Μη-Παραμετρική Στατιστική

Φροντιστήριο #4

Συνέχεια στο γενικό πρόβλημα δύο δειγμάτων

Ο έλεγχος Mann-Whitney (U-test) σε συνδυασμό με τον έλεγχο αθροισμάτων τάξεων μεγέθους του Wilcoxon (Mann-Whitney U test and Wilcoxon Rank sum test)

Οι δύο αυτοί έλεγχοι είναι ισοδύναμοι. Χρησιμοποιούνται για δύο ανεξάρτητα δείγματα για να ελέγξουμε αν αυτά προέρχονται από την ίδια συνεχή κατανομή χωρίς αυτή να προσδιορίζεται.

Έστω $X_1, X_2, \dots, X_m$ τ.δ. μεγέθους m από έναν πληθυσμό X με συνάρτηση κατανομής $F_X(x)$ και $Y_1, Y_2, \dots, Y_n$ ένα άλλο τυχαίο δείγμα μεγέθους n από έναν άλλο πληθυσμό με συνάρτηση κατανομής $F_Y(x)$.

Η μηδενική υπόθεση H_0, υποδηλώνει ότι οι δύο μεταβλητές X και Y έχουν την ίδια κατανομή ενώ η εναλλακτική υπόθεση υποδηλώνει ότι η Y τείνει να είναι "μεγαλύτερη" (ή "μικρότερη") από την X δηλ.

$$H_1: F_X(x) \geq F_Y(x) \ \forall x \text{ ή } F_X(x) > F_Y(x) \text{ για κάποιο } x$$

$$\text{ή } H_1: F_X(x) \leq F_Y(x) \ \forall x \text{ ή } F_X(x) < F_Y(x) \text{ για κάποιο } x$$

$$\text{ή } H_1: F_X(x) \neq F_Y(x)$$

Για το μοντέλο θέσης, ισοδύναμα μπορούμε να γράφουμε:

$$H_0: \theta = 0 \quad vs \quad H_1: \theta > 0$$
$$H_1: \theta < 0$$
$$H_1: \theta \neq 0$$

Για παράδειγμα η υπόθεση $H_1: F_X(x) \geq F_Y(x)$ μπορεί να γραφεί: $1 - P(X \leq x) \geq 1 - P(Y \leq x)$

$\Rightarrow P(X \leq x) \geq P(Y \leq x) \Rightarrow 1 - P(X > x) \geq 1 - P(Y > x)$

$\Rightarrow P(X > x) \leq P(Y > x) \; \forall x$ άρα μπορούμε να πούμε ισοδύναμα ότι η πιθανότητα $Y > x$ είναι μεγαλύτερη ή ίση με την πιθανότητα $X > x$ για κάθε τιμή του x. Σε μία τέτοια περίπτωση λέμε ότι η τ.μ. Y είναι "στοχαστικά μεγαλύτερη" από την τ.μ. X και γράφουμε:

$Y \overset{st}{>} X$.

Πρακτικά, ο ερευνητής πραγματοποιώντας έναν τέτοιο έλεγχο επιθυμεί να ανιχνεύσει διαφορές μεταξύ των δύο πληθυσμών με βάση τα τ.δ. από αυτούς τους πληθυσμούς. Εάν τα δείγματα αποτελούνται από δεδομένα σε διατεταγμένη κλίμακα η πιο ενδιαφέρουσα διαφορά για έναν ερευνητή θα ήταν μία διαφορά στην θέση των δύο πληθυσμών. Τείνουν οι τιμές του ενός πληθυσμού να είναι μεγαλύτερες από τις τιμές του άλλου πληθυσμού; Έχουν οι δύο πληθυσμοί ίσες διαμέσους; Έχουν οι δύο πληθυσμοί ίσες μέσες τιμές;

Το πιο συνηθισμένο παραμετρικό ανάλογο των δύο αυτών ελέγχων είναι ο έλεγχος t για δύο ανεξάρτητα δείγματα (independent samples t-test).

Πολλές στατιστικές μέθοδοι που εφαρμόζονται στο πρόβλημα των **δύο** δειγμάτων βασίζονται σε σ.σ. ελέγχου που σχετίζονται με το άθροισμα των βαθμών των παρατηρήσεων στο ενιαίο δείγμα που προκύπτει από την συνένωση των δύο δειγμάτων. Αυτές οι στατιστικές συναρτήσεις είναι γνωστές ως rank-order statistics και μπορούν να παρέχουν "πληροφορίες" για πιθανές διαφορές μεταξύ των πληθυσμών. Για παράδειγμα, αν ο πληθυσμός X έχει μεγαλύτερη μέση τιμή από τον πληθυσμό Y οι δειγματικές μετρήσεις θα αντανακλούν αυτήν την διαφορά όταν οι περισσότεροι βαθμοί στις μετρήσεις της X ξεπερνούν τους βαθμούς των μετρήσεων της Y.

Μία σημαντική κατηγορία σ.σ. ελέγχου αυτού του τύπου είναι οι συναρτήσεις γραμμικού βαθμού (linear rank statistics) που ορίζονται ως <u>γραμμική συνάρτηση</u> (linear function) των δείκτριων τυχαίων μεταβλητών Z_i του διανύσματος Z, όπου

$$Z = (Z_1, Z_2, \ldots, Z_N)^*.$$

* Θα εξηγήσουμε παρακάτω τον ορισμό του διανύσματος Z.

Αν έχουμε δύο τυχαία δείγματα, $X_1, X_2, \dots, X_m$ και $Y_1, Y_2, \dots, Y_n$ από πληθυσμούς με <u>συνεχείς</u> σ.κ. $F_X(x)$ και $F_Y(x)$, αντίστοιχα, τότε κάτω από την $H_0: F_X(x) = F_Y(x) = F(x)\ \forall x$, με F απροσδιόριστη, τότε έχουμε ένα σύνολο $m+n=N$ παρατηρήσεων από έναν κοινό (άγνωστο) πληθυσμό για το οποίον (αν δεν υπάρχουν ισοβαθμίες) μπορούν να τεθούν οι ακέραιοι βαθμοί $1, 2, \dots, N$.

Έστω $Z = (Z_1, Z_2, \dots, Z_N)$, όπου

$Z_i = 1$, αν η i-οστή τ.μ. στο ενιαίο διατεταγμένο δείγμα είναι μία X και

$Z_i = 0$, αν η i-οστή τ.μ. στο ενιαίο διατεταγμένο δείγμα είναι μία Y.

για $i = 1, 2, \dots, N$ με $N = n+m$. Συνεπώς, ο βαθμός της παρατήρησης για την οποία Z_i είναι η δείκτρια συνάρτηση είναι i και το διάνυσμα Z ορίζει τη σ.σ. διατεταγμένων τάξεων (βαθμών) των ενιαίων δειγμάτων και επιπλέον αναγνωρίζει το δείγμα από το οποίο προέρχεται η παρατήρηση.

Για παράδειγμα, δοθέντων των μετρήσεων $(X_1, X_2, X_3, X_4) = (2, 9, 3, 4)$ και $(Y_1, Y_2, Y_3) = (1, 6, 10)$ το ενιαίο διατεταγμένο δείγμα είναι:

$(1, 2, 3, 4, 6, 9, 10)$ ή $(Y_1, X_1, X_3, X_4, Y_2, X_2, Y_3)$

και το αντίστοιχο διάνυσμα Z είναι:

$(0, 1, 1, 1, 0, 1, 0)$. Αφού, για παράδειγμα $Z_6 = 1$, μία παρατήρηση X (συγκεκριμένα η X_2) έχει βαθμό 6 στο ενιαίο διατεταγμένο δείγμα.

Γενικά, λοιπόν μία σ.σ. γραμμικού βαθμού μπορεί να γραφεί στη μορφή:

$$T_N(Z) = \sum_{i=1}^{N} a_i Z_i$$

όπου τα a_i είναι κατάλληλες σταθερές, που καλούνται βάρη ή σκορς (weights or scores).

Το παραμετρικό t-test έχει αποδειχθεί ισχυρό test κάτω από τις υποθέσεις της κανονικότητας των πληθυσμών και των ίσων διασπορών. Τα δύο μη-παραμετρικά tests που παρουσιάζουμε εδώ, δεν απαιτούν την ισχύ της υπόθεσης της κανονικότητας.

Έλεγχος Wilcoxon Rank-sum

Η σ.σ. ελέγχου που χρησιμοποιείται είναι η:

$$W_N = \sum_{i=1}^{N} i Z_i$$

όπου οι Z_i είναι οι δείκτριες τ.μ. που έχουμε ορίσει παραπάνω. Αν δεν υπάρχουν ισοβαθμίες η ακριβής κατανομή και η διακύμανση της τ.μ. W_N κάτω από τη μηδενική υπόθεση μπορούν να υπολογιστούν. Έχει αποδειχθεί ότι:

$$E(W_N) = \frac{m(N+1)}{2}, \quad var(W_N) = \frac{mn(N+1)}{12}.$$

Η κατανομή της W_N είναι διακριτή, γνωστή ως Wilcoxon-Rank-Sum κατανομή και είναι συμμετρική.

Για μεγάλα δείγματα (πρακτικά για $n, m \geq 10$) η κανονική προσέγγιση για την τ.μ. W_N μπορεί να χρησιμοποιηθεί και να δώσει αρκετά "καλές" προσεγγίσεις για ενιαία δείγματα με μέγεθος $N \geq 12$. Αποδεικνύεται ότι η σ.σ.

$$W_N^* = \frac{W_N - E(W_N)}{\sqrt{var(W_N)}} \sim N(0,1)$$ κάτω από την H_0.

Αν συμβολίσουμε με w_N^* την παρατηρούμενη τιμή της W_N^*, τότε ο προσεγγιστικός υπολογισμός της p-τιμής (p-value) γίνεται σύμφωνα με τον ακόλουθο πίνακα:

$$H_0: F_X(x) = F_Y(x) \text{ ή } H_0: \theta = 0$$

X: τδ. μεγέθους m
Y: τδ μεγέθους n

Εναλλακτική	p-value
$H_1: \theta \neq 0$ ή $F_X(x) \neq F_Y(x)$	$2P(W_N^* \geq \lvert w_N^* \rvert) = 2(1 - \Phi(\lvert w_N^* \rvert))$
$H_1: \theta > 0$ ή $F_X(x) \geq F_Y(x)$	$P(W_N^* \geq w_N^*) = 1 - \Phi(w_N^*)$
$H_1: \theta < 0$ ή $F_X(x) \leq F_Y(x)$	$P(W_N^* \leq w_N^*) = \Phi(w_N^*)$, όπου

$\Phi(x)$ είναι η σ.κ. της τυπικής κανονικής κατανομής $N(0,1)$.

Αν υπάρχουν ισοβαθμίες (ties), τότε η διασπορά της τ.μ. W_N στην κανονική προσέγγιση μπορεί να διορθωθεί ώστε να λάβει υπόψη αυτές τις ισοβαθμίες. Σε περιπτώσεις ισοβαθμιών, χρησιμοποιείται η σ.σ.

$$W_N^* = \frac{W_N - E(W_N)}{\sqrt{var_c(W)}}, \text{ όπου}$$

$$var_c(W) = \frac{nm(N+1)}{12} - \frac{mn\sum_{j=1}^{r}(t_j^3 - t_j)}{12N(N-1)}$$

η οποία ασυμπτωτικά καθώς $n,m \to \infty$ ακολουθεί την κατανομή $N(0,1)$ κάτω από την H_0. Το r είναι το πλήθος των ομάδων με ισοβαθμίες και t_j είναι το μέγεθος της ομάδας j, $1 \le j \le r$.

<u>**Παρατήρηση:**</u> Έστω τα τ.δ. $X_1, X_2, \ldots, X_m$ και $Y_1, Y_2, \ldots, Y_n$. Έχουμε γράψει:

$$W_N = \sum_{i=1}^{N} i \cdot Z_i, \text{ όπου } Z_i = \begin{cases} 1, & \text{αν η } i\text{-οστή τ.μ. είναι } X \\ 0, & \text{αν η } i\text{-οστή τ.μ. είναι } Y \end{cases}$$

Αν συνενώσουμε τα δύο δείγματα:

$X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n$ και $R_i = \text{rank}(X_i)$

$=$ βαθμός(X_i) μεταξύ όλων των παρατηρήσεων

μπορούμε ισοδύναμα να γράψουμε:

$$W_N = \sum_{i=1}^{m} R_i .$$

Υπάρχει ένας ωραίος τρόπος να ξαναγράψουμε την σ.σ. W_N.

Έχουμε, εξ ορισμού,

$$R_i = \sum_{k=1}^{m} I(X_k \leq X_i) + \sum_{j=1}^{n} I(Y_j < X_i)$$

Άρα

$$W_N = \sum_{i=1}^{m} R_i = \frac{m(m+1)}{2} + \sum_{i=1}^{m} \sum_{j=1}^{n} I(Y_j < X_i)$$

Τότε η σ.σ.

$$U_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} I(Y_j < X_i) = W_N - \frac{m(m+1)}{2}$$

καλείται σ.σ. Mann-Whitney. Η σ.σ. Mann-Whitney απορρίπτει την Ho για υψηλές τιμές της σ.σ. $U_{m,n}$ και προφανώς προκύπτει ένα ισοδύναμο test με το Wilcoxon-Rank-sum test.

Μπορούμε εναλλακτικά να γράψουμε την $U_{m,n}$ ως εξής:

$$U_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} D_{ij} \quad \text{όπου} \quad D_{ij} = \begin{cases} 1, & \text{αν } Y_j < X_i \\ 0, & \text{αν } Y_j > X_i \end{cases}$$

τότε το $\sum_{j=1}^{n} D_{ij}$ είναι ο αριθμός των τιμών j για τις οποίες $Y_j < X_i$, δηλαδή η τιμή της $U_{m,n}$ είναι: το πλήθος των Y_j που είναι μικρότερα από X_1 συν το πλήθος των Y_j που είναι μικρότερα από X_2, ..., συν το πλήθος των Y_j που είναι μικρότερα από το X_m.

Στην ουσία, η σ.σ. Mann-Whitney $U_{m,n}$ είναι μία γραμμική συνάρτηση της σ.σ. Wilcoxon rank-sum W_N.

Για μικρά δείγματα, αν το δείγμα με τις λιγότερες παρατηρήσεις πάντοτε ορίζεται ως το δείγμα που προέρχεται από τον πληθυσμό X κατάλληλοι πίνακες υπάρχουν που για $m \leq n$ δίνουν τα κατάλληλα κρίσιμα σημεία για τη διενέργεια του ελέγχου. Συνεπώς, η ακριβής κατανομή της U μπορεί τότε να προσδιοριστεί. Όταν m και n είναι μεγάλα για αυτούς τους πίνακες, η ασυμπτωτική προσέγγιση για την κατανομή της U μπορεί να χρησιμοποιηθεί.

Αποδεικνύεται ότι, κάτω από την H_0:

$$E(U_{m,n}) = \frac{mn}{2}, \quad var(U_{m,n}) = \frac{mn(N+1)}{12}$$

Επομένως, η σ.σ. ελέγχου $U_{m,n}$ για μεγάλα δείγματα

είναι:

$$Z = \frac{U - \frac{mn}{2}}{\sqrt{\frac{mn(N+1)}{12}}} \sim N(0,1) \text{ (προσεγγιστικά)}$$

Μία διόρθωση συνέχειας του 0.5 μπορεί να χρησιμοποιηθεί επειδή η U λαμβάνει μόνο ακέραιες τιμές. Όπως και στον έλεγχο Wilcoxon Rank sum μία διόρθωση για τις ισοβαθμίες μπορεί να ενσωματωθεί στην σ.σ. ελέγχου, ως εξής:

$$var_c(U) = \frac{mn(N+1)}{12}\left[1 - \frac{\sum_{j=1}^{r}(t_j^3 - t_j)}{N(N^2-1)}\right]$$

Σχόλια:
- Τα δύο test όπως είδαμε είναι ισοδύναμα.
- Το Wilcoxon Rank sum είναι πιο εύκολο στη χρήση.
- Μόνο ανεξαρτησία δειγμάτων και συνεχείς κατανομές χρειάζεται να **υποτεθούν** για τους δύο πληθυσμούς.
- Τα test μπορούν να υλοποιηθούν και για μικρά δείγματα (ακριβείς κατανομές) και για μεγάλα δείγματα (με χρήση της κανονικής προσέγγισης).
- Διορθώσεις για τις ισοβαθμίες επίσης μπορούν να χρησιμοποιηθούν.

• Όταν οι πληθυσμοί υποτίθεται ότι διαφέρουν μόνο ως προς τη θέση (location) το Mann-Whitney test είναι άμεσα συγκρίσιμο με το Student t-test για έλεγχο μέσων τιμών.

• Για κάποιες μη-κανονικές κατανομές το Mann-Whitney test μοιάζει πιο αξιόπιστο. Πολλοί Στατιστικοί το θεωρούν (μαζί με το Wilcoxon Rank sum test) ως το καλύτερο μη-παραμετρικό test για έλεγχο της θέσης μεταξύ δύο πληθυσμών.

Στην R, ο έλεγχος Wilcoxon Rank-sum (ισοδύναμα ο Mann-Whitney-U) για δύο ανεξάρτητα δείγματα πραγματοποιείται με τη συνάρτηση wilcox.test.

Ας δούμε ένα παράδειγμα.

<u>Παράδειγμα</u> Οι βαθμοί στο μάθημα της Στατιστικής για ένα δείγμα $n=15$ φοιτητών από την Α.Σ.Ο.Ε.Ε δίνονται παρακάτω στη στήλη X. Οι βαθμοί στο μάθημα της Στατιστικής για ένα δείγμα $m=17$ φοιτητών σε ένα επαρχιακό Πανεπιστήμιο της Χώρας δίνονται στη στήλη Y. Να ελεγχθεί σε εσσ $\alpha=5\%$, αν οι βαθμοί των φοιτητών της επαρχίας τείνουν να είναι μεγαλύτεροι από αυτούς των φοιτητών της πρωτεύουσας.

<u>Λύση</u> Έστω $F_X(x)$ και $F_Y(x)$ οι σ.κ για τα δύο δείγματα

X: = βαθμοί φοιτητών Στατιστικής στην Αθήνα ($n=15$)

Y: = βαθμοί φοιτητών Στατιστικής στην επαρχία ($m=17$)

<u>Παρατήρηση</u>: Οι έλεγχοι:

A. $H_0: F_X(x) = F_Y(x) \quad \forall x$
$H_1: F_X(x) \neq F_Y(x)$ για κάποιο x

B. $H_0: F_X(x) \leq F_Y(x)$
$H_1: F_X(x) > F_Y(x)$ για κάποιο x

Γ. $H_0: F_X(x) \geq F_Y(x) \quad \forall x$
$H_1: F_X(x) < F_Y(x)$ για κάποιο x

μπορούν να γραφούν ισοδύναμα:

A. $H_0: P(X > x) = P(Y > x) = 1/2$
$H_1: P(X > x) \neq P(Y > x)$ για ένα τουλ. x
ή $P(X < Y) \neq 1/2$

Ομοίως, ισοδύναμα:

B. $H_0: P(X < Y) \leq 1/2$
$H_1: P(X < Y) > 1/2$

και

Γ. $H_0: P(X < Y) \geq 1/2$
$H_1: P(X < Y) < 1/2$

Η σ.σ. ελέγχου $T = \sum_{i=1}^{n} R(X_i)$ δηλ. άθροισμα των βαθμών του δείγματος X στο ενιαίο δείγμα παρατηρήσεων (συνήθως) χρησιμοποιείται όταν δεν υπάρχουν ισοβαθμίες ή όταν έχουμε λίγες ισοβαθμίες. Αν υπάρχουν <u>πολλές</u> ισοβαθμίες χρησιμοποιούμε την τυποποιημένη τιμή της T, $T_1 = \dfrac{T - E(T)}{\sqrt{var(T)}}$ με $E(T) = \dfrac{n(N+1)}{2}$, $N = n+m$

$$var(T) = \frac{nm}{N(N-1)} \sum_{i=1}^{N} R_i^2 - \frac{nm(N+1)^2}{4(N-1)} \quad \text{όπου}$$

το $\sum_{i=1}^{N} R_i^2$ αναφέρεται στο άθροισμα των τετραγώνων των βαθμών στο ενιαίο δείγμα.

$T_1 \sim N(0,1)$. Τα ποσοστιαία σημεία της T δίνονται από πίνακες (π.χ. βλέπε Πίνακας 9, βιβλίο Ξεκαλάκη, παράρτημα), συμβολίζονται με w_p, όπου $p = 0.001, 0.005, 0.01, 0.025, 0.05, 0.1$). Για τα άνω-ποσοστιαία σημεία ισχύει: $w_{1-p} = n(N+1) - w_p$, $N = n+m$.

Οι περιοχές απόρριψης για κάθε έλεγχο καθορίζονται ως εξής:

A. $\cancel{H_0}$ σε εσσ α αν $T > w_{1-\alpha/2}$ ή $T < w_{\alpha/2}$

ή $T_1 > z_{1-\alpha/2}$ ή $T_1 < z_{\alpha/2}$

B: $\cancel{H_0}$ σε εσσ α αν $T < w_\alpha$ ή $T_1 < z_\alpha$

δηλ. μικρές τιμές της T ένδειξη για $\cancel{H_0}$.

Γ. $\cancel{H_0}$ σε εσσ α αν $T > w_{1-\alpha}$ ή $T_1 < z_{1-\alpha}$

Μεγάλες τιμές της T ένδειξη για $\cancel{H_0}$. Συχνά ο έλεγχος Wilcoxon-Mann-Whitney (WMW-test) αναφέρεται στις μέσες τιμές των δύο πληθυσμών (αν αυτές υπάρχουν) και στις διαμέσους τους. Σε τέτοιες περιπτώσεις ισοδύναμα μπορούμε να γράψουμε:

A. $H_0: E(X) = E(Y)$
$H_1: E(X) \neq E(Y)$

B. $H_0: E(X) \geq E(Y)$
$H_1: E(X) < E(Y)$

Γ. $H_0: E(X) \leq E(Y)$
$H_1: E(X) > E(Y)$

Συνεπώς, τότε το test χρησιμοποιείται για την ύπαρξη ή όχι στατιστικά σημαντικών διαφορών στη "Θέση" (μέση τιμή ή διάμεσο) δύο πληθυσμών.

X: 7 8 6 9 4 2 6 5 8 9 10 4
5 7 4 $(n=15)$

Y: 6 5 3 7 9 8 6 9 10 5 2 3 7
9 9 7 4 $(m=17)$

Ενιαίο δείγμα: 7 8 6 9 4 2 6 5 8 9 10 4 5 7 4
6 5 3 7 9 8 6 9 10 5 2 3 7 8 9
7 4 $(N=n+m=32)$

Βαθμοί στο
ενιαίο δείγμα: 19, 23.5, 14.5, 28, 6.5, 1.5, 14.5
10.5, 23.5, 28, 31.5, 6.5, 10.5, 19,
6.5, 14.5, 10.5, 3.5, 19, 28, 23.5
14.5, 28, 31.5, 10.5, 1.5, 3.5, 19, 23.5
28, 19, 6.5

Ελέγχουμε σε εσσ $\alpha=5\%$,

$H_0: P(X<Y)\leq 1/2$ ή $H_0: E(X)\geq E(Y)$

$H_1: P(X<Y)>1/2$ $\quad$ $H_1: E(X)<E(Y)$

$$T=\sum_{i=1}^{n} R(X_i)=243.5 \text{ και } \sum_{i=1}^{N} R_i^2=11398.5$$

Χρησιμοποιώ την κανονική προσέγγιση (σ.σ. T_1):

$$T_1=\frac{T-\frac{n(N+1)}{2}}{\sqrt{\frac{nm}{N(N-1)}\sum_{i=1}^{N}R_i^2-\frac{nm(N+1)^2}{4(N-1)}}}=\frac{243.5-\frac{15\cdot 33}{2}}{\sqrt{\frac{15\cdot 17}{32\cdot 31}11398.5-\frac{15\cdot 17\cdot 33^2}{4\cdot 31}}}$$

$$=\frac{243.5-247.5}{\sqrt{690.58}}=-0.15.$$

Απορρίπτω H_0 αν $T_1 < z_\alpha = -z_{1-\alpha} = -z_{0.95} = -1.645$

αφού $T_1 = -0.15$ η H_0 <u>δεν</u> απορρίπτεται.

<u>Υπολογισμός p-value</u>

$$\hat{\alpha} = P(T_1 \leq -0.15 | H_0) = P(Z \leq -0.15) \simeq 0.44$$

άρα σε $\alpha = 5\%$ δεν απορρίπτεται η H_0, αφού $\alpha = 5\%$ δεν είναι μεγαλύτερο $\hat{\alpha} = 0.44$.

Λύση με την R, επισυνάπτεται στην επόμενη σελίδα. Με διάφορα γραφήματα βλέπουμε ότι έχουμε (εμπειρικές) "ισχυρές" ενδείξεις αποκλίσεων από την υπόθεση της κανονικότητας για τα δύο δείγματα.

Η τιμή W που δίνει η R είναι η τιμή του στατιστικού $U = \sum_{i=1}^{n} R(X_i) - \frac{n(n+1)}{2}$

$$= 243.5 - \frac{15 \cdot 16}{2} = 243.5 - 120 = 123.5 = W.$$

```
library(ggplot2)
x <- c(7,8,6,9,4,2,6,5,8,9,10,4,5,7,4)
y <- c(6,5,3,7,9,8,6,9,10,5,2,3,7,8,9,7,4)
#hist(x)
#boxplot(x)
qqnorm(x)
qqline(x)
#hist(y)
#boxplot(y)
#qqnorm(y)
#qqline(y)
wilcox.test(x,y,alternative="l", EXACT=false)
```

```
 Wilcoxon rank sum test with continuity correction

data:  x and y
W = 123.5, p-value = 0.447
alternative hypothesis: true location shift is less than 0

Warning message:
In wilcox.test.default(x, y, alternative = "l", EXACT = false) :
  cannot compute exact p-value with ties
```