
Έλεγχος-t με εφαρμογές

Οι έλεγχοι-t (t-tests) χρησιμοποιούνται για να απαντήσουμε σε ερευνητικά ερωτήματα τα οποία μετατρέπονται σε έλεγχο υποθέσεων μέσω τιμών (βλ. σημειώσεις «Ανασκόπηση μερικών βασικών μεθόδων της Στατιστικής Συμπερασματολογίας») σε έναν ή δύο πληθυσμούς. Στη συνέχεια παρουσιάζονται ο έλεγχος-t για ένα δείγμα, ο έλεγχος-t για ανεξάρτητα δείγματα, και ο έλεγχος-t κατά ζεύγη δειγμάτων. Για τις εφαρμογές, οι έλεγχοι πραγματοποιούνται σε ε.σ. $\alpha = 5\%$.

1 Έλεγχος-t για ένα δείγμα (one-sample t-test)

Ένας φοιτητής του Τμήματος Μαθηματικών έχει ακούσει ότι ο μέσος χρόνος για την απόκτηση πτυχίου είναι 5,5 χρόνια. Επειδή δεν έχει πρόσβαση στα επίσημα στοιχεία της Γραμματείας, σκέφτηκε να πραγματοποιήσει μία έρευνα με

ερευνητικό ερώτημα (1): «Ο μέσος χρόνος απόκτησης πτυχίου στο Τμήμα Μαθηματικών είναι 5,5 χρόνια».

Αυτό το ερευνητικό ερώτημα μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu = 5,5 \text{ κατά } H_1 : \mu \neq 5,5,$$

όπου μ αναπαριστά το μέσο χρόνο απόκτησης πτυχίου στο Τμήμα Μαθηματικών (μέση τιμή πληθυσμού, δηλ. όλων των αποφοίτων). Για να προχωρήσει επομένως στον παραπάνω έλεγχο πρέπει να πάρει ένα δείγμα αποφοίτων καταγράφοντας το χρόνο που χρειάστηκε κάθε ένας από αυτούς μέχρι την απόκτηση του πτυχίου του.

Γενικά, για να χρησιμοποιήσουμε τον έλεγχο-t για ένα δείγμα, πρέπει να δούμε αν ικανοποιούνται τα εξής.

Προϋποθέσεις:

- να έχουμε ανεξάρτητες παρατηρήσεις,
- να ακολουθούν κανονική κατανομή.

Ο έλεγχος-t είναι ανεκτικός ή εύρωστος (robust) σε μικρές ως μέτριες παραβιάσεις της κανονικότητας.

Αρχείο δεδομένων

Για την εφαρμογή στο SPSS και στην R: **degree years.sav**

Εφαρμογή με το SPSS

Analyze → *Compare Means* → *One-Sample T Test*: στο *Test Variable* εισάγουμε την εξαρτημένη μεταβλητή (εδώ years) και στο *Test Value* την τιμή που θέλουμε να ελέγξουμε (εδώ 5,5), και πατάμε OK.

Στο Σχήμα 1 δίνονται τα παραγόμενα αποτελέσματα από το SPSS. Ένας επίσημος (επιστημονικός) τρόπος να παρουσιάσουμε τα αποτελέσματα της παραπάνω έρευνάς μας π.χ. σε μία εργασία, είναι ο εξής.

Ο έλεγχος-t έδειξε ότι ο μέσος χρόνος απόκτησης του πτυχίου στο Τμήμα Μαθηματικών ($M = 5,656$, $SD = 1,088$) δεν είναι στατιστικά σημαντικά διαφορετικός από 5,5 χρόνια, $t(31) = 0,812$, $p\text{-value} = 0,423$.

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
Time to get the Degree	32	5,656	1,0883	,1924

One-Sample Test						
Test Value = 5.5						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Time to get the Degree	,812	31	,423	,1563	-,236	,549

Σχήμα 1: Αποτελέσματα SPSS.

Σημειώστε ότι στο παραπάνω συμπέρασμά μας για την σημαντικότητα¹ του ελέγχου, εκτός από την τιμή του p-value, αναφέρουμε και την παρατηρούμενη τιμή της ελεγχουσυνάρτησης t με τους βαθμούς ελευθερίας εντός παρενθέσεων, καθώς και τις τιμές της (δειγματικής) μέσης τιμής και τυπικής απόκλισης.

Το SPSS εκτελεί δίπλευρους ελέγχους-t όπως ο παραπάνω. Αν θέλουμε να εκτελέσουμε κάποιον από τους μονόπλευρους ελέγχους στηρίζομαστε στα αποτελέσματα του δίπλευρου. Έτσι, αν το ερευνητικό ερώτημα που είχαμε θέσει εξ αρχής ήταν

ερευνητικό ερώτημα (2): «Ο μέσος χρόνος απόκτησης πτυχίου στο Τμήμα Μαθηματικών είναι περισσότερο από 5,5 χρόνια»,

το οποίο μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu \leq 5,5 \quad \text{κατά} \quad H_1 : \mu > 5,5,$$

το μόνο που αλλάζει είναι ο υπολογισμός του p-value. Σε αυτή την περίπτωση είναι

$$\text{p-value} = \frac{\text{p-value δίπλευρου}}{2} = \frac{0,423}{2} = 0,2115,$$

και θα παρουσιάζαμε τα αποτελέσματα ως εξής. Ο έλεγχος-t έδειξε ότι ο μέσος χρόνος απόκτησης του πτυχίου ($M = 5,656$, $SD = 1,088$) δεν είναι στατιστικά σημαντικά μεγαλύτερος από 5,5 χρόνια, $t(31) = 0,812$, p-value = 0,2115.

Από την άλλη πλευρά, αν το ερευνητικό ερώτημα που είχαμε θέσει εξ αρχής ήταν

ερευνητικό ερώτημα (3): «Ο μέσος χρόνος απόκτησης πτυχίου στο Τμήμα Μαθηματικών είναι λιγότερο από 5,5 χρόνια»,

το οποίο μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu \geq 5,5 \quad \text{κατά} \quad H_1 : \mu < 5,5,$$

το p-value σε αυτή την περίπτωση είναι

$$\text{p-value} = 1 - \frac{\text{p-value δίπλευρου}}{2} = 1 - \frac{0,423}{2} = 0,7885,$$

και θα παρουσιάζαμε τα αποτελέσματα ως εξής. Ο έλεγχος-t έδειξε ότι ο μέσος χρόνος απόκτησης του πτυχίου ($M = 5,656$, $SD = 1,088$) δεν είναι στατιστικά σημαντικά μικρότερος από 5,5 χρόνια, $t(31) = 0,812$, p-value = 0,7885.

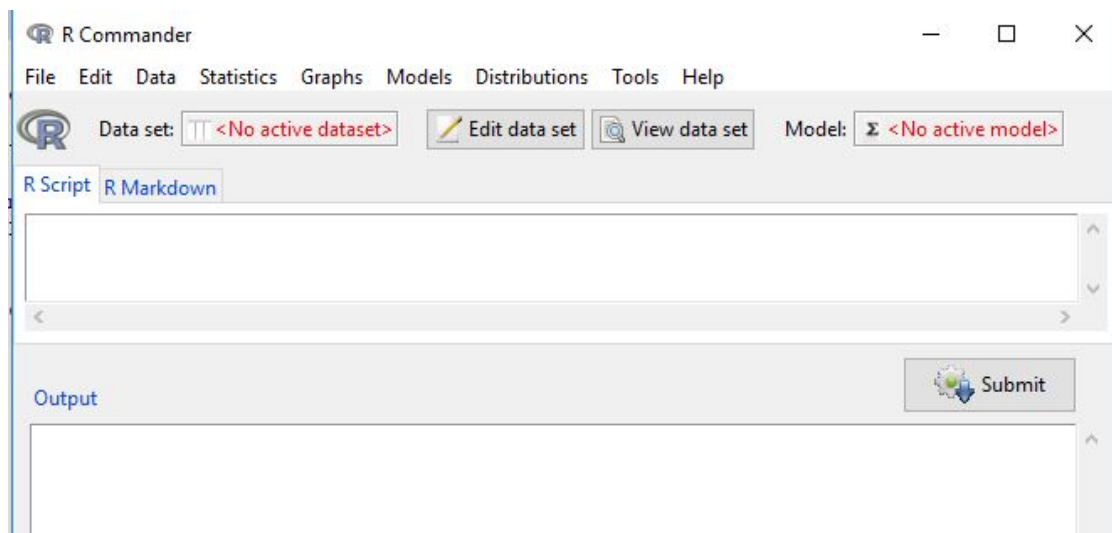
Ιδιαίτερη προσοχή χρειάζεται κάθε φορά που θέλουμε να εκτελέσουμε κάποιον μονόπλευρο έλεγχο-t αν θα πάρουμε (p-value δίπλευρου)/2, ή αν θα πάρουμε $1 - (\text{p-value δίπλευρου})/2$.

¹Ένας έλεγχος αναφέρεται ως *στατιστικά σημαντικός* (significant) στην περίπτωση που απορρίπτουμε την μηδενική υπόθεση, δηλ. αν p-value < α, διαφορετικά αναφέρεται ως *στατιστικά μη σημαντικός*.

Εφαρμογή με την R

Η συνάρτηση που εκτελεί ελέγχους-t στην R είναι η *t.test()*. Σημειώνεται ότι εδώ έχουμε τη δυνατότητα να εκτελέσουμε άμεσα και τους μονόπλευρους ελέγχους.

Στη συνέχεια παρουσιάζουμε την διαδικασία μέσα από ένα γραφικό περιβάλλον που παρέχεται με το πακέτο R Commander καλώντας το από την κονσόλα της R με την εντολή *library('Rcmdr')*. Πρέπει να σημειωθεί ότι αν δεν υπάρχει ήδη στο σύστημά μας, χρησιμοποιούμε πρώτα τη διαδικασία *install.packages('Rcmdr')* ώστε να το εγκαταστήσουμε. Ο ενδιαφερόμενος αναγνώστης μπορεί να μελετήσει τις εντολές της R που εμφανίζονται στο *R Script* για κάθε βήμα που εκτελούμε στο R Commander (βλ. Σχήμα 2).



Σχήμα 2: Περιβάλλον του R Commander.

Για να προχωρήσουμε στην ανάλυση πρέπει πρώτα να κάνουμε εισαγωγή τα δεδομένα μας. Από *Data* → *Import Data* → *from SPSS data set:* στο *Enter name for data set* δίνουμε μια ονομασία για τα δεδομένα που θέλουμε να εισαγάγουμε, και πατάμε *OK*, ενώ στη συνέχεια επιλέγουμε το αρχείο δεδομένων μας από το φάκελο στον οποίο το έχουμε αποθηκεύσει (εδώ *degree years.sav*). Υπάρχει η δυνατότητα να επεξεργαστούμε τα δεδομένα (*Edit data set*) ή απλά να τα δούμε (*View data set*).

Για να εκτελέσουμε τον έλεγχο-t, *Statistics* → *Means* → *Single-sample t-test:* στο *null hypothesis mu* την τιμή που θέλουμε να ελέγχουμε (εδώ 5,5), ενώ μπορούμε να επιλέξουμε αν θα εκτελέσουμε τον δίπλευρο έλεγχο ή κάποιον από τους μονόπλευρους, και πατάμε *OK*. Τα αποτελέσματα εμφανίζονται στο *Output* του R Commander, όπως φαίνονται στο Σχήμα 3.

2 Έλεγχος-t για ανεξάρτητα δείγματα (independent-samples t-test)

Ο φοιτητής του Τμήματος Μαθηματικών σκέφτηκε να πραγματοποιήσει μία έρευνα για να απαντήσει στο εξής ερευνητικό ερώτημα.

ερευνητικό ερώτημα (1): «Ο μέσος χρόνος απόκτησης πτυχίου στο Τμήμα Μαθηματικών για τους άνδρες και τις γυναίκες απόφοιτους είναι διαφορετικός».

Αυτό το ερευνητικό ερώτημα μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu_1 = \mu_2 \quad \text{κατά} \quad H_1 : \mu_1 \neq \mu_2,$$

όπου μ_1 , μ_2 αναπαριστούν το μέσο χρόνο απόκτησης πτυχίου στο Τμήμα Μαθηματικών ανδρών και γυναικών αποφοίτων, αντίστοιχα (μέσες τιμές πληθυσμών, δηλ. ανδρών και γυναικών αποφοίτων,

```

One Sample t-test

data:  years
t = 0.8121, df = 31, p-value = 0.4229
alternative hypothesis: true mean is not equal to 5.5
95 percent confidence interval:
 5.263863 6.048637
sample estimates:
mean of x
 5.65625
(a)

One Sample t-test

data:  years
t = 0.8121, df = 31, p-value = 0.2115
alternative hypothesis: true mean is greater than 5.5
95 percent confidence interval:
 5.330045      Inf
sample estimates:
mean of x
 5.65625
(β)

One Sample t-test

data:  years
t = 0.8121, df = 31, p-value = 0.7885
alternative hypothesis: true mean is less than 5.5
95 percent confidence interval:
 -Inf 5.982455
sample estimates:
mean of x
 5.65625
(γ)

```

Σχήμα 3: Αποτελέσματα R για (α) ερευνητικό ερώτημα (1), (β) ερευνητικό ερώτημα (2), (γ) ερευνητικό ερώτημα (3).

αντίστοιχα). Για να προχωρήσει επομένως στον παραπάνω έλεγχο πρέπει να πάρει ένα δείγμα αποφοίτων καταγράφοντας και το φύλο εκτός από το χρόνο που χρειάστηκε κάθε ένας από αυτούς μέχρι την απόκτηση του πτυχίου του.

Γενικά, για να χρησιμοποιήσουμε τον έλεγχο-t για ανεξάρτητα δείγματα, πρέπει να δούμε αν ικανοποιούνται τα εξής.

Προϋποθέσεις:

- να έχουμε ανεξάρτητες παρατηρήσεις (πέραν των ανεξάρτητων δειγμάτων/ομάδων),
- να ακολουθούν κανονικές κατανομές ανά ομάδα.

Θυμίζουμε ότι ο έλεγχος-t είναι ανεκτικός σε μικρές ως μέτριες παραβιάσεις της κανονικότητας.

Το παραπάνω ερευνητικό ερώτημα μπορεί να διατυπωθεί ισοδύναμα ως εξής.

ερευνητικό ερώτημα (ισοδύναμο) : «Το φύλο είναι παράγοντας που επηρεάζει το χρόνο απόκτησης πτυχίου στο Τμήμα Μαθηματικών».

Αυτό που γίνεται ξεκάθαρο από την ισοδύναμη διατύπωση του ερευνητικού ερωτήματος είναι ότι στην ουσία με αυτόν τον έλεγχο-t εξετάζουμε την επίδραση ενός παράγοντα (factor), εδώ το φύλο, με δύο επίπεδα (levels), εδώ άνδρας και γυναίκα, (δηλ. πρόκειται για μια κατηγορική μεταβλητή με δύο

κατηγορίες) που έχει το ρόλο της ανεξάρτητης μεταβλητής, πάνω σε μία ποσοτική μεταβλητή, εδώ χρόνος απόκτησης πτυχίου, που έχει το ρόλο της εξαρτημένης μεταβλητής².

Αρχείο δεδομένων

Για την εφαρμογή στο SPSS και στην R: **degree years.sav**

Εφαρμογή με το SPSS

Analyze → Compare Means → Independent-Samples T Test: στο *Test Variable* εισάγουμε την εξαρτημένη μεταβλητή (εδώ years) και στο *Grouping Variable* την ανεξάρτητη μεταβλητή (εδώ gender), ενώ στο *Define Groups* εισάγουμε τις τιμές που αντιστοιχούν στις δύο ομάδες που θέλουμε να συγκρίνουμε (εδώ 1 και 2, που αντιστοιχούν σε άνδρα και γυναίκα) και πατάμε OK.

Group Statistics				
	gender	N	Mean	Std. Deviation
Time to get the Degree	male	16	6,000	1,0954
	female	16	5,313	,9979

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
Time to get the Degree	Equal variances assumed	,093	,763	1,856	30	,073	,6875	,3705	Lower: -,0691 Upper: 1,4441
	Equal variances not assumed			1,856	29,743	,073	,6875	,3705	Lower: -,0694 Upper: 1,4444

Σχήμα 4: Αποτελέσματα SPSS.

Στο Σχήμα 4 δίνονται τα παραγόμενα αποτελέσματα από το SPSS. Η πρώτη γραμμή στον πίνακα Independent Samples Test είναι ο έλεγχος-t για ανεξάρτητα δείγματα στην περίπτωση που μπορούμε να θεωρήσουμε ότι έχουμε ίδια διασπορά στους δύο πληθυσμούς ($\sigma_1^2 = \sigma_2^2 = \sigma^2$). Με το Levene's test (βλ. Παρατήρηση 1 παρακάτω) εξετάζουμε αν μπορούμε να θεωρήσουμε ίσες διασπορές στους δύο πληθυσμούς. Η δεύτερη γραμμή είναι ο έλεγχος-Welch's t με τους «διορθωμένους» βαθμούς ελευθερίας ώστε να αντισταθμιστεί ενδεχόμενη ανισότητα διασπορών στους δύο πληθυσμούς. Ένας τρόπος να παρουσιάσουμε τα αποτελέσματα της έρευνάς μας είναι ο εξής.

Ο έλεγχος-t για ανεξάρτητα δείγματα έδειξε ότι δεν υπάρχει στατιστικά σημαντική διαφορά ανάμεσα στους μέσους χρόνους απόκτησης του πτυχίου των ανδρών ($M = 6$, $SD = 1,095$) και των γυναικών ($M = 5,313$, $SD = 0.998$), $t(30) = 1,856$, $p\text{-value} = 0,073$, ή

ο έλεγχος-t για ανεξάρτητα δείγματα έδειξε ότι το φύλο δεν είναι στατιστικά σημαντικός παράγοντας για το χρόνο απόκτησης πτυχίου (άνδρες: $M = 6$, $SD = 1,095$, γυναίκες: $M = 5,313$, $SD = 0.998$), $t(30) = 1,856$, $p\text{-value} = 0,073$.

Αν θέλουμε να εκτελέσουμε κάποιον από τους μονόπλευρους ελέγχους στηρίζομαστε στα αποτελέσματα του δίπλευρου. Έτσι, αν το ερευνητικό ερώτημα που είχαμε θέσει εξ αρχής ήταν

ερευνητικό ερώτημα (2): «Ο μέσος χρόνος απόκτησης πτυχίου των ανδρών στο Τμήμα Μαθηματικών είναι μεγαλύτερος από των γυναικών»,

το οποίο μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu_1 \leq \mu_2 \quad \text{κατά} \quad H_1 : \mu_1 > \mu_2,$$

²Με άλλα λόγια, σε ένα πείραμα, η μεταβλητή που θεωρούμε ότι συνιστά το *αίτιο* καλείται ανεξάρτητη μεταβλητή, με την έννοια ότι οι τιμές που παίρνει δεν εξαρτώνται από άλλες μεταβλητές. Η μεταβλητή που θεωρούμε ότι συνιστά το *αποτέλεσμα* ονομάζεται εξαρτημένη μεταβλητή, με την έννοια ότι οι τιμές της εξαρτώνται από το αίτιο, δηλαδή την ανεξάρτητη μεταβλητή.

το p-value σε αυτή την περίπτωση είναι

$$p\text{-value} = \frac{\text{p-value δίπλευρου}}{2} = \frac{0,073}{2} = 0,0365,$$

και θα παρουσιάζαμε τα αποτελέσματα ως εξής. Ο έλεγχος-t για ανεξάρτητα δείγματα έδειξε ότι ο μέσος χρόνος απόκτησης του πτυχίου των ανδρών ($M = 6, SD = 1,095$) είναι στατιστικά σημαντικά μεγαλύτερος από των γυναικών ($M = 5,313, SD = 0.998$), $t(30) = 1,856$, $p\text{-value} = 0,0365$.

Από την άλλη πλευρά, αν το ερευνητικό ερώτημα που είχαμε θέσει εξ αρχής ήταν

ερευνητικό ερώτημα (3): «Ο μέσος χρόνος απόκτησης πτυχίου των ανδρών στο Τμήμα Μαθηματικών είναι μικρότερος από των γυναικών»,

το οποίο μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu_1 \geq \mu_2 \quad \text{κατά} \quad H_1 : \mu_1 < \mu_2,$$

το p-value σε αυτή την περίπτωση είναι

$$p\text{-value} = 1 - \frac{\text{p-value δίπλευρου}}{2} = 1 - \frac{0,073}{2} = 0,9635,$$

και θα παρουσιάζαμε τα αποτελέσματα ως εξής. Ο έλεγχος-t για ανεξάρτητα δείγματα έδειξε ότι ο μέσος χρόνος απόκτησης του πτυχίου των ανδρών ($M = 6, SD = 1,095$) δεν είναι στατιστικά σημαντικά μικρότερος από των γυναικών ($M = 5,313, SD = 0.998$), $t(30) = 1,856$, $p\text{-value} = 0,9665$.

Θυμίζουμε ότι χρειάζεται ιδιαίτερη προσοχή κάθε φορά που θέλουμε να εκτελέσουμε κάποιον μονόπλευρο έλεγχο-t αν θα πάρουμε $(p\text{-value δίπλευρου})/2$, ή αν θα πάρουμε $1 - (p\text{-value δίπλευρου})/2$.

Παρατήρηση 1 (Levene's test). Το Levene's test για τον έλεγχο ισότητας διασπορών σε πληθυσμούς είναι πιο γενικό από τον έλεγχο-F για το λόγο δύο διασπορών με την έννοια ότι μπορεί να εφαρμοστεί για τον έλεγχο ισότητας διασπορών σε περισσότερους από δύο πληθυσμούς. Έτσι, αν Y_{ij} είναι η j παρατήρηση από την i ομάδα, $i = 1, \dots, k$, $j = 1, \dots, n_i$, με $n = \sum_{i=1}^k n_i$, $\bar{Y}_i$ ο δειγματικός μέσος της i ομάδας, $D_{ij} = |Y_{ij} - \bar{Y}_i|$, $\bar{D}_i = (1/n_i) \sum_{j=1}^{n_i} D_{ij}$, $\bar{D} = (1/n) \sum_{i=1}^k \sum_{j=1}^{n_i} D_{ij}$, τότε

$$\mathcal{F}_{Levene} = \frac{\sum_{i=1}^k n_i (\bar{D}_i - \bar{D})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (D_{ij} - \bar{D}_i)^2 / (n-k)} \sim \mathcal{F}_{k-1, n-k},$$

ενώ απορρίπτουμε τη μηδενική υπόθεση για ισότητα όλων των διασπορών σε ε.σ. α αν $\mathcal{F}_{Levene} > \mathcal{F}_{k-1, n-k; \alpha}$.

Εφαρμογή με την R

Θυμίζουμε ότι η συνάρτηση που εκτελεί ελέγχους-t στην R είναι η $t.test()$.

Από το περιβάλλον R Commander κάνουμε αρχικά εισαγωγή τα δεδομένα μας. Για να εκτελέσουμε τον έλεγχο-t, *Statistics* → *Means* → *Independent samples t-test*: στο *Response Variable* επιλέγουμε την εξαρτημένη μεταβλητή (εδώ years), στο *Groups* επιλέγουμε την ανεξάρτητη μεταβλητή (εδώ gender), ενώ στο *Options* έχουμε τη δυνατότητα να επιλέξουμε αν θα εκτελέσουμε τον δίπλευρο έλεγχο ή κάποιον από τους μονόπλευρους, ή αν θα θεωρήσουμε ίσες διασπορές ή όχι (δηλ. να εκτελέσουμε το Welch's t-test το οποίο είναι προεπιλογή), και πατάμε OK. Τα αποτελέσματα για το ερευνητικό ερώτημα (1) εμφανίζονται στο *Output* του R Commander, όπως φαίνονται στο Σχήμα 5. Όμοια αποτελέσματα προκύπτουν για τα ερευνητικά ερωτήματα (2) και (3).

Παρατήρηση 2. Το γεγονός ότι η R εκτελεί το Welch's t-test ως προεπιλογή στηρίζεται στην λογική ότι δεν χρειάζεται να μπούμε στην διαδικασία να εξετάσουμε την προϋπόθεση της ισότητας των διασπορών τη στιγμή που αυτός ο έλεγχος είναι ικανός να «διορθώσει» ενδεχόμενη παραβίαση αυτής της προϋπόθεσης.

```

Two Sample t-test

data:  years by gender
t = 1.8558, df = 30, p-value = 0.07333
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.06907742  1.44407742
sample estimates:
 mean in group male mean in group female
(a)                6.0000                5.3125

Welch Two Sample t-test

data:  years by gender
t = 1.8558, df = 29.743, p-value = 0.07341
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.06935179  1.44435179
sample estimates:
 mean in group male mean in group female
(β)                6.0000                5.3125

```

Σχήμα 5: Αποτελέσματα R για το ερευνητικό ερώτημα (1) θεωρώντας (α) ίσες διασπορές, και (β) άνισες διασπορές (Welch's t-test).

3 Έλεγχος-t για κατά ζεύγη δείγματα (paired-samples t-test)

Έστω ότι μία ιατρική ομάδα θέλει να αξιολογήσει αν υπάρχει διαφορετική επίδραση δύο διαφορετικών αγωγών σε ασθενείς με υψηλή συστολική πίεση όσον αφορά το χρόνο αντίδρασής τους στην κάθε αγωγή. Προς αυτή την κατεύθυνση, σε ένα δείγμα ασθενών που πληρούν τις προϋποθέσεις για να συμμετέχουν στην έρευνα καταγράφεται ο χρόνος αντίδρασής τους στις δύο αγωγές. Το ερευνητικό ερώτημα για την ιατρική ομάδα σε αυτή την περίπτωση είναι το εξής.

ερευνητικό ερώτημα (1): «Ο χρόνος αντίδρασης των ασθενών στις δύο αγωγές είναι διαφορετικός».

Αυτό το ερευνητικό ερώτημα μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu_1 = \mu_2 \quad \text{κατά} \quad H_1 : \mu_1 \neq \mu_2,$$

όπου μ_1, μ_2 αναπαριστούν το μέσο χρόνο αντίδρασης των ασθενών του πληθυσμού στην αγωγή Α και Β, αντίστοιχα. Αν και πρόκειται για έλεγχο υποθέσεων μέσων τιμών όπως στην Ενότητα 2 είναι λάθος να χρησιμοποιήσουμε εκείνο τον έλεγχο-t. Η διαφορά οφείλεται στο γεγονός ότι οι παρατηρήσεις τώρα είναι εξαρτημένες ανά ζεύγη, αφού ουσιαστικά πρόκειται για το ίδιο δείγμα.

Αξίζει να σημειώσουμε ότι εκτός από το να πάρουμε μετρήσεις στο ίδιο δείγμα έχοντας εφαρμόσει μια αγωγή/δράση/χειραγώγηση Α και στη συνέχεια να πάρουμε μετρήσεις έχοντας εφαρμόσει μία αγωγή/δράση/χειραγώγηση Β με σκοπό να εντοπίσουμε ενδεχόμενες διαφορές, μία άλλη συνηθισμένη κατάσταση είναι να πάρουμε μετρήσεις στο ίδιο δείγμα «πριν» και «μετά» από κάποια αγωγή/δράση/χειραγώγηση. Επίσης, γίνεται κατανοητό πως στην ουσία ενδιαφερόμαστε να εξετάσουμε την επίδραση ενός παράγοντα με δύο επίπεδα, λ.χ. παρουσία της αγωγής Α και παρουσία της αγωγής Β, ή απουσία της αγωγής και παρουσία της αγωγής, που έχει το ρόλο της ανεξάρτητης μεταβλητής, πάνω σε μία ποσοτική μεταβλητή που έχει το ρόλο της εξαρτημένης μεταβλητής, χρησιμοποιώντας όμως το ίδιο δείγμα³.

³Ο πειραματικός σχεδιασμός που αντιστοιχεί στον έλεγχο-t για ανεξάρτητα δείγματα αναφέρεται ως *πειραματικός σχεδιασμός μεταξύ συμμετεχόντων* (δηλ. υπάρχουν δύο πειραματικές συνθήκες και σε κάθε συνθήκη χρησιμοποιούνται διαφορετικοί συμμετέχοντες), ενώ για τον έλεγχο-t για κατά ζεύγη δείγματα αναφέρεται ως *πειραματικός σχεδιασμός εντός συμμετεχόντων* (δηλ. υπάρχουν δύο πειραματικές συνθήκες και οι συμμετέχοντες είναι οι ίδιοι και στις δύο συνθήκες του πειράματος, δηλαδή έχουμε *επαναλαμβανόμενες μετρήσεις*).

$$t = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} \sim t_{n-1},$$

- να έχουμε ανεξάρτητα ζεύγη παρατηρήσεων (οπότε τα D_i ανεξάρτητες παρατηρήσεις),
- τα D_i να ακολουθούν κανονική κατανομή.

Π. Μπομποιάς

Ο έλεγχος-t κατά ζεύγη δειγμάτων έδειξε ότι δεν υπάρχει στατιστικά σημαντική διαφορά ανάμεσα στους μέσους χρόνους απόκρισης των ασθενών στην αγωγή Α ($M = 42, 43$, $SD = 7, 26$) και στην αγωγή Β ($M = 44, 93$, $SD = 8, 80$), $t(29) = -1, 822$, $p\text{-value} = 0, 079$.

Αν θέλουμε να εκτελέσουμε κάποιον από τους μονόπλευρους ελέγχους στηρίζομαστε στα αποτελέσματα του δίπλευρου. Έτσι, αν το ερευνητικό ερώτημα που είχαμε θέσει εξ αρχής ήταν

ερευνητικό ερώτημα (2): «Ο χρόνος αντίδρασης των ασθενών στην αγωγή Α είναι μεγαλύτερος από το χρόνο αντίδρασης στην αγωγή Β»,

το οποίο μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu_1 \leq \mu_2 \quad \text{κατά} \quad H_1 : \mu_1 > \mu_2,$$

το $p\text{-value}$ σε αυτή την περίπτωση είναι

$$p\text{-value} = 1 - \frac{p\text{-value δίπλευρου}}{2} = 1 - \frac{0, 079}{2} = 0, 9605,$$

και θα παρουσιάζαμε τα αποτελέσματα ως εξής. Ο έλεγχος-t κατά ζεύγη δειγμάτων έδειξε ότι ο μέσος χρόνος απόκρισης των ασθενών στην αγωγή Α ($M = 42, 43$, $SD = 7, 26$) δεν είναι στατιστικά σημαντικά μεγαλύτερος από μέσο χρόνο απόκρισης στην αγωγή Β ($M = 44, 93$, $SD = 8, 80$), $t(29) = -1, 822$, $p\text{-value} = 0, 9605$.

Από την άλλη πλευρά, αν το ερευνητικό ερώτημα που είχαμε θέσει εξ αρχής ήταν

ερευνητικό ερώτημα (3): «Ο χρόνος αντίδρασης των ασθενών στην αγωγή Α είναι μικρότερος από το χρόνο αντίδρασης στην αγωγή Β»,

το οποίο μετατρέπεται στον έλεγχο υποθέσεων

$$H_0 : \mu_1 \geq \mu_2 \quad \text{κατά} \quad H_1 : \mu_1 < \mu_2,$$

το $p\text{-value}$ σε αυτή την περίπτωση είναι

$$p\text{-value} = \frac{p\text{-value δίπλευρου}}{2} = \frac{0, 079}{2} = 0, 0395,$$

και θα παρουσιάζαμε τα αποτελέσματα ως εξής. Ο έλεγχος-t κατά ζεύγη δειγμάτων έδειξε ότι ο μέσος χρόνος απόκρισης των ασθενών στην αγωγή Α ($M = 42, 43$, $SD = 7, 26$) είναι στατιστικά σημαντικά μικρότερος από μέσο χρόνο απόκρισης στην αγωγή Β ($M = 44, 93$, $SD = 8, 80$), $t(29) = -1, 822$, $p\text{-value} = 0, 0395$.