κατανομή της διαφοράς των μέσων όρων δειγμάτων n_1 και n_2 , δηλαδή η μεταβλητή $\bar{x}_1 - \bar{x}_2$, έχει τα ακόλουθα χαρακτηριστικά:

- είναι επίσης κανονική,
- έχει μέσο όρο $\mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$
- και τυπική απόκλιση $\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$.

Με βάση αυτά τα θεωρητικά ευρήματα, κατασκευάζουμε τους στατιστικούς ελέγχους που περιγράφονται στις επόμενες δυο ενότητες.

Στο φυλλάδιο αυτό δεν εξετάζουμε ελέγχους βασισμένους στην κανονική κατανομή (z) γιατί στερούνται πρακτικού ενδιαφέροντος.

3. Έλεγχος t με ίσες τυπικές αποκλίσεις (pooled t-test)

Στην ενότητα αυτή υποθέτουμε ότι οι τυπικές αποκλίσεις των δυο υπό έλεγχο πληθυσμών είναι ίσες.

Για την ορθή εφαρμογή του ελέγχου πρέπει να τηρούνται οι ακόλουθες προϋποθέσεις:

1. Τα δείγματα να είναι ανεξάρτητα.
2. Οι πληθυσμοί από τους οποίους προέρχονται τα δείγματα να είναι κανονικά κατανεμημένοι ή τα δείγματα να είναι μεγάλα.
3. Οι τυπικές αποκλίσεις των πληθυσμών να είναι ίσες.

Το στατιστικό μέτρο που χρησιμοποιείται για τον έλεγχο (test statistic) δίνεται από τον ακόλουθο τύπο

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

όπου με s_p συμβολίζουμε την κοινή τυπική απόκλιση (pooled standard deviation) που ισούται με

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Εάν υποθέσουμε ότι ορίζουμε επίπεδο σημαντικότητας (significance level) α , οι κρίσιμες τιμές του t βρίσκονται από τον πίνακα τιμών της κατανομής t ως εξής

- $\pm t_{\frac{\alpha}{2}}$ εάν πρόκειται για αμφίπλευρο έλεγχο
- $-t_{\alpha}$ εάν πρόκειται για αριστερόπλευρο (left-tailed) έλεγχο
- $+t_{\alpha}$ εάν πρόκειται για δεξιόπλευρο (right-tailed) έλεγχο

Οι βαθμοί ελευθερίας, df , στην περίπτωση αυτού του ελέγχου ισούνται με

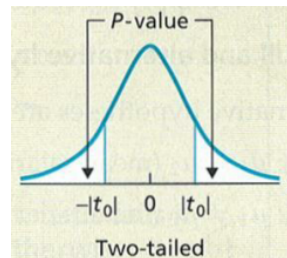
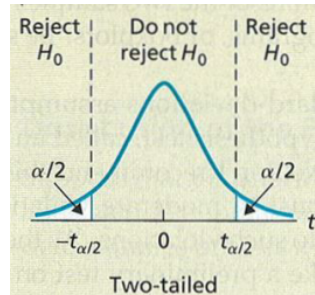
$$df = n_1 + n_2 - 2$$

Ας δούμε πως απεικονίζει ο **Weiss (2002)** τις περιοχές αποδοχής και απόρριψης καθώς και την αντίστοιχη πιθανότητα (p) σε διαφορετικές περιπτώσεις ελέγχων.

1. Αμφίπλευρος έλεγχος:

$H_a: \mu_1 \neq \mu_2$
(Two-tailed)

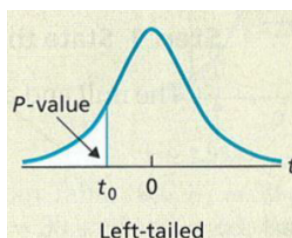
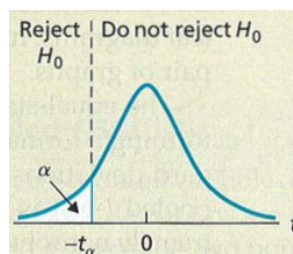
$\pm t_{\alpha/2}$
(Two-tailed)



2. Αριστερόπλευρος έλεγχος:

$H_a: \mu_1 < \mu_2$
(Left-tailed)

$-t_{\alpha}$
(Left-tailed)



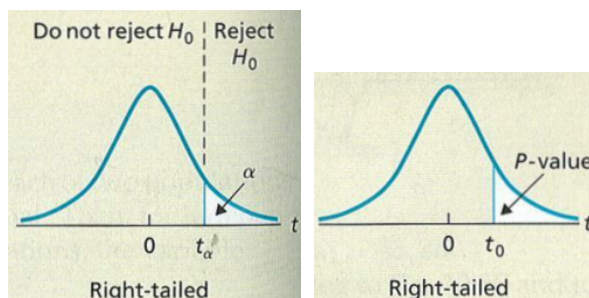
3. Δεξιόπλευρος έλεγχος:

$$H_a: \mu_1 > \mu_2$$

(Right-tailed)

$$t_\alpha$$

(Right-tailed)



3.1. Εφαρμογή στο παράδειγμα

Βασικά στατιστικά μεγέθη του δείγματος του πρώτου πληθυσμού:

$$\bar{x}_1 = 57.48$$

$$s_1 = 23.95$$

$$n_1 = 30$$

Βασικά στατιστικά μεγέθη του δείγματος του δεύτερου πληθυσμού:

$$\bar{x}_2 = 66.39$$

$$s_2 = 22.26$$

$$n_2 = 35$$

Όπως αναφέραμε προηγουμένως, στη συγκεκριμένη περίπτωση θέλουμε να ελέγξουμε κατά πόσον οι μέσοι όροι των πληθυσμών είναι διαφορετικοί, επομένως έχουμε αμφίπλευρο έλεγχο.

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

Οι προϋποθέσεις φαίνεται να τηρούνται:

1. Τα δυο δείγματα είναι πράγματι ανεξάρτητα το ένα από το άλλο.
2. Αν και δεν έχουμε γνώση των πληθυσμών, τα δυο δείγματα φαίνεται να είναι κανονικά κατανοημένα, όπως φαίνεται από τα ιστογράμματα που είδαμε προηγουμένως. Μάλιστα, όπως γράφει ο Weiss (2002), εφόσον τα δείγματα είναι μεγάλα (δηλαδή έχουν ≥ 30 δεδομένα έκαστο), αρκεί να επιβεβαιώσουμε ότι δεν περιέχουν ακραίες τιμές (outliers) και η μια ακραία τιμή που φαίνεται να περιέχει το ένα δείγμα (από τα box plots) δεν αποτελεί σπουδαίο στοιχείο.
3. Οι τυπικές αποκλίσεις των δυο δειγμάτων, 23.95 και 22.26, πράγματι φαίνεται να είναι περίπου ίσες.

Ας θεωρήσουμε επίπεδο εμπιστοσύνης (confidence level) 95%, που σημαίνει ότι έχουμε επίπεδο σημαντικότητας (significance level), α , ίσο με 100% μείον το επίπεδο εμπιστοσύνης:

$$\alpha = 100\% - 95\% = 5\% = 0.05$$

Για να υπολογίσουμε την τιμή του στατιστικού μέτρου t πρέπει πρώτα να υπολογίσουμε την s_p :

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(30 - 1)23.95^2 + (35 - 1)22.26^2}{30 + 35 - 2}} = 23.053$$

Οπότε:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{57.48 - 66.39}{23.053 \sqrt{\frac{1}{30} + \frac{1}{35}}} = -1.553$$

Σε βαθμούς ελευθερίας

$$df = 30 + 35 - 2 = 63$$

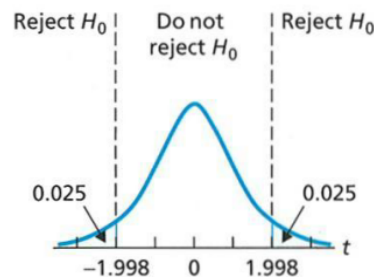
επίπεδο σημαντικότητας $\alpha=0.05$ και αμφίπλευρο έλεγχο:

df	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	df
			...			
63	1.295	1.669	1.998	2.387	2.656	63

αντιστοιχούν οι ακόλουθες κρίσιμες τιμές:

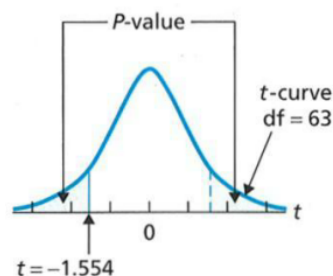
$$\pm t_{\frac{\alpha}{2}} = \pm t_{\frac{0.05}{2}} = \pm t_{0.025} = \pm 1.998$$

Παρατηρούμε ότι η τιμή -1.553 πέφτει εντός της περιοχής αποδοχής («Do not reject H_0 »)

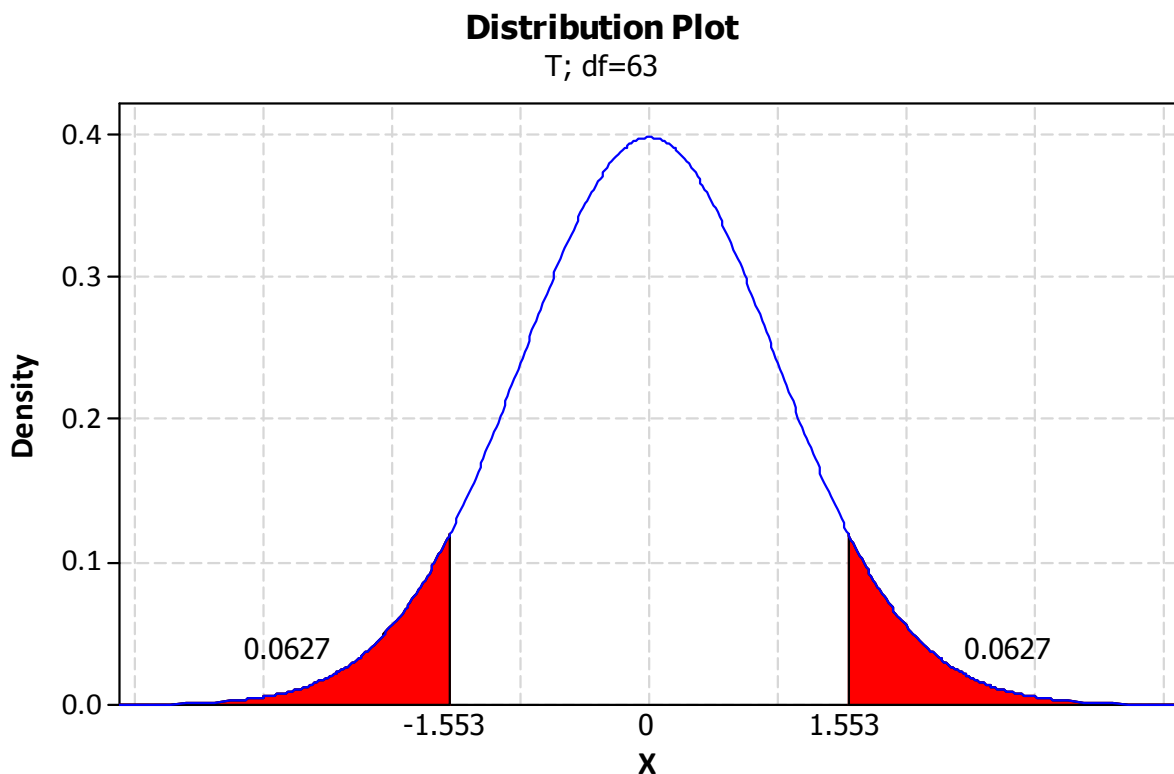


συνεπώς δεν απορρίπτουμε («αποδεχόμαστε» αλλά δεν το λέμε έτσι) την μηδενική υπόθεση σε επίπεδο εμπιστοσύνης 95% (που είναι το ίδιο σαν να λέμε σε επίπεδο σημαντικότητας 5% ή 0.05).

Εάν έχουμε πρόσβαση και σε υπολογιστή, υπολογίζουμε και την ακριβή τιμή της πιθανότητας που αντιστοιχεί στη τιμή t που βρήκαμε



Αυτή ισούται με $0.0627 + 0.0627 = 0.125$ όπως φαίνεται και από το επόμενο σχήμα του Minitab:



Εφόσον $p = 0.125 > 0.05$ και πάλι εξάγουμε το ίδιο συμπέρασμα, δηλαδή ότι δεν απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας 5%.

4. Έλεγχος t με άνισες τυπικές αποκλίσεις (nonpooled t-test)

Στην ενότητα αυτή υποθέτουμε ότι οι τυπικές αποκλίσεις των δυο υπό έλεγχο πληθυσμό είναι διάφορες.

ΥΠΟ ΚΑΤΑΣΚΕΥΗ

Βιβλιογραφικές αναφορές

Weiss, N.A. (2002): *Elementary Statistics*, 5th edition, Addison Wesley.