Έλεγχοι τάξης

Μη Παραμετρική Στατιστική

Παναγιώτης Παπασταμούλης
Αναπληρωτής Καθηγητής
Τμήμα Στατιστικής ΟΠΑ

papastamoulis@aueb.gr

2026

Περιεχόμενα : Έλεγχοι τάξης

1. Εισαγωγή
2. Έλεγχος προσημασμένων τάξεων μεγέθους Wilcoxon (Signed ranks test)
   - Ακριβές τεστ (για μικρά δείγματα)
   - Ασυμπτωτικό τέστ
3. Έλεγχος ισότητας δύο πληθυσμιακών διαμέσων
   - Ανεξάρτητα δείγματα : έλεγχος Wilcoxon-Mann-Whitney rank sum
   - Εξαρτημένα δείγματα
4. Έλεγχος ισότητας περισσότερων των δύο πληθυσμιακών διαμέσων
   - Ανεξάρτητα δείγματα : έλεγχος Kruskal-Wallis
   - Εξαρτημένα δείγματα : έλεγχος Friedman

Σκοπός

Μας ενδιαφέρει το πρόβλημα ελέγχου ισότητας

- της πληθυσμιακής διαμέσου με δοθείσα τιμή
- δύο ή περισσότερων πληθυσμιακών διαμέσων
    - ανεξάρτητα δείγματα
    - εξαρτημένα δείγματα
- δύο ή περισσότερων πληθυσμιακών διακυμάνσεων.

με χρήση μη παραμετρικών τεχνικών.

Έστω ένα σύνολο δεδομένων

$$X_1, X_2, \ldots, X_n.$$

Ορισμός 1

Τάξη (rank) μιας παρατήρησης ενός συνόλου δεδομένων είναι ο αριθμός που δίνει τη θέση που η παρατήρηση έχει στο διατεταγμένο κατά αύξουσα τάξη μεγέθους δείγμα. Δηλαδή,

$$R_i = R(X_i) = \sum_{j=1}^{n} \mathrm{I}(X_j \leqslant X_i), \quad i = 1, \ldots, n.$$

Αν δύο ή περισσότερες τιμές ταυτίζονται, οπότε και λέμε ότι έχουμε **δεσμούς ή ισοβαθμίες**, τότε ως τάξη καθεμίας θεωρείται ο μέσος όρος των τάξεων που αυτές θα είχαν αν ήταν διαφορετικές (midranks).

Παράδειγμα

Παράδειγμα 2

Υπολογίστε τις τάξεις για κάθε παρατήρηση των ακόλουθων συνόλων δεδομένων:

(α) 126 142 156 228 245 246 370 419 433 454 478 503

(β) 126 142 370 228 245 245 370 370 433 454 433 142

(γ) «καθόλου» «λίγο» «πολύ» «λίγο» «μέτρια» «πολύ» «λίγο» «καθόλου» «πολύ»

Λύση (α)

- Σε αυτό το σύνολο δεδομένων οι παρατηρήσεις είναι ήδη διατεταγμένες κατά αύξουσα τάξη μεγέθους.
- Παρατηρούμε ότι δεν υπάρχουν δεσμοί (καθώς δεν υπάρχουν παρατηρήσεις με την ίδια τιμή).
- Έτσι προκύπτει ότι οι τάξεις τους είναι:

$$1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,$$

αντίστοιχα.

Λύση (β)

- Σε αυτό το σύνολο δεδομένων οι παρατηρήσεις δεν είναι διατεταγμένες κατά αύξουσα τάξη μεγέθους.
- Τις διατάσσουμε κατά αύξουσα τάξη μεγέθους. Είναι τότε:

$$126,\ 142,\ 142,\ 228,\ 245,\ 245,\ 370,\ 370,\ 370,\ 433,\ 433,\ 454.$$

- Παρατηρούμε ότι υπάρχουν δεσμοί:

$$R(142) = (2+3)/2 = 2.5$$
$$R(245) = (5+6)/2 = 5.5$$
$$R(370) = (7+8+9)/3 = 8$$
$$R(433) = (10+11)/2 = 10.5$$

- Επομένως, οι τάξεις των αρχικών παρατηρήσεων είναι αυτές που δίνονται στον πίνακα που ακολουθεί:

i	1	2	3	4	5	6	7	8	9	10	11	12
X_i	126	142	370	228	245	245	370	370	433	454	433	142
$R(X_i)$	1	2.5	8	4	5.5	5.5	8	8	10.5	12	10.5	2.5

Λύση (γ)

- Σε αυτό το σύνολο δεδομένων οι παρατηρήσεις δεν είναι διατεταγμένες κατά αύξουσα τάξη μεγέθους.
- Τις διατάσσουμε κατά αύξουσα τάξη μεγέθους.
  «καθόλου» «καθόλου» «λίγο» «λίγο» «λίγο» «μέτρια» «πολύ» «πολύ» «πολύ».
- Οι τάξεις αυτών είναι: 1.5, 1.5, 4, 4, 4, 6, 8, 8, 8, αντίστοιχα.
- Επομένως

i	1	2	3	4	5	6	7	8	9
X_i	«καθόλου»	«λίγο»	«πολύ»	«λίγο»	«μέτρια»	«πολύ»	«λίγο»	«καθόλου»	«πολύ»
$R(X_i)$	1.5	4	8	4	6	8	4	1.5	8

Ιδιότητες τάξεων

Πρόταση 1.1

Αν $R_1, ..., R_n$ *είναι οι τάξεις ενός τυχαίου δείγματος* $X_1, \ldots, X_n$ *από μία συνεχή αθροιστική συνάρτηση κατανομής τότε, υποθέτοντας ότι δεν υπάρχουν δεσμοί, ισχύει ότι:*

α) $\mathrm{P}\left(R_i = j\right) = \frac{1}{n}, j = 1, ..., n.$

β) $\mathrm{E}\left(R_i\right) = \frac{n+1}{2}$ *και* $\mathrm{Var}\left(R_i\right) = \frac{n^2-1}{12}.$

γ) $\mathrm{Cov}\left(R_i, R_j\right) = -\frac{n+1}{12}$, *για* $i \neq j, i, j = 1, ..., n.$

Απόδειξη: στον Πίνακα

Έλεγχος προσημασμένων τάξεων μεγέθους Wilcoxon (Wilcoxon signed ranks test)

Τι είναι το Wilcoxon Signed-Rank Test·

Το Wilcoxon signed-rank test είναι ένας μη παραμετρικός έλεγχος τάξεων για στατιστικό έλεγχο υποθέσεων. Χρησιμοποιείται είτε:

- για να ελέγξει τη θέση (location) ενός πληθυσμού με βάση ένα δείγμα, είτε
- για να συγκρίνει τη θέση δύο πληθυσμών χρησιμοποιώντας δύο συζευγμένα δείγματα.

Σχέση με το t-test

- Η εκδοχή με ένα δείγμα έχει παρόμοιο σκοπό με το one-sample Student's t-test.
- Για δύο συζευγμένα δείγματα, λειτουργεί όπως το paired Student's t-test (γνωστό και ως t-test for matched pairs ή t-test for dependent samples).

Πότε χρησιμοποιείται;

Το Wilcoxon test είναι καλή εναλλακτική του t-test όταν:

- δεν μπορεί να θεωρηθεί ότι οι διαφορές ακολουθούν κανονική κατανομή (normal distribution).

Αντί αυτού, υποθέτει ότι:

- η κατανομή των διαφορών είναι συμμετρική γύρω από μία κεντρική τιμή.

Στόχος του ελέγχου

Ο στόχος είναι να ελεγχθεί αν:

- η κεντρική τιμή των διαφορών διαφέρει σημαντικά από το μηδέν.

Σύγκριση με το Sign Test

- Το Wilcoxon test είναι πιο ισχυρό από το sign test,
- επειδή λαμβάνει υπόψη και το μέγεθος των διαφορών.

Ωστόσο:

- απαιτεί την υπόθεση συμμετρίας της κατανομής.

Έλεγχοι ισότητας της πληθυσμιακής διαμέσου με δοθείσα τιμή

- Έστω $X_1, \ldots, X_n$, ένα τυχαίο δείγμα από έναν πληθυσμό με συνεχή αθροιστική συνάρτηση κατανομής $F(\cdot)$, με διάμεσο m.
- Επιπλέον, υποθέτουμε ότι η F είναι συμμετρική γύρω από το σημείο m.
- Ενδιαφερόμαστε να ελέγξουμε μία εκ των τριών ακόλουθων υποθέσεων:
  (Α) $H_0 : m = m_0$, έναντι της $H_1 : m \neq m_0$.
  (Β) $H_0 : m = m_0$, έναντι της $H_1 : m > m_0$.
  (Γ) $H_0 : m = m_0$, έναντι της $H_1 : m < m_0$.
- Διαθέσιμα μη παραμετρικά τεστ
  - Προσημικό τεστ (διωνυμικός έλεγχος)
  - Wilcoxon Signed Rank test (1945)

Έλεγχος προσημασμένων τάξεων (Signed Ranks Test)

- Παρουσιάστηκε από τον Frank Wilcoxon (1945).
- Βασίζεται στις τάξεις των διαφορών

$$D_i = X_i - m_0, \quad i = 1, \ldots, n,$$

  με εξαίρεση τις μηδενικές διαφορές.
- Δημιουργούμε τις διαφορές $D_i = X_i - m_0$ και αφαιρούμε τις μηδενικές.
- Υποθέτουμε συμμετρία γύρω από το 0: θετικές και αρνητικές διαφορές είναι ισοπίθανες.
- Για κάθε $c > 0$:

$$F_D(-c) = P(D \leq -c) = 1 - F_D(c) = P(D \geq c).$$

Τάξεις των διαφορών

- Απόλυτες τιμές: $|D_1|, \ldots, |D_n|$ διατάσσονται κατά αύξουσα τάξη.
- Ορίζουμε τις τάξεις $R(|D_i|)$, $i = 1, \ldots, n$.
- Στατιστικές T^+ και T^-:

$$T^+ = \sum_{i=1}^{n} I(D_i > 0) R(|D_i|), \quad T^- = \sum_{i=1}^{n} I(D_i < 0) R(|D_i|).$$

- $T^+ + T^- = \frac{n(n+1)}{2}$.

Προσημασμένες τάξεις

- Ορίζουμε:

$$R_i^* = \begin{cases} +R(|D_i|), & D_i > 0 \\ -R(|D_i|), & D_i < 0 \end{cases}, \quad i = 1, \ldots, n$$

- Άθροισμα προσημασμένων τάξεων:

$$S = \sum_{i=1}^{n} R_i^* = T^+ - T^- = 2T^+ - \frac{n(n+1)}{2}$$

- Η R_i^* ονομάζεται **προσημασμένος βαθμός** ή **προσημασμένη τάξη μεγέθους**.

Προσημασμένες τάξεις

- Ορίζουμε:

$$R_i^* = \begin{cases} +R(|D_i|), & D_i > 0 \\ -R(|D_i|), & D_i < 0 \end{cases}, \quad i = 1, \ldots, n$$

- Άθροισμα προσημασμένων τάξεων:

$$S = \sum_{i=1}^{n} R_i^* = T^+ - T^- = 2T^+ - \frac{n(n+1)}{2}$$

- Η R_i^* ονομάζεται **προσημασμένος βαθμός** ή **προσημασμένη τάξη μεγέθους**.
- Για τη διεξαγωγή των ελέγχων η χρήση των T^+, T^-, S είναι ισοδύναμη.

Προσημασμένες τάξεις

- Ορίζουμε:

$$R_i^* = \begin{cases} +R(|D_i|), & D_i > 0 \\ -R(|D_i|), & D_i < 0 \end{cases}, \quad i = 1, \ldots, n$$

- Άθροισμα προσημασμένων τάξεων:

$$S = \sum_{i=1}^{n} R_i^* = T^+ - T^- = 2T^+ - \frac{n(n+1)}{2}$$

- Η R_i^* ονομάζεται **προσημασμένος βαθμός** ή **προσημασμένη τάξη μεγέθους**.
- Για τη διεξαγωγή των ελέγχων η χρήση των T^+, T^-, S είναι ισοδύναμη.
- Ας συμφωνήσουμε ότι χρησιμοποιούμε την T^+.

Κριτήρια απόρριψης

- Υπό $H_0 : m = m_0$, αναμένεται $T^+ \approx T^-$.
- Κρίσιμες περιοχές για $H_1 : m > m_0$:

$$T^+ \text{ μεγάλες}, \quad T^- \text{ μικρές}, \quad S \text{ μεγάλες}.$$

- Κρίσιμες περιοχές για $H_1 : m < m_0$:

$$T^+ \text{ μικρές}, \quad T^- \text{ μεγάλες}, \quad S \text{ μικρές}.$$

- Διπλής κατεύθυνσης: απορρίπτεται για ακραίες τιμές T^+, T^- ή S.

Θεώρημα 3

Έστω $X_1, \ldots, X_n$ *ένα τυχαίο δείγμα από έναν πληθυσμό με συνεχή αθροιστική συνάρτηση κατανομής* $F(\cdot)$. *Ενδιαφερόμαστε να ελέγξουμε τη μηδενική υπόθεση ότι η διάμεσος* m *της άγνωστης κατανομής είναι ίση με* m_0, *όπου* m_0 *είναι ένας γνωστός αριθμός. Αν* $D_i = X_i - m_0$, $i = 1, ..., n$, *είναι οι μη μηδενικές διαφορές, τότε, υπό τη μηδενική υπόθεση, η κατανομή της στατιστικής συνάρτησης* T^+ *που παριστάνει το άθροισμα των τάξεων που αντιστοιχούν στις θετικές διαφορές δίνεται από τη σχέση:*

$$P\left(T^+ = k\right) = \frac{C_n(k)}{2^n}, \text{ για } k = 0, ..., \frac{n(n+1)}{2},$$

όπου $C_n(k)$ *ο αριθμός των υποσυνόλων του συνόλου* $\{1, 2, ..., n\}$ *των οποίων τα στοιχεία αθροίζουν στο* k.

Απόδειξη: Randles and Wolfe (1979).

Παρατηρήσεις

- Η ακριβής κατανομή της στατιστικής T^+ υπό τη μηδενική υπόθεση εξαρτάται από το μέγεθος του δείγματος, αλλά όχι από την κατανομή των X_i.
- Επομένως, η διαδικασία ελέγχου που βασίζεται στη στατιστική T^+ είναι **απαλλαγμένη παραμέτρων** (distribution free).
- Δεν μπορεί να δοθεί σε κλειστή μορφή γενικά, καθώς δεν υπάρχει κλειστή μορφή για το $C_n(k)$, μόνο αναδρομική σχέση

$$C_n(k) = C_{n-1}(k) + C_{n-1}(k-n).$$

Κριτήρια απόρριψης για μικρά δείγματα

- Χρησιμοποιώντας το Θεώρημα 3 για μικρές τιμές του μεγέθους δείγματος n, σε επίπεδο σημαντικότητας a, προκύπτει ότι:
  - **(A)** Για το πρόβλημα $H_0 : m = m_0$, $H_1 : m \neq m_0$, απορρίπτεται η μηδενική υπόθεση αν
  $$T^+ \leq T_{n,1-a/2} \quad \text{ή} \quad T^+ > T_{n,a/2}.$$
  - **(B)** Για το πρόβλημα $H_0 : m = m_0$, $H_1 : m > m_0$, απορρίπτεται η μηδενική υπόθεση αν
  $$T^+ > T_{n,a}.$$
  - **(Γ)** Για το πρόβλημα $H_0 : m = m_0$, $H_1 : m < m_0$, απορρίπτεται η μηδενική υπόθεση αν
  $$T^+ \leq T_{n,1-a}.$$

  όπου $T_{n,p}$ είναι το p-ποσοστιαίο σημείο της κατανομής της T^+.
- Τα σημεία $T_{n,p}$ για συνήθεις τιμές του επιπέδου σημαντικότητας δίνονται στον Πίνακα 1.

Πρόταση 2.1

Έστω $X_1, \ldots, X_n$ είναι ένα τυχαίο δείγμα από έναν πληθυσμό με συνεχή αθροιστική συνάρτηση κατανομής $F(\cdot)$. Ενδιαφερόμαστε να ελέγξουμε την υπόθεση ότι η διάμεσος m της άγνωστης κατανομής είναι ίση με m_0, όπου m_0 είναι ένας γνωστός αριθμός. Αν $D_i = X_i - m_0$, $i = 1, ..., n$, είναι οι μη μηδενικές διαφορές, τότε, υπό τη μηδενική υπόθεση, η κατανομή της στατιστικής συνάρτησης T^+ που παριστάνει το άθροισμα των τάξεων που αντιστοιχούν στις θετικές διαφορές, είναι συμμετρική με κέντρο συμμετρίας την τιμή $\frac{n(n+1)}{4}$.

Πρόταση 2.2

Έστω $T_{n,p}$ είναι το p-ποσοστιαίο σημείο της κατανομής της T^+, τότε $T_{n,1-p} = \frac{n(n+1)}{2} - T_{n,p}$.

Πίνακας: Ποσοστιαία σημεία για τον προσημικό έλεγχο τάξης Wilcoxon (μέσω της T^+).

n	$T_{n,0.995}$	$T_{n,0.99}$	$T_{n,0.975}$	$T_{n,0.95}$	$T_{n,0.90}$	$T_{n,0.80}$	$T_{n,0.70}$	$T_{n,0.60}$	$T_{n,0.50}$	$\frac{n(n+1)}{2}$
4	0	0	0	0	1	3	3	4	5	10
5	0	0	0	1	3	4	5	6	7.5	15
6	0	0	1	3	4	6	8	9	10.5	21
7	0	1	3	4	6	9	11	12	14	28
8	1	2	4	6	9	12	14	16	18	36
9	2	4	6	9	11	15	18	20	22.5	45
10	4	6	9	11	15	19	22	25	27.5	55
11	6	8	11	14	18	23	27	30	33	66
12	8	10	14	18	22	28	32	36	39	78
13	10	13	18	22	27	33	38	42	45.5	91
14	13	16	22	26	32	39	44	48	52.5	105
15	16	20	26	31	37	45	51	55	60	120
16	20	24	30	36	43	51	58	63	68	136
17	24	28	35	42	49	58	65	71	76.5	153
18	28	33	41	48	56	66	73	80	85.5	171
19	33	38	47	54	63	74	82	89	95	190
20	38	44	53	61	70	83	91	98	105	210
21	44	50	59	68	78	91	100	108	115.5	231
22	49	56	67	76	87	100	110	119	126.5	253
23	55	63	74	84	95	110	120	130	138	276
24	62	70	82	92	105	120	131	141	150	300
25	69	77	90	101	114	131	143	153	162.55	325
26	76	85	99	111	125	142	155	165	175.5	351
27	84	94	108	120	135	154	167	178	189	378
28	92	102	117	131	146	166	180	192	203	406
29	101	111	127	141	158	178	193	206	217.5	435
30	110	121	138	152	170	191	207	220	232.5	465

Πηγή: Μπατσίδης, Παπασταμούλης, Πετρόπουλος, Ρακιτζής (2022): Πίνακας Π.17.

Παράδειγμα 4

Στον πίνακα που ακολουθεί καταγράφεται ο αριθμός των σελίδων 12 τυχαία επιλεγμένων βιβλίων από μία βιβλιοθήκη ενός μαθηματικού τμήματος:

$$126\ 142\ 156\ 228\ 245\ 246\ 370\ 419\ 433\ 454\ 478\ 503.$$

Χρησιμοποιώντας το τεστ του Wilcoxon και αφού κάνετε τις κατάλληλες υποθέσεις, να ελέγξετε, σε επίπεδο σημαντικότητας 5%, την υπόθεση ότι η πληθυσμιακή διάμεσος του αριθμού των σελίδων των βιβλίων της βιβλιοθήκης είναι ίση με 220 σελίδες.

Λύση

- Δημιουργούμε τις διαφορές $D_i = X_i - m_0$, $i = 1, ..., 12$, οι οποίες είναι οι:

  -94, -78, -64, 8, 25, 26, 150, 199, 213, 234, 258, 283.

- Οι απόλυτες τιμές των διαφορών $|D_1|, \ldots, |D_{12}|$, διατάσσονται κατά αύξουσα τάξη μεγέθους και υπολογίζονται οι τάξεις τους:

$\lvert D_i \rvert$	8	25	26	64	78	94	150	199	213	234	258	383
$R(\lvert D_i \rvert)$	1	2	3	4	5	6	7	8	9	10	11	12

- $T^- = 4 + 5 + 6 = 15$
- $T^+ = \frac{n(n+1)}{2} - T^- = \frac{12 \cdot 13}{2} - 15 = 63$
- Κρίσιμη περιοχή, σε επίπεδο σημαντικότητας 5%: $T^+ \leqslant T_{12,1-0.05/2}$ ή $T^+ > T_{12,0.05/2}$
- Από Πίνακα 1: $T_{12,1-0.05/2} = 14$ και

  $T_{12,0.05/2} = \frac{12(12+1)}{2} - T_{12,0.975} = 78 - 14 = 64.$
- Καθώς $T^+ = 63 \nleq 14$ και $T^+ = 63 \not> 64$, συμπεραίνουμε ότι δεν απορρίπτεται η μηδενική υπόθεση $H_0 : m = 220$ έναντι της $H_1 : m \neq 220$ σε επίπεδο σημαντικότητας 5%.

Ασυμπτωτική κατανομή

Πρόταση 2.3

Έστω $X_1, \ldots, X_n$ *είναι ένα τυχαίο δείγμα από έναν πληθυσμό με συνεχή αθροιστική συνάρτηση κατανομής* $F(\cdot)$. *Ενδιαφερόμαστε να ελέγξουμε την υπόθεση ότι η διάμεσος* m *της άγνωστης κατανομής είναι ίση με* m_0, *όπου* m_0 *είναι ένας γνωστός αριθμός. Υπό τη μηδενική υπόθεση* $H_0 : m = m_0$ *και την πρόσθετη υπόθεση ότι δεν υπάρχουν δεσμοί, για μεγάλο μέγεθος δείγματος, ισχύει ότι:*

$$W = \frac{T^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \xrightarrow{d} \mathcal{N}\left(0,1\right),$$

όπου $T^+ = \sum_{i=1}^{n} I\left(D_i > 0\right) R\left(|D_i|\right)$, *με* $D_i = X_i - m_0$, $i = 1, ..., n$, *και* $R\left(|D_i|\right)$, $i = 1, ..., n$, *είναι οι τάξεις των απόλυτων τιμών των θετικών διαφορών.*

Απόδειξη: στον πίνακα

Χρησιμοποιώντας την παραπάνω πρόταση, έχουμε τις ακόλουθες κρίσιμες περιοχές για τους αντίστοιχους ελέγχους:

(A) $|W| \geq z_{a/2}$ για τον έλεγχο της $H_0 : m = m_0$ έναντι της $H_1 : m \neq m_0$.

(B) $W \geq z_a$ για τον έλεγχο της $H_0 : m = m_0$ έναντι της $H_1 : m > m_0$.

(Γ) $W \leq -z_a$ για τον έλεγχο της $H_0 : m = m_0$ έναντι της $H_1 : m < m_0$.

Σε όσα προαναφέρθηκαν είχε υποτεθεί η μη ύπαρξη δεσμών μεταξύ των απόλυτων διαφορών.

Παράδειγμα 5

Θα απαντήσουμε στο πρόβλημα του Παραδείγματος 4 με χρήση ασυμπτωτικής προσέγγισης.

- Είναι $T^+ = 63$, $n = 12$
- Υπολογισμός W:

$$W = \frac{T^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \approx 1.8827$$

- Επειδή $|W| = 1.8827 > z_{\alpha/2} = z_{0.025} = 1.96$ απορρίπτεται η H_0 έναντι της δίπλευρης εναλλακτικής σε επίπεδο σημαντικότητας $\alpha = 5\%$
- Ασυμπτωτικό p-value:

$$\text{p-value} = 2\mathrm{P}(Z > |w|) \approx 0.0597.$$

Η ασυμπτωτική κατανομή όταν υπάρχουν δεσμοί

Πρόταση 2.4

Έστω $X_1, \ldots, X_n$ ένα τυχαίο δείγμα από έναν πληθυσμό με αθροιστική συνάρτηση κατανομής $F(\cdot)$. Ενδιαφερόμαστε να ελέγξουμε την υπόθεση ότι η διάμεσος m της άγνωστης κατανομής είναι ίση με m_0, όπου m_0 είναι ένας γνωστός αριθμός. Έστω, επίσης, $D_i = X_i - m_0$, $i = 1, ..., n$, οι μη μηδενικές διαφορές και $d_1, ..., d_c$, ο αριθμός των παρατηρήσεων σε καθεμία από τις c διαφορετικές απόλυτες διαφορές (σε αύξουσα τάξη μεγέθους), με $d_i \geq 1$, και $\sum_{i=1}^{c} d_i = n$. Τότε, υπό τη μηδενική υπόθεση, η προσεγγιστική κατανομή της στατιστικής συνάρτησης T^+ που παριστάνει το άθροισμα των τάξεων που αντιστοιχούν στις θετικές διαφορές είναι

$$W = \frac{T^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \sum_{i=1}^{c} \frac{d_i\left(d_i^2-1\right)}{48}}} \xrightarrow{d} \mathcal{N}(0,1).$$

Wilcoxon-Mann-Whitney rank sum test

Εισαγωγή

- Το τεστ που θα μελετήσουμε χρησιμοποιείται για ανίχνευση διαφορών στην κατανομή *ανεξάρτητων* παρατηρήσεων $(X_1, \ldots, X_m)$ και $(Y_1, \ldots, Y_n)$.
- Μελετήθηκε για πρώτη φορά από τον Wilcoxon (1945), για την ειδική περίπτωση $m = n$.
- Η περίπτωση $m \neq n$ μελετήθηκε από τους Mann and Whitney (1947) και από τον Wilcoxon (1949).
- Για τους λόγους αυτούς, ο έλεγχος αναφέρεται συχνά και ως Wilcoxon-Mann-Whitney.

Εισαγωγή

- Το τεστ που θα μελετήσουμε χρησιμοποιείται για ανίχνευση διαφορών στην κατανομή *ανεξάρτητων* παρατηρήσεων $(X_1, \ldots, X_m)$ και $(Y_1, \ldots, Y_n)$.
- Μελετήθηκε για πρώτη φορά από τον Wilcoxon (1945), για την ειδική περίπτωση $m = n$.
- Η περίπτωση $m \neq n$ μελετήθηκε από τους Mann and Whitney (1947) και από τον Wilcoxon (1949).
- Για τους λόγους αυτούς, ο έλεγχος αναφέρεται συχνά και ως Wilcoxon-Mann-Whitney.
- Η συνήθης περίπτωση είναι αυτή όπου ο ερευνητής έχει ανεξάρτητα δείγματα από δύο πληθυσμούς και ενδιαφέρεται να ελέγξει αν οι πληθυσμοί ταυτίζονται, πχ
  - τείνουν οι τιμές ενός πληθυσμού να είναι μεγαλύτερες;
  - έχουν οι δύο πληθυσμοί ίσες διαμέσους;
  - έχουν οι δύο πληθυσμοί ίσες μέσες τιμές;

Wilcoxon Rank Sum test (Mann Whitney U)

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. με συνάρτηση κατανομής F.
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. με συνάρτηση κατανομής G.
- Υποθέτουμε οτι X_i και Y_j ανεξάρτητες για κάθε i, j.
- Ενδιαφέρει ο έλεγχος υποθέσεων για την ανιχνευση διαφορών μεταξύ F και G:
  - Π_1 $H_0 : F(t) = G(t), \forall t$ έναντι $H_1 : F(t) \neq G(t)$ για τουλάχιστον ένα t
  - Π_2 $H_0 : F(t) \leqslant G(t), \forall t$ έναντι $H_1 : F(t) > G(t)$ για τουλάχιστον ένα t
  - Π_3 $H_0 : F(t) \geqslant G(t), \forall t$ έναντι $H_1 : F(t) < G(t)$ για τουλάχιστον ένα t

Wilcoxon Rank Sum test (Mann Whitney U)

Οι υποθέσεις αυτές είναι ισοδύναμες με

- Π_1 $H_0 : \mathrm{P}(X_1 < Y_1) = 1/2$ έναντι $H_1 : \mathrm{P}(X_1 < Y_1) \neq 1/2$
- Π_2 $H_0 : \mathrm{P}(X_1 < Y_1) \leqslant 1/2$ έναντι $H_1 : \mathrm{P}(X_1 < Y_1) > 1/2$
- Π_3 $H_0 : \mathrm{P}(X_1 < Y_1) \geqslant 1/2$ έναντι $H_1 : \mathrm{P}(X_1 < Y_1) < 1/2$

Ένα μοντέλο που περιγράφει αυτές τις υποθέσεις είναι όταν η F προέρχεται από μετατόπιση θέσης της G:

$$\exists \Delta : G(t) = F(t - \Delta) \quad \forall t \in \mathbb{R}.$$

Αυτό ισοδυναμεί με το ότι η X έχει την ίδια κατανομή με την $Y + \Delta$.

Στατιστική συνάρτηση αθροίσματος τάξεων (Wilcoxon Rank Sum)

$$W_{m,n} = \sum_{j=1}^{n} R_{N,j}$$

όπου

- $N = m + n$
- $R_{N,j} = \sum_{k=1}^{n} \mathrm{I}_{\{Y_k \leqslant Y_j\}} + \sum_{i=1}^{m} \mathrm{I}_{\{X_i < Y_j\}}$
- Με άλλα λόγια: το $R_{N,j}$ είναι η τάξη του Y_j στο ενιαίο δείγμα $(X_1, \ldots, X_m, Y_1, \ldots, Y_n)$, $j = 1, \ldots, n$.
- Περιπτώσεις ισοβαθμίας αντιμετωπίζονται κατά τον συνήθη τρόπο.
- Η ακριβής κατανομή της $W_{m,n}$ μπορεί να υπολογιστεί αναδρομικά (πχ: πίνακας 9, σημειώσεις Ξεκαλάκη, είτε μέσω της `pwilcox()` στην R)
- Για μεγάλα δείγματα χρησιμοποιούμε προσέγγιση από την κανονική κατανομή.

Παράδειγμα υπολογισμού της $W_{m,n}$

- $x = (2, -1, 1)$ και $y = (3, -2, 0, 2)$
- $m = 3$ και $n = 4$
- $N = 3 + 4 = 7$
- Δημιουργώ το ενιαίο δείγμα $(2, -1, 1, 3, -2, 0, 2)$
- Διατάσσω το ενιαίο δείγμα $(-2, -1, 0, 1, 2, 2, 3)$
- Υπολογίζω τις τάξεις $(1, 2, 3, 4, 5.5, 5.5, 7)$
- $R_7 = (1, 3, 5.5, 7)$
- $W_{3,4} = \sum_{j=1}^{4} R_{7,j} = 1 + 3 + 5.5 + 7 = 16.5$

Διαίσθηση πίσω από την $W_{m,n}$

- Αν οι «κόκκινες» τιμές είναι πολύ μικρότερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα δεξιά

Διαίσθηση πίσω από την $W_{m,n}$

- Αν οι «κόκκινες» τιμές είναι πολύ μικρότερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα δεξιά
  - $R_N = \{1, 2, \ldots, m, m+1, m+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μεγάλες» τιμές

Διαίσθηση πίσω από την $W_{m,n}$

- Αν οι «κόκκινες» τιμές είναι πολύ μικρότερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα δεξιά
  - $R_N = \{1, 2, \ldots, m, m+1, m+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μεγάλες» τιμές
- Αν οι «κόκκινες» τιμές είναι πολύ μεγαλύτερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα αριστερά

Διαίσθηση πίσω από την $W_{m,n}$

- Αν οι «κόκκινες» τιμές είναι πολύ μικρότερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα δεξιά
  - $R_N = \{1, 2, \ldots, m, m+1, m+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μεγάλες» τιμές
- Αν οι «κόκκινες» τιμές είναι πολύ μεγαλύτερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα αριστερά
  - $R_N = \{1, 2, \ldots, n, n+1, n+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μικρές» τιμές

Διαίσθηση πίσω από την $W_{m,n}$

- Αν οι «κόκκινες» τιμές είναι πολύ μικρότερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα δεξιά
  - $R_N = \{1, 2, \ldots, m, m+1, m+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μεγάλες» τιμές
- Αν οι «κόκκινες» τιμές είναι πολύ μεγαλύτερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα αριστερά
  - $R_N = \{1, 2, \ldots, n, n+1, n+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μικρές» τιμές
- Αν οι «κόκκινες» τιμές δεν διαφέρουν «πολύ» από τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο (διατεταγμένο) δείγμα να εναλλάσσονται τυχαία με αυτές των X.

Διαίσθηση πίσω από την $W_{m,n}$

- Αν οι «κόκκινες» τιμές είναι πολύ μικρότερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα δεξιά
  - $R_N = \{1, 2, \ldots, m, m+1, m+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μεγάλες» τιμές
- Αν οι «κόκκινες» τιμές είναι πολύ μεγαλύτερες σε σχέση με τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο δείγμα να είναι στριμωγμένες στα αριστερά
  - $R_N = \{1, 2, \ldots, n, n+1, n+2, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει «μικρές» τιμές
- Αν οι «κόκκινες» τιμές δεν διαφέρουν «πολύ» από τις «μπλε» τιμές, τότε περιμένουμε οι τάξεις των Y στο ενιαίο (διατεταγμένο) δείγμα να εναλλάσσονται τυχαία με αυτές των X.
  - $R_N = \{1, 2, 3, 4, 5, \ldots, m+1, m+2, m+3, \ldots, N\}$
  - Αυτό σημαίνει ότι το $W_{m,n}$ θα λάβει τιμές που θα «μοιάζουν» με τυχαία δειγματοληψία n τιμών (χωρίς επανάθεση) από το σύνολο $\{1, 2, \ldots, N\}$

Σχέση με Mann-Whitney U

- Ισχύει ότι: $R_{N,j} = \sum_{k=1}^{n} \mathrm{I}_{\{Y_k \leqslant Y_j\}} + \sum_{i=1}^{m} \mathrm{I}_{\{X_i < Y_j\}}$
- Συνεπώς

$$\begin{aligned} W_{m,n} &= \sum_{j=1}^{n}\sum_{k=1}^{n} \mathrm{I}_{\{Y_k \leqslant Y_j\}} + \sum_{j=1}^{n}\sum_{i=1}^{m} \mathrm{I}_{\{X_i < Y_j\}} \\ &= \frac{n(n+1)}{2} + U_{m,n} \end{aligned}$$

  όπου $U_{m,n} = \sum_{j=1}^{n}\sum_{i=1}^{m} \mathrm{I}_{\{X_i < Y_j\}}$
- Η στατιστική συνάρτηση $U_{m,n}$ ονομάζεται **Mann-Whitney U**.
- Οι $W_{m,n}$ και $U_{m,n}$ είναι ισοδύναμες όσον αφορά τη συμπερασματαλογία
- Θα χρησιμοποιούμε την $W_{m,n}$

Παράδειγμα υπολογισμού του $U_{m,n}$

- $x = (2, -1, 1)$ και $y = (3, -2, 0, 2)$
- $m = 3$ και $n = 4$
- $N = 3 + 4 = 7$
- Δημιουργώ το ενιαίο δείγμα $(2, -1, 1, 3, -2, 0, 2)$
- Διατάσσω το ενιαίο δείγμα $(-2, -1, 0, 1, 2, 2, 3)$
- Υπολογίζω τις τάξεις $(1, 2, 3, 4, 5.5, 5.5, 7)$
- $R_7 = (1, 3, 5.5, 7)$
- $W_{3,4} = \sum_{j=1}^{4} R_{7,j} = 1 + 3 + 5.5 + 7 = 16.5$
- $U_{3,4} = W_{3,4} - 4 \times 5/2 = 6.5$

Προσεγγιστική κατανομή για $W_{m,n}$

Πρόταση : Υπό την H_0 ισχύει ότι

$$\mathrm{E}W_{m,n} = \frac{n(N+1)}{2} \tag{1}$$

$$\mathrm{Var}W_{m,n} = \frac{n(N+1)(N-n)}{12}. \tag{2}$$

- Απόδειξη: επόμενη σελίδα.
- Κανονική προσέγγιση

$$\begin{aligned} Z &= \frac{W_{m,n} - \mathrm{E}W_{m,n}}{\sqrt{\mathrm{Var}W_{m,n}}} \\ &= \frac{W_{m,n} - n(N+1)/2}{\sqrt{n(N+1)(N-n)/12}} \stackrel{m,n\to\infty}{\to} \mathcal{N}(0,1) \end{aligned}$$

Απόδειξη των (1) και (2)

- Το πρόβλημα είναι ισοδύναμο με την εύρεση της μέσης τιμής και διασποράς της τ.μ. που ισούται με το άθροισμα n ακεραίων που έχουν επιλεγεί τυχαία και χωρίς επανάθεση από το σύνολο $\{1, 2, \ldots, N\}$

Απόδειξη των (1) και (2)

- Το πρόβλημα είναι ισοδύναμο με την εύρεση της μέσης τιμής και διασποράς της τ.μ. που ισούται με το άθροισμα n ακεραίων που έχουν επιλεγεί τυχαία και χωρίς επανάθεση από το σύνολο $\{1, 2, \ldots, N\}$
- Έστω $Z_1, Z_2, \ldots, Z_n$ τ.μ που προκύπτουν τυχαία και χωρίς επανάθεση από το σύνολο $\{1, 2, \ldots, N\}$.

Απόδειξη των (1) και (2)

- Το πρόβλημα είναι ισοδύναμο με την εύρεση της μέσης τιμής και διασποράς της τ.μ. που ισούται με το άθροισμα n ακεραίων που έχουν επιλεγεί τυχαία και χωρίς επανάθεση από το σύνολο $\{1, 2, \ldots, N\}$
- Έστω $Z_1, Z_2, \ldots, Z_n$ τ.μ που προκύπτουν τυχαία και χωρίς επανάθεση από το σύνολο $\{1, 2, \ldots, N\}$.
- Προφανώς $W_{m,n} \overset{H_0}{=} \sum_{j=1}^{n} Z_j$

Απόδειξη των (1) και (2)

- Το πρόβλημα είναι ισοδύναμο με την εύρεση της μέσης τιμής και διασποράς της τ.μ. που ισούται με το άθροισμα n ακεραίων που έχουν επιλεγεί τυχαία και χωρίς επανάθεση από το σύνολο $\{1, 2, \ldots, N\}$
- Έστω $Z_1, Z_2, \ldots, Z_n$ τ.μ που προκύπτουν τυχαία και χωρίς επανάθεση από το σύνολο $\{1, 2, \ldots, N\}$.
- Προφανώς $W_{m,n} \stackrel{H_0}{=} \sum_{j=1}^{n} Z_j$
- Συνεπώς

$$\begin{aligned} \mathrm{E}W_{m,n} &= \mathrm{E}\left(\sum_{j=1}^{n} Z_j\right) = \sum_{j=1}^{n} \mathrm{E}Z_j = n\mathrm{E}Z_1 \\ &= n\sum_{j=1}^{N} j\frac{1}{N} = \frac{n}{N}\frac{N(N+1)}{2} = \frac{n(N+1)}{2} \end{aligned}$$

και η απόδειξη της (1) ολοκληρώθηκε.

Απόδειξη των (1) και (2)

Για τη διασπορά έχουμε

$$
\begin{aligned}
\mathrm{Var} W_{m,n} &= \mathrm{Var}\left(\sum_{j=1}^{n} Z_j\right) \\
&= \sum_{j=1}^{n} \mathrm{Var} Z_j + 2 \sum_{j=1}^{n} \sum_{i<j} \mathrm{Cov}(Z_i, Z_j) \qquad (3)
\end{aligned}
$$

Απόδειξη των (1) και (2)

Για τη διασπορά έχουμε

$$\begin{aligned}\mathrm{Var}W_{m,n} &= \mathrm{Var}\left(\sum_{j=1}^{n} Z_j\right)\\ &= \sum_{j=1}^{n}\mathrm{Var}Z_j + 2\sum_{j=1}^{n}\sum_{i<j}\mathrm{Cov}(Z_i, Z_j) \qquad (3)\end{aligned}$$

- Υπολογισμός $\mathrm{Var}Z_j$

$$\begin{aligned}\mathrm{Var}Z_j = \mathrm{E}(Z_j^2) - (\mathrm{E}Z_j)^2 &= \sum_{j=1}^{N} j^2\frac{1}{N} - \left(\frac{N+1}{2}\right)^2\\ &= \frac{1}{N}\frac{N(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2 = \ldots\\ &= \frac{N^2-1}{12} \qquad (4)\end{aligned}$$

Απόδειξη των (1) και (2)

Για τη συνδιασπορά $\mathrm{Cov}(Z_i, Z_j)$ έχουμε ($i \neq j$):

$$
\begin{aligned}
\mathrm{Cov}(Z_i, Z_j) &= \mathrm{E}(Z_i Z_j) - (\mathrm{E}Z_i)(\mathrm{E}Z_j) \\
&= \sum_{i=1}^{N} \sum_{j \neq i} ij \frac{1}{N} \frac{1}{N-1} - \left(\frac{N+1}{2} \right)^2 \\
&= \frac{1}{N(N-1)} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} ij - \sum_{i=1}^{N} i^2 \right) - \frac{(N+1)^2}{4} \\
&= \frac{1}{N(N-1)} \left(\sum_{i=1}^{N} i \sum_{j=1}^{N} j - \sum_{i=1}^{N} i^2 \right) - \frac{(N+1)^2}{4} = \dots \\
&= -\frac{N+1}{12} \qquad (5)
\end{aligned}
$$

Αντικαθιστώντας τις (4) και (5) στην (3) προκύπτει η (2), και η αποδειξη ολοκληρώθηκε.

Παράδειγμα υπολογισμού του Z

- $x = (2, -1, 1)$ και $y = (3, -2, 0, 2)$
- $m = 3$ και $n = 4$
- $N = 3 + 4 = 7$
- Δημιουργώ το ενιαίο δείγμα $(2, -1, 1, 3, -2, 0, 2)$
- Διατάσσω το ενιαίο δείγμα $(-2, -1, 0, 1, 2, 2, 3)$
- Υπολογίζω τις τάξεις $(1, 2, 3, 4, 5.5, 5.5, 7)$
- $R_7 = (1, 3, 5.5, 7)$
- $W_{3,4} = \sum_{j=1}^{4} R_{7,j} = 1 + 3 + 5.5 + 7 = 16.5$
- $U_{3,4} = W_{3,4} - 4 \times 5/2 = 6.5$

Παράδειγμα υπολογισμού του Z

- $x = (2, -1, 1)$ και $y = (3, -2, 0, 2)$
- $m = 3$ και $n = 4$
- $N = 3 + 4 = 7$
- Δημιουργώ το ενιαίο δείγμα $(2, -1, 1, 3, -2, 0, 2)$
- Διατάσσω το ενιαίο δείγμα $(-2, -1, 0, 1, 2, 2, 3)$
- Υπολογίζω τις τάξεις $(1, 2, 3, 4, 5.5, 5.5, 7)$
- $R_7 = (1, 3, 5.5, 7)$
- $W_{3,4} = \sum_{j=1}^{4} R_{7,j} = 1 + 3 + 5.5 + 7 = 16.5$
- $U_{3,4} = W_{3,4} - 4 \times 5/2 = 6.5$
- Υπολογισμός Z:
  - $\mathrm{E}W_{3,4} = 4(7+1)/2 = 16$
  - $\mathrm{Var}W_{3,4} = 4(7+1)(7-4)/12 = 8$
  - $Z = \frac{W_{m,n} - \mathrm{E}W_{m,n}}{\sqrt{\mathrm{Var}W_{m,n}}} = \frac{16.5-16}{\sqrt{8}} \approx 0.177$

(Προσεγγιστική) συμπερασματολογία για τις υποθέσεις

Π_1 $H_0 : F(t) = G(t), \forall t$ έναντι $H_1 : F(t) \neq G(t)$ για τουλάχιστον ένα t

Π_2 $H_0 : F(t) \leqslant G(t), \forall t$ έναντι $H_1 : F(t) > G(t)$ για τουλάχιστον ένα t

Π_3 $H_0 : F(t) \geqslant G(t), \forall t$ έναντι $H_1 : F(t) < G(t)$ για τουλάχιστον ένα t

όπου

- z_α το άνω α ποσοστιαίο σημείο της $\mathcal{N}(0,1)$, $0 < \alpha < 1$
- z η παρατηρηθείσα τιμή της Z στο δείγμα
- $\Phi(\cdot)$ η συνάρτηση κατανομής της $\mathcal{N}(0,1)$

(Προσεγγιστική) συμπερασματολογία για τις υποθέσεις

Π_1 $H_0 : F(t) = G(t), \forall t$ έναντι $H_1 : F(t) \neq G(t)$ για τουλάχιστον ένα t

- Περιοχή απόρριψης σε επίπεδο σημαντικότητας α: $|Z| > z_{\alpha/2}$
- p-value: $2\mathrm{P}(Z > |z|) = 2(1 - \Phi(|z|))$

Π_2 $H_0 : F(t) \leqslant G(t), \forall t$ έναντι $H_1 : F(t) > G(t)$ για τουλάχιστον ένα t

Π_3 $H_0 : F(t) \geqslant G(t), \forall t$ έναντι $H_1 : F(t) < G(t)$ για τουλάχιστον ένα t

όπου

- z_α το άνω α ποσοστιαίο σημείο της $\mathcal{N}(0, 1)$, $0 < \alpha < 1$
- z η παρατηρηθείσα τιμή της Z στο δείγμα
- $\Phi(\cdot)$ η συνάρτηση κατανομής της $\mathcal{N}(0, 1)$

(Προσεγγιστική) συμπερασματολογία για τις υποθέσεις

Π_1 $H_0 : F(t) = G(t), \forall t$ έναντι $H_1 : F(t) \neq G(t)$ για τουλάχιστον ένα t

- Περιοχή απόρριψης σε επίπεδο σημαντικότητας α: $|Z| > z_{\alpha/2}$
- p-value: $2\mathrm{P}(Z > |z|) = 2(1 - \Phi(|z|))$

Π_2 $H_0 : F(t) \leqslant G(t), \forall t$ έναντι $H_1 : F(t) > G(t)$ για τουλάχιστον ένα t

- Περιοχή απόρριψης σε επίπεδο σημαντικότητας α: $Z > z_\alpha$
- p-value: $\mathrm{P}(Z > z) = 1 - \Phi(z)$

Π_3 $H_0 : F(t) \geqslant G(t), \forall t$ έναντι $H_1 : F(t) < G(t)$ για τουλάχιστον ένα t

όπου

- z_α το άνω α ποσοστιαίο σημείο της $\mathcal{N}(0,1)$, $0 < \alpha < 1$
- z η παρατηρηθείσα τιμή της Z στο δείγμα
- $\Phi(\cdot)$ η συνάρτηση κατανομής της $\mathcal{N}(0,1)$

(Προσεγγιστική) συμπερασματολογία για τις υποθέσεις

Π_1 $H_0 : F(t) = G(t), \forall t$ έναντι $H_1 : F(t) \neq G(t)$ για τουλάχιστον ένα t

- Περιοχή απόρριψης σε επίπεδο σημαντικότητας α: $|Z| > z_{\alpha/2}$
- p-value: $2\mathrm{P}(Z > |z|) = 2(1 - \Phi(|z|))$

Π_2 $H_0 : F(t) \leqslant G(t), \forall t$ έναντι $H_1 : F(t) > G(t)$ για τουλάχιστον ένα t

- Περιοχή απόρριψης σε επίπεδο σημαντικότητας α: $Z > z_\alpha$
- p-value: $\mathrm{P}(Z > z) = 1 - \Phi(z)$

Π_3 $H_0 : F(t) \geqslant G(t), \forall t$ έναντι $H_1 : F(t) < G(t)$ για τουλάχιστον ένα t

- Περιοχή απόρριψης σε επίπεδο σημαντικότητας α: $Z < -z_\alpha$
- p-value: $\mathrm{P}(Z < z) = \Phi(z)$

όπου

- z_α το άνω α ποσοστιαίο σημείο της $\mathcal{N}(0,1)$, $0 < \alpha < 1$
- z η παρατηρηθείσα τιμή της Z στο δείγμα
- $\Phi(\cdot)$ η συνάρτηση κατανομής της $\mathcal{N}(0,1)$

Εφαρμογή

Οι βαθμοί στο μάθημα Στατιστικής σε δύο διαφορετικά Τμήματα είναι:

- τμήμα 1: $(7, 8, 6, 9, 4, 2, 6, 5, 8, 9, 10, 4, 5, 7, 4)$
- τμήμα 2: $(6, 5, 3, 7, 9, 8, 6, 9, 10, 5, 2, 3, 7, 8, 9, 7, 4)$.

Να ελεγχθεί σε επίπεδο σημαντικότητας $\alpha = 5\%$ αν οι βαθμοί στο Τμήμα 2 τείνουν να είναι μεγαλύτεροι από αυτούς των φοιτητών στο Τμήμα 1.

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 $(m = 15)$.

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 $(m = 15)$.
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 2 $(n = 17)$.

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 $(m = 15)$.
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 2 $(n = 17)$.
- Υποθέτουμε οτι X_i και Y_j ανεξάρτητες για κάθε i, j.

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 $(m = 15)$.
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 2 $(n = 17)$.
- Υποθέτουμε οτι X_i και Y_j ανεξάρτητες για κάθε i, j.
- Έστω $F(t)$ η συνάρτηση κατανομής του X_i, $i = 1, \ldots, m$.

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 ($m = 15$).
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 2 ($n = 17$).
- Υποθέτουμε οτι X_i και Y_j ανεξάρτητες για κάθε i, j.
- Έστω $F(t)$ η συνάρτηση κατανομής του X_i, $i = 1, \ldots, m$.
- Έστω $G(t)$ η συνάρτηση κατανομής του Y_j, $j = 1, \ldots, n$.

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 $(m = 15)$.
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 2 $(n = 17)$.
- Υποθέτουμε οτι X_i και Y_j ανεξάρτητες για κάθε i, j.
- Έστω $F(t)$ η συνάρτηση κατανομής του X_i, $i = 1, \ldots, m$.
- Έστω $G(t)$ η συνάρτηση κατανομής του Y_j, $j = 1, \ldots, n$.
- $H_0 : F(t) \leqslant G(t), \forall t$ έναντι της $H_1 : F(t) > G(t)$, για τουλάχιστον ένα t (πρόβλημα Π_2)

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 ($m = 15$).
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 2 ($n = 17$).
- Υποθέτουμε οτι X_i και Y_j ανεξάρτητες για κάθε i, j.
- Έστω $F(t)$ η συνάρτηση κατανομής του X_i, $i = 1, \ldots, m$.
- Έστω $G(t)$ η συνάρτηση κατανομής του Y_j, $j = 1, \ldots, n$.
- $H_0 : F(t) \leqslant G(t), \forall t$ έναντι της $H_1 : F(t) > G(t)$, για τουλάχιστον ένα t (πρόβλημα Π_2)
- Ισοδυναμεί με $H_0 : \mathrm{P}(X_1 < Y_1) \leqslant 1/2$ έναντι της $H_1 : \mathrm{P}(X_1 < Y_1) > 1/2$

Εφαρμογή

- Έστω $X = (X_1, \ldots, X_m)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 1 $(m = 15)$.
- Έστω $Y = (Y_1, \ldots, Y_n)$ ανεξάρτητες και ισόνομες τ.μ. που περιγράφουν την βαθμολογία των φοιτητών στο τμήμα 2 $(n = 17)$.
- Υποθέτουμε οτι X_i και Y_j ανεξάρτητες για κάθε i, j.
- Έστω $F(t)$ η συνάρτηση κατανομής του X_i, $i = 1, \ldots, m$.
- Έστω $G(t)$ η συνάρτηση κατανομής του Y_j, $j = 1, \ldots, n$.
- $H_0 : F(t) \leqslant G(t), \forall t$ έναντι της $H_1 : F(t) > G(t)$, για τουλάχιστον ένα t (πρόβλημα Π_2)
- Ισοδυναμεί με $H_0 : \mathrm{P}(X_1 < Y_1) \leqslant 1/2$ έναντι της $H_1 : \mathrm{P}(X_1 < Y_1) > 1/2$
- Η H_0 σημαίνει ότι οι βαθμολογίες στο τμήμα 1 (X_i) είναι στοχαστικά μεγαλύτερες-ίσες (*stochastically dominant*) από τις βαθμολογίες στο τμήμα 2 (Y_j)

Εφαρμογή

	βαθμός	τμήμα	τάξη	$R_{N,i}$
1	2.00	1.00	1.50	—
2	2.00	2.00	1.50	1.50
3	3.00	2.00	3.50	3.50
4	3.00	2.00	3.50	3.50
5	4.00	1.00	6.50	—
6	4.00	1.00	6.50	—
7	4.00	1.00	6.50	—
8	4.00	2.00	6.50	6.50
9	5.00	1.00	10.50	—
10	5.00	1.00	10.50	—
11	5.00	2.00	10.50	10.50
12	5.00	2.00	10.50	10.50
13	6.00	1.00	14.50	—
14	6.00	1.00	14.50	—
15	6.00	2.00	14.50	14.50
16	6.00	2.00	14.50	14.50
17	7.00	1.00	19.00	—
18	7.00	1.00	19.00	—
19	7.00	2.00	19.00	19.00
20	7.00	2.00	19.00	19.00
21	7.00	2.00	19.00	19.00
22	8.00	1.00	23.50	—
23	8.00	1.00	23.50	—
24	8.00	2.00	23.50	23.50
25	8.00	2.00	23.50	23.50
26	9.00	1.00	28.00	—
27	9.00	1.00	28.00	—
28	9.00	2.00	28.00	28.00
29	9.00	2.00	28.00	28.00
30	9.00	2.00	28.00	28.00
31	10.00	1.00	31.50	—
32	10.00	2.00	31.50	31.50

Εφαρμογή

	βαθμός	τμήμα	τάξη	$R_{N,i}$
1	2.00	1.00	1.50	—
2	2.00	2.00	1.50	1.50
3	3.00	2.00	3.50	3.50
4	3.00	2.00	3.50	3.50
5	4.00	1.00	6.50	—
6	4.00	1.00	6.50	—
7	4.00	1.00	6.50	—
8	4.00	2.00	6.50	6.50
9	5.00	1.00	10.50	—
10	5.00	1.00	10.50	—
11	5.00	2.00	10.50	10.50
12	5.00	2.00	10.50	10.50
13	6.00	1.00	14.50	—
14	6.00	1.00	14.50	—
15	6.00	2.00	14.50	14.50
16	6.00	2.00	14.50	14.50
17	7.00	1.00	19.00	—
18	7.00	1.00	19.00	—
19	7.00	2.00	19.00	19.00
20	7.00	2.00	19.00	19.00
21	7.00	2.00	19.00	19.00
22	8.00	1.00	23.50	—
23	8.00	1.00	23.50	—
24	8.00	2.00	23.50	23.50
25	8.00	2.00	23.50	23.50
26	9.00	1.00	28.00	—
27	9.00	1.00	28.00	—
28	9.00	2.00	28.00	28.00
29	9.00	2.00	28.00	28.00
30	9.00	2.00	28.00	28.00
31	10.00	1.00	31.50	—
32	10.00	2.00	31.50	31.50

- Wilcoxon rank sum:

$$\begin{aligned} W_{m,n} &= \sum_{i=1}^{n} R_{N,i} \\ &= 1.50 + 3.50 + \ldots + 31.5 = 284.5 \end{aligned}$$

Εφαρμογή

	βαθμός	τμήμα	τάξη	$R_{N,i}$
1	2.00	1.00	1.50	–
2	2.00	2.00	1.50	1.50
3	3.00	2.00	3.50	3.50
4	3.00	2.00	3.50	3.50
5	4.00	1.00	6.50	–
6	4.00	1.00	6.50	–
7	4.00	1.00	6.50	–
8	4.00	2.00	6.50	6.50
9	5.00	1.00	10.50	–
10	5.00	1.00	10.50	–
11	5.00	2.00	10.50	10.50
12	5.00	2.00	10.50	10.50
13	6.00	1.00	14.50	–
14	6.00	1.00	14.50	–
15	6.00	2.00	14.50	14.50
16	6.00	2.00	14.50	14.50
17	7.00	1.00	19.00	–
18	7.00	1.00	19.00	–
19	7.00	2.00	19.00	19.00
20	7.00	2.00	19.00	19.00
21	7.00	2.00	19.00	19.00
22	8.00	1.00	23.50	–
23	8.00	1.00	23.50	–
24	8.00	2.00	23.50	23.50
25	8.00	2.00	23.50	23.50
26	9.00	1.00	28.00	–
27	9.00	1.00	28.00	–
28	9.00	2.00	28.00	28.00
29	9.00	2.00	28.00	28.00
30	9.00	2.00	28.00	28.00
31	10.00	1.00	31.50	–
32	10.00	2.00	31.50	31.50

- Wilcoxon rank sum:

$$\begin{aligned} W_{m,n} &= \sum_{i=1}^{n} R_{N,i} \\ &= 1.50 + 3.50 + \ldots + 31.5 = 284.5 \end{aligned}$$

- Mann-Whitney U:

$$U_{m,n} = W_{m,n} - \frac{n(n+1)}{2} = 131.5$$

Εφαρμογή

	βαθμός	τμήμα	τάξη	$R_{N,i}$
1	2.00	1.00	1.50	–
2	2.00	2.00	1.50	1.50
3	3.00	2.00	3.50	3.50
4	3.00	2.00	3.50	3.50
5	4.00	1.00	6.50	–
6	4.00	1.00	6.50	–
7	4.00	1.00	6.50	–
8	4.00	2.00	6.50	6.50
9	5.00	1.00	10.50	–
10	5.00	1.00	10.50	–
11	5.00	2.00	10.50	10.50
12	5.00	2.00	10.50	10.50
13	6.00	1.00	14.50	–
14	6.00	1.00	14.50	–
15	6.00	2.00	14.50	14.50
16	6.00	2.00	14.50	14.50
17	7.00	1.00	19.00	–
18	7.00	1.00	19.00	–
19	7.00	2.00	19.00	19.00
20	7.00	2.00	19.00	19.00
21	7.00	2.00	19.00	19.00
22	8.00	1.00	23.50	–
23	8.00	1.00	23.50	–
24	8.00	2.00	23.50	23.50
25	8.00	2.00	23.50	23.50
26	9.00	1.00	28.00	–
27	9.00	1.00	28.00	–
28	9.00	2.00	28.00	28.00
29	9.00	2.00	28.00	28.00
30	9.00	2.00	28.00	28.00
31	10.00	1.00	31.50	–
32	10.00	2.00	31.50	31.50

- Wilcoxon rank sum:

$$\begin{aligned} W_{m,n} &= \sum_{i=1}^{n} R_{N,i} \\ &= 1.50 + 3.50 + \ldots + 31.5 = 284.5 \end{aligned}$$

- Mann-Whitney U:

$$U_{m,n} = W_{m,n} - \frac{n(n+1)}{2} = 131.5$$

- Προσεγγιστική ελεγχοσυνάρτηση:

$$\begin{aligned} z &= \frac{W_{m,n} - n(N+1)/2}{\sqrt{n(N+1)(N-n)/12}} \\ &= \frac{284.5 - 280.5}{\sqrt{701.25}} \approx 0.15 < z_{0.05} \approx 1.64 \end{aligned}$$

Εφαρμογή

	βαθμός	τμήμα	τάξη	$R_{N,i}$
1	2.00	1.00	1.50	–
2	2.00	2.00	1.50	1.50
3	3.00	2.00	3.50	3.50
4	3.00	2.00	3.50	3.50
5	4.00	1.00	6.50	–
6	4.00	1.00	6.50	–
7	4.00	1.00	6.50	–
8	4.00	2.00	6.50	6.50
9	5.00	1.00	10.50	–
10	5.00	1.00	10.50	–
11	5.00	2.00	10.50	10.50
12	5.00	2.00	10.50	10.50
13	6.00	1.00	14.50	–
14	6.00	1.00	14.50	–
15	6.00	2.00	14.50	14.50
16	6.00	2.00	14.50	14.50
17	7.00	1.00	19.00	–
18	7.00	1.00	19.00	–
19	7.00	2.00	19.00	19.00
20	7.00	2.00	19.00	19.00
21	7.00	2.00	19.00	19.00
22	8.00	1.00	23.50	–
23	8.00	1.00	23.50	–
24	8.00	2.00	23.50	23.50
25	8.00	2.00	23.50	23.50
26	9.00	1.00	28.00	–
27	9.00	1.00	28.00	–
28	9.00	2.00	28.00	28.00
29	9.00	2.00	28.00	28.00
30	9.00	2.00	28.00	28.00
31	10.00	1.00	31.50	–
32	10.00	2.00	31.50	31.50

- Wilcoxon rank sum:

$$\begin{aligned} W_{m,n} &= \sum_{i=1}^{n} R_{N,i} \\ &= 1.50 + 3.50 + \ldots + 31.5 = 284.5 \end{aligned}$$

- Mann-Whitney U:

$$U_{m,n} = W_{m,n} - \frac{n(n+1)}{2} = 131.5$$

- Προσεγγιστική ελεγχοσυνάρτηση:

$$\begin{aligned} z &= \frac{W_{m,n} - n(N+1)/2}{\sqrt{n(N+1)(N-n)/12}} \\ &= \frac{284.5 - 280.5}{\sqrt{701.25}} \approx 0.15 < z_{0.05} \approx 1.64 \end{aligned}$$

- p-value: $\mathrm{P}(Z > 0.15) \approx 0.44 > \alpha = 0.05$

Εφαρμογή

	βαθμός	τμήμα	τάξη	$R_{N,i}$
1	2.00	1.00	1.50	–
2	2.00	2.00	1.50	1.50
3	3.00	2.00	3.50	3.50
4	3.00	2.00	3.50	3.50
5	4.00	1.00	6.50	–
6	4.00	1.00	6.50	–
7	4.00	1.00	6.50	–
8	4.00	2.00	6.50	6.50
9	5.00	1.00	10.50	–
10	5.00	1.00	10.50	–
11	5.00	2.00	10.50	10.50
12	5.00	2.00	10.50	10.50
13	6.00	1.00	14.50	–
14	6.00	1.00	14.50	–
15	6.00	2.00	14.50	14.50
16	6.00	2.00	14.50	14.50
17	7.00	1.00	19.00	–
18	7.00	1.00	19.00	–
19	7.00	2.00	19.00	19.00
20	7.00	2.00	19.00	19.00
21	7.00	2.00	19.00	19.00
22	8.00	1.00	23.50	–
23	8.00	1.00	23.50	–
24	8.00	2.00	23.50	23.50
25	8.00	2.00	23.50	23.50
26	9.00	1.00	28.00	–
27	9.00	1.00	28.00	–
28	9.00	2.00	28.00	28.00
29	9.00	2.00	28.00	28.00
30	9.00	2.00	28.00	28.00
31	10.00	1.00	31.50	–
32	10.00	2.00	31.50	31.50

- Wilcoxon rank sum:

$$\begin{aligned} W_{m,n} &= \sum_{i=1}^{n} R_{N,i} \\ &= 1.50 + 3.50 + \ldots + 31.5 = 284.5 \end{aligned}$$

- Mann-Whitney U:

$$U_{m,n} = W_{m,n} - \frac{n(n+1)}{2} = 131.5$$

- Προσεγγιστική ελεγχοσυνάρτηση:

$$\begin{aligned} z &= \frac{W_{m,n} - n(N+1)/2}{\sqrt{n(N+1)(N-n)/12}} \\ &= \frac{284.5 - 280.5}{\sqrt{701.25}} \approx 0.15 < z_{0.05} \approx 1.64 \end{aligned}$$

- p-value: $\mathrm{P}(Z > 0.15) \approx 0.44 > \alpha = 0.05$
- Δεν απορριπτω την H_0 έναντι της H_1: οι βαθμοί στο τμήμα 2 δεν είναι μεγαλύτεροι από το τμήμα 1 ($\alpha = 0.05$).

Εφαρμογή: Στην R

- α′ τρόπος: με δικές μας εντολές
- β′ τρόπος: μέσω της base function: `wilcox.test()`
- δείτε το αρχείο `wilcoxon_sum_rank.R`
- Παρατήρηση
  - Η R υπολογίζει το Mann-Whitney U ($U_{m,n}$) και όχι το Wilcoxon rank sum ($W_{m,n}$)
  - Το p-value του ελέγχου δεν επηρεάζεται από αυτό

Εφαρμογή: ακριβής κατανομή

- Υπολογισμός ποσοστιαίου σημείου της ακριβούς κατανομής του $U_{m,n}$
  - `qwilcox(p = 0.05, m = 17, n = 15, lower.tail = F)`
  - 171
- Προηγουμένως όμως υπολογίσαμε ότι $U_{m,n} = 131.5$
- Επειδή $131.5 < 171$, δεν απορρίπτουμε την H_0, έναντι της H_1.
- Υπολογισμός ακριβούς p-value:
  - `pwilcox(q = 131.5, m = 15, n = 17,lower.tail = F)`
  - `0.4408149`
- Παρατηρήστε ότι το προσεγγιστικό p-value που υπολογίσαμε πριν είναι αρκετά κοντά στο (ακριβές) p-value
- Προσοχή: οι παραπάνω συναρτήσεις μπορούν να κρασάρουν την `R` αν τουλάχιστον ένα εκ των m, n είναι «μεγάλο».

Εξαρτημένα δείγματα

Έστω δύο τυχαία δείγματα:

- $X_1, \ldots, X_n$, από πληθυσμό με διάμεσο m_X
- $Y_1, \ldots, Y_n$, από πληθυσμό με διάμεσο m_Y

Τα δύο δείγματα θεωρούνται **εξαρτημένα**.

Σκοπός: Έλεγχος υποθέσεων σχετικά με τη διαφορά των διαμέσων.

i) $H_0 : m_X = m_Y$ έναντι $H_1 : m_X > m_Y$

ii) $H_0 : m_X = m_Y$ έναντι $H_1 : m_X < m_Y$

iii) $H_0 : m_X = m_Y$ έναντι $H_1 : m_X \neq m_Y$

Τυχαίο δείγμα διαφορών

Ορίζουμε τις διαφορές:

$$D_i = X_i - Y_i, \quad i = 1, \ldots, n$$

Τότε οι έλεγχοι ανάγονται σε έλεγχο της διάμεσου m_D του πληθυσμού διαφορών:

i) $H_0 : m_D = 0$ έναντι $H_1 : m_D > 0$

ii) $H_0 : m_D = 0$ έναντι $H_1 : m_D < 0$

iii) $H_0 : m_D = 0$ έναντι $H_1 : m_D \neq 0$

Μπορούν να εφαρμοστεί το Wilcoxon signed rank test για $m_0 = 0$.

Παρατήρηση

Αν το δείγμα $D_1, \ldots, D_n$ προέρχεται από συμμετρικό πληθυσμό:

$$m_D = \mu_D$$

όπου μ_D η μέση τιμή της κατανομής της διαφοράς $X - Y$.
Συνεπώς, η δειγματική διάμεσος προσεγγίζει τη δειγματική μέση τιμή.
Αυτό επιτρέπει γενίκευση από τη διάμεσο στη μέση τιμή σε συμμετρικούς πληθυσμούς.

Παράδειγμα 6

Ένας γιατρός καταγράφει το βάρος 11 ασθενών πριν και μετά από μια τρίμηνη δίαιτα. Τα αποτελέσματα είναι τα ακόλουθα:

Ασθενής	1	2	3	4	5	6	7	8	9	10	11
Πριν	89	84	94	91	90	81	84	89	109	121	78
Μετά	82	79	82	94	95	79	70	71	90	100	79

Με επίπεδο σημαντικότητας 5% ελέγξτε αν η δίαιτα επιφέρει στατιστικά σημαντική διαφοροποίηση στο βάρος.

Λύση

Έστω X και Y οι τυχαίες μεταβλητές που παριστάνουν το βάρος **πριν** και **μετά** τη δίαιτα.
Δεδομένα:

- $X_1, \ldots, X_{11}$ από πληθυσμό με διάμεσο m_X
- $Y_1, \ldots, Y_{11}$ από πληθυσμό με διάμεσο m_Y

Τα δείγματα είναι **εξαρτημένα**, καθώς οι μετρήσεις γίνονται στους ίδιους ασθενείς.
Δημιουργούμε τις διαφορές:

$$D_i = X_i - Y_i, \quad i = 1, \ldots, 11$$

και εξετάζουμε τη διάμεσο m_D των διαφορών.

Δεδομένα δειγμάτων διαφορών

Οι διαφορές καταγράφονται στον πίνακα:

Διαφορά	7	5	12	-3	-5	2	14	18	19	21	-1

Στοιχεία δειγματος:

- Δειγματική διάμεσος: 7
- Δειγματική μέση τιμή: 8.09

Υπόθεση ελέγχου:

$$H_0 : m_D = 0 \quad \text{εναντίον} \quad H_1 : m_D \neq 0$$

Ταξινόμηση απόλυτων διαφορών

Απόλυτες τιμές των διαφορών:

$$|D_1|, \ldots, |D_{11}|$$

Ταξινομημένες και με τάξεις:

$\lvert D_i \rvert$	$\underline{1}$	2	$\underline{3}$	$\underline{5}$	5	7	12	14	18	19	21
$R(\lvert D_i \rvert)$	1	2	3	4.5	4.5	6	7	8	9	10	11

Άθροισμα τάξεων:

$$T^- = 1 + 3 + 4.5 = 8.5, \quad T^+ = \frac{11 \cdot 12}{2} - 8.5 = 57.5$$

Στατιστική για τον έλεγχο:

$$T = \min\{T^+, T^-\} = 8.5$$

Στατιστική συνάρτηση W

Λαμβάνοντας υπόψη έναν δεσμό ($c = 1, d_1 = 2$), χρησιμοποιούμε:

$$W = \frac{T^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum_{l=1}^{c}(d_l^3 - d_l)}{48}}}$$

Με $n = 11$:

$$W = \frac{57.5 - \frac{11 \cdot 12}{4}}{\sqrt{\frac{11 \cdot 12 \cdot 23}{24} - \frac{2^3 - 2}{48}}} = \frac{24.5}{11.24} = 2.18$$

Με διόρθωση συνέχειας:

$$W = \frac{24}{11.24} = 2.14$$

Κριτήριο απόρριψης ($\alpha = 0.05$):

$$|W| \geq z_{0.025} = 1.96$$

Συμπέρασμα

Καθώς $2.14 > 1.96$, **απορρίπτουμε** τη μηδενική υπόθεση H_0.
Συμπέρασμα:

- Υπάρχει στατιστικά σημαντική διαφορά στο βάρος πριν και μετά τη δίαιτα.
- Η απώλεια βάρους είναι σημαντική σε επίπεδο σημαντικότητας 5%.

Έλεγχος ισότητας περισσότερων των δύο πληθυσμιακών διαμέσων
Kruskal-Wallis test

- Έστω ότι έχουμε k το πλήθος πληθυσμούς, με αθροιστικές συναρτήσεις κατανομής F_i, $i = 1, ..., k$, $k \geqslant 3$.
- Επιπλέον, λαμβάνουμε k το πλήθος, ανεξάρτητα μεταξύ τους τυχαία δείγματα, ένα από καθέναν από αυτούς τους $i = 1, ..., k$, πληθυσμούς, μεγέθους n_i, $i = 1, ..., k$, με $n_1 + ... + n_k = n$.
- Έστω $X_{i1}, \ldots, X_{i,n_i}$ οι δειγματικές τιμές από τον i-οστό πληθυσμό, $i = 1, ..., k$.
- Θέλουμε να ελέγξουμε τη μηδενική υπόθεση:

$$H_0 : F_1 (x) = ... = F_k (x) \,,$$

  για κάθε $x \in \mathbb{R}$, όπου $k \geqslant 3$, έναντι της εναλλακτικής υπόθεσης H_1 : ότι υπάρχει τουλάχιστον ένα ζεύγος i, j με $i \neq j$, $i, j = 1, ..., k$, $k \geqslant 3$, τέτοιο ώστε $m_i \neq m_j$, όπου m_l η διάμεσος του l-οστού πληθυσμού, $l = 1, ..., k$.

- Έστω $R(X_{ij})$ $i = 1, ..., k$, $j = 1, ..., n_i$ οι τάξεις των διαθέσιμων δειγματικών τιμών των k δειγμάτων στο σύνολο των n παρατηρήσεων.
- $R_i = \sum_{j=1}^{n_i} R(X_{ij})$, $i = 1, ..., k$
- $\sum_{i=1}^{k} R_i = \frac{n(n+1)}{2}$
- Αν υποθέσουμε ότι η H_0 είναι αληθής, αναμένουμε ότι

$$\frac{R_1}{n_1} = \frac{R_2}{n_2} = ... = \frac{R_k}{n_k}, k \geqslant 3.$$

- Λαμβάνοντας υπόψη ότι $\sum_{i=1}^{k} R_i = \frac{n(n+1)}{2}$, έχουμε ότι:

$$R_1 + \frac{n_2}{n_1} R_1 + ... + \frac{n_k}{n_1} R_1 = \frac{n(n+1)}{2},$$

οπότε:

$$\frac{R_1}{n_1}(n_1 + n_2 + ... + n_k) = \frac{n(n+1)}{2} \Rightarrow \frac{R_1}{n_1} = \frac{R_2}{n_2} = ... = \frac{R_k}{n_k} = \frac{n+1}{2}.$$

Διαίσθηση για το τεστ των Kruskall-Wallis

Για να αποφανθούμε για την αποδοχή ή απόρριψη της μηδενικής υπόθεσης, ότι τα τυχαία δείγματα προέρχονται από τον ίδιο πληθυσμό, είναι να εξετάζουμε αν οι ποσότητες $\frac{R_i}{n_i}$, $i = 1, ..., k$, $k \geqslant 3$, είναι περίπου ίσες μεταξύ τους και ίσες με $(n+1)/2$ ή, εναλλακτικά, αν η ποσότητα:

$$\sum_{i=1}^{k} \left(\frac{R_i}{n_i} - \frac{n+1}{2} \right)^2,$$

είναι κοντά στο μηδέν. Γίνεται αντιληπτό ότι μεγάλες τιμές της παραπάνω ποσότητας θα υποδεικνύουν απόκλιση από την υπόθεση ότι τα k το πλήθος τυχαία δείγματα προέρχονται από τον ίδιο πληθυσμό.

Στηριζόμενοι σε αυτήν την ιδέα, οι Kruskall and Wallis (1952)[1] πρότειναν στην περίπτωση μη ύπαρξης δεσμών τη στατιστική συνάρτηση:

$$\mathrm{KW} = \frac{12}{n\left(n+1\right)} \sum_{i=1}^{k} \frac{1}{n_i} \left(R_i - \frac{n_i \left(n+1\right)}{2} \right)^2 . \tag{6}$$

[1] Kruskal, William H., and W. Allen Wallis. "Use of ranks in one-criterion variance analysis." Journal of the American statistical Association 47.260 (1952): 583-621.

Πρόταση 4.1

Στην περίπτωση μη ύπαρξης δεσμών, δύο ισοδύναμες εκφράσεις της στατιστικής συνάρτησης KW που δόθηκε στη σχέση (6) είναι οι:

$$\mathrm{KW} = \frac{12}{n(n+1)} \sum_{i=1}^{k} n_i \left(\frac{R_i}{n_i} - \frac{(n+1)}{2} \right)^2, \tag{7}$$

και

$$\mathrm{KW} = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1). \tag{8}$$

- Προφανώς, καθώς η στατιστική συνάρτηση KW δίνεται από τη σχέση (7) προκύπτει ότι λαμβάνει την τιμή μηδέν, όταν

$$\frac{R_1}{n_1} = \frac{R_2}{n_2} = \ldots = \frac{R_k}{n_k} = \frac{n+1}{2}.$$

- Η υπόθεση H_0 απορρίπτεται για μεγάλες τιμές της στατιστικής συνάρτησης KW, δηλαδή αν $\mathrm{KW} \geqslant c$, όπου c είναι ένας αριθμός τέτοιος, ώστε:

$$P\left(\mathrm{KW} \geqslant c \,|\, H_0 \text{ αληθής}\right) = a.$$

- Η εύρεση της ακριβούς κατανομής της στατιστικής συνάρτησης KW υπό τη μηδενική υπόθεση είναι πολύ δύσκολη και έχει επιτευχθεί από τους Kruskall-Wallis στην περίπτωση των τριών πληθυσμών $k = 3$ και για μικρά σε μέγεθος δείγματα τέτοια, ώστε $n_i \leq 5$, για $i = 1, 2, 3$.
- Για τη γενικότερη περίπτωση στηριζόμαστε σε ασυμπτωτική προσέγγιση.

Ασυμπτωτικό τεστ των Kruskal-Wallis

- Λαμβάνοντας υπόψη ότι $R_i = \sum_{j=1}^{n_i} R\left(X_{ij}\right)$, $i = 1, ..., k$, $k \geqslant 3$, δηλαδή ότι είναι στην ουσία άθροισμα n_i το πλήθος τυχαίων (όχι ανεξάρτητων) μεταβλητών, μπορεί να εφαρμοστεί το Κεντρικό Οριακό Θεώρημα (για ισόνομες τυχαίες μεταβλητές)
- Υπό τη μηδενική υπόθεση και υποθέτοντας μη ύπαρξη δεσμών στις δειγματικές παρατηρήσεις:

$$\frac{R_i - \mathrm{E}(R_i)}{\sqrt{\mathrm{Var}(R_i)}} \xrightarrow{d} \mathcal{N}\left(0, 1\right).$$

- Επομένως, υπό τη μηδενική υπόθεση,

$$\sum_{i=1}^{k} \left(\frac{R_i - \mathrm{E}(R_i)}{\sqrt{\mathrm{Var}(R_i)}} \right)^2 = \sum_{i=1}^{k} \frac{\left(R_i - \mathrm{E}(R_i)\right)^2}{\mathrm{Var}(R_i)} \xrightarrow{d} \chi^2_{k-1},$$

- $k - 1$ το πλήθος τάξεις R_i ορίζονται ανεξάρτητα, καθώς ισχύει η ισότητα $\sum_{i=1}^{k} R_i = \frac{n(n+1)}{2}$.

Πρόταση 4.2

Υπό τη μηδενική υπόθεση και υποθέτοντας ότι δεν υπάρχουν δεσμοί, αποδεικνύεται ότι:

$$\mathrm{E}\left(R_i\right) = n_i \frac{n+1}{2} \ \textit{και} \ \mathrm{Var}\left(R_i\right) = n_i \frac{\left(n+1\right)\left(n-n_i\right)}{12}.$$

- Υπό τη μηδενική υπόθεση:

$$\sum_{i=1}^{k} \left(\frac{R_i - \mathrm{E}(R_i)}{\sqrt{\mathrm{Var}(R_i)}} \right)^2 = \frac{12}{(n+1)} \sum_{i=1}^{k} \frac{\left(R_i - n_i \frac{n+1}{2}\right)^2}{n_i\left(n-n_i\right)} \xrightarrow{d} \chi^2_{k-1}.$$

- Επί πλέον

$$\mathrm{KW} = \frac{12}{n\left(n+1\right)} \sum_{i=1}^{k} \frac{1}{n_i} \left(R_i - \frac{n_i\left(n+1\right)}{2} \right)^2 \xrightarrow{d} \chi^2_{k-1}$$

- Επομένως, η υπό μελέτη μηδενική υπόθεση απορρίπτεται, αν $\mathrm{KW} \geqslant \chi^2_{k-1,a}$, όπου $\chi^2_{k-1,a}$ είναι το σημείο εκείνο για το οποίο ισχύει ότι $\mathrm{P}\left(\chi^2_{k-1} \geqslant \chi^2_{k-1,a}\right) = a$.

Παράδειγμα

Παράδειγμα 7

Σε ένα εργοστάσιο τρεις μηχανές χρησιμοποιούνται για την παραγωγή δοχείων εμφιαλώσεως. Κατά τη διάρκεια μιας εργάσιμης εβδομάδας καταγράφεται ο αριθμός των δοχείων που κατασκευάστηκαν από κάθε μηχανή και τα αποτελέσματα παρατίθενται στον πίνακα που ακολουθεί, με την επιπλέον επισήμανση ότι κάποιες μέρες δεν παρήχθησαν δοχεία από κάποιες μηχανές λόγω μη λειτουργίας των αντίστοιχων μηχανών.

Μηχανή 1	340	345	330	342	338
Μηχανή 2	339	333	344		
Μηχανή 3	347	343	349	355	

Να ελέγξετε αν υπάρχει στατιστικά σημαντική διαφορά ως προς την παραγωγή δοχείων των τριών μηχανών.

Περιγραφή δεδομένων

Έχουμε διαθέσιμα τρία ανεξάρτητα δείγματα:

$$n_1 = 5, \quad n_2 = 3, \quad n_3 = 4, \quad n = 12.$$

Οι παρατηρήσεις X_{ij}, $j = 1, \ldots, n_i$, $i = 1, 2, 3$, αναμειγνύονται και διατάσσονται σε αύξουσα σειρά, και υπολογίζονται οι τάξεις $R(X_{ij})$.

Παρατήρηση	Τάξη	Δείγμα	Τιμή
1	1	1	330
2	2	2	333
3	3	1	338
4	4	2	339
5	5	1	340
6	6	1	342
7	7	3	343
8	8	2	344
9	9	1	345
10	10	3	347
11	11	3	349
12	12	3	355

Αθροίσματα τάξεων

$$R_1 = 5 + 9 + 1 + 6 + 3 = 24$$

$$R_2 = 4 + 2 + 8 = 14$$

$$R_3 = 10 + 7 + 11 + 12 = 40$$

Στατιστική Kruskal–Wallis

$$
\begin{aligned}
\mathrm{KW} &= \frac{12}{n(n+1)} \sum_{i=1}^{3} \frac{R_i^2}{n_i} - 3(n+1) \\
&= \frac{1}{13} \left(\frac{24^2}{5} + \frac{14^2}{3} + \frac{40^2}{4} \right) - 39 \\
&= 5.6564
\end{aligned}
$$

Συμπέρασμα

Υπό τη μηδενική υπόθεση:

$$\mathrm{KW} \sim \chi^2_2$$

Κρίσιμη τιμή:

$$\chi^2_{2,0.05} = 5.99$$

Εφόσον:

$$5.6564 < 5.99$$

Δεν απορρίπτεται η H_0 στο επίπεδο σημαντικότητας 5%.

$$\text{p-value} = P(\chi^2_2 > 5.6564) \approx 0.059$$

Εξαρτημένα δείγματα: έλεγχος Friedman

Το πρόβλημα

Πειραματική μονάδα	Δείγματα					
	1	2	…	i	…	k
1	X_{11}	X_{21}	…	X_{i1}	…	X_{k1}
2	X_{12}	X_{22}	…	X_{i2}	…	X_{k2}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
j	X_{1j}	X_{2j}	…	X_{ij}	…	X_{kj}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
n	X_{1n}	X_{2n}	…	X_{in}	…	X_{kn}

- k εξαρτημένα δείγματα μεγέθους n
- Παρατηρήσεις X_{ij}, $i = 1, \dots, k$, $j = 1, \dots, n$

Έλεγχος:

$$H_0 : m_1 = \cdots = m_k$$

$$H_1 : m_i \neq m_j \text{ για κάποιο } i \neq j$$

Η ιδέα του Friedman

Πειραματική μονάδα	Δείγματα					
	1	2	$\dots$	i	$\dots$	k
1	R_{11}	R_{21}	$\dots$	R_{i1}	$\dots$	R_{k1}
2	R_{12}	R_{22}	$\dots$	R_{i2}	$\dots$	R_{k2}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
j	R_{1j}	R_{2j}	$\dots$	R_{ij}	$\dots$	R_{kj}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
n	R_{1n}	R_{2n}	$\dots$	R_{in}	$\dots$	R_{kn}

- Οι παρατηρήσεις οργανώνονται σε πίνακα
- Κάθε γραμμή: πειραματική μονάδα
- Κάθε στήλη: δείγμα
- Υπολογίζουμε τάξεις **ανά γραμμή**

Υπολογισμός τάξεων

Για κάθε γραμμή:

- Αντιστοιχούμε τάξεις $1, \ldots, k$
- Συμβολισμός: $R(X_{ij}) = R_{ij}$

Υπολογίζουμε:

$$R_i = \sum_{j=1}^{n} R_{ij}, \quad \bar{R}_i = \frac{R_i}{n}$$

Υπό τη μηδενική υπόθεση

Ισχύει:

$$\mathrm{E}(\bar{R}_i) = \frac{k+1}{2}, \quad \mathrm{Var}(\bar{R}_i) = \frac{k^2-1}{12n}$$

Άρα εξετάζουμε πόσο αποκλίνουν οι $\bar{R}_i$ από το $\frac{k+1}{2}$.

Η στατιστική Friedman

- Ο Friedman πρότεινε:

$$Q = \frac{12n}{k(k+1)} \sum_{i=1}^{k} \left(\bar{R}_i - \frac{k+1}{2} \right)^2$$

- Ισοδύναμα:

$$Q = \frac{12}{nk(k+1)} \sum_{i=1}^{k} R_i^2 - 3n(k+1)$$

- Για μεγάλο n:

$$Q \sim \chi^2_{k-1}$$

  Κριτήριο απόρριψης:

$$Q \geq \chi^2_{k-1,\alpha}$$

- Πότε ισχύει η προσέγγιση·
  - $k = 3$: $n \geq 9$
  - $k \geq 4$: $n \geq 6$
  - Σύμφωνα με άλλες πηγές: $k > 7$

Δεσμοί (ties)

Αν υπάρχουν δεσμοί, χρησιμοποιούμε:

$$Q' = \frac{k-1}{\sum R_{ij}^2 - 0.25nk(k+1)^2} \left(\sum R_i^2 - 0.25n^2k(k+1)^2 \right)$$

$$Q' \sim \chi^2_{k-1}$$

Παράδειγμα

Παράδειγμα 8

Σε έναν διαγωνισμό μαγειρικής, 5 κριτές αξιολόγησαν τα πιάτα που έφτιαξαν οι 8 συμμετέχοντες στον τελικό του διαγωνισμού. Τα αποτελέσματα των αξιολογήσεων δίνονται στον επόμενο πίνακα, όπου περιέχονται οι βαθμολογίες που έδωσε κάθε κριτής. Να ελέγξετε την υπόθεση, σε ε.σ. 10%, ότι ο τρόπος βαθμολογίας των κριτών είναι ανεξάρτητος των φιναλίστ.

Κριτής	Σ1	Σ2	Σ3	Σ4	Σ5	Σ6	Σ7	Σ8
1	92	75	68	80	88	90	82	85
2	60	60	75	88	88	90	92	85
3	85	89	89	60	70	70	82	92
4	90	60	78	78	85	88	65	92
5	80	70	83	83	60	65	90	90

Πίνακας: Βαθμολογίες κριτών για τους 8 συμμετέχοντες.

Λύση

- Χρησιμοποιούμε το τεστ του Friedman με στατιστική Q', λόγω δεσμών.
- $k = 8$ δείγματα, $n = 5$ πειραματικές μονάδες.
- Οι τάξεις $R(X_{ij})$ δίνονται στον ακόλουθο πίνακα.

Υπολογισμός τάξεων ανά γραμμή

	Σ1	Σ2	Σ3	Σ4	Σ5	Σ6	Σ7	Σ8
1	8	2	1	3	6	7	4	5
2	1.5	1.5	3	6	6	7	8	5
3	5	6.5	6.5	1	2.5	2.5	4	8
4	7	1	3.5	3.5	5	6	2	8
5	4	3	5.5	5.5	1	2	7.5	7.5

Άθροισμα τάξεων για κάθε δείγμα

$$
\begin{aligned}
R_1 &= 8 + 1.5 + 5 + 7 + 4 = 25.5 \\
R_2 &= 2 + 1.5 + 6.5 + 1 + 3 = 14 \\
R_3 &= 1 + 3 + 6.5 + 3.5 + 5.5 = 19.5 \\
R_4 &= 3 + 6 + 1 + 3.5 + 5.5 = 19 \\
R_5 &= 6 + 6 + 2.5 + 5 + 1 = 20.5 \\
R_6 &= 7 + 7 + 2.5 + 6 + 2 = 24.5 \\
R_7 &= 4 + 8 + 4 + 2 + 7.5 = 25.5 \\
R_8 &= 5 + 5 + 8 + 8 + 7.5 = 33.5
\end{aligned}
$$

Άλλα αθροίσματα και σταθερές

$$\sum_{i=1}^{8}\sum_{j=1}^{5} R(X_{ij})^2 = 1037$$

$$\sum_{i=1}^{8} R_i^2 = 4380.5$$

$$0.25nk(k+1)^2 = 810, \quad 0.25n^2k(k+1)^2 = 4050$$

Υπολογισμός στατιστικής Q'

Η στατιστική:

$$Q' = \frac{k-1}{\sum R_{ij}^2 - 0.25nk(k+1)^2}\left(\sum R_i^2 - 0.25n^2k(k+1)^2\right)$$

Αντικατάσταση:

$$Q' = \frac{8-1}{1037-810}(4380.5 - 4050) \approx 10.192$$

Έλεγχος υπόθεσης

- Προσεγγιστική κατανομή: $Q' \sim \chi^2_7$
- Επίπεδο σημαντικότητας: $\alpha = 0.10$
- Άνω 0.10-ποσοστιαίο σημείο: $\chi^2_{7,0.10} = 12.017$
- Απόφαση:
$$Q' = 10.192 < 12.017$$

Συμπέρασμα

- Δεν απορρίπτουμε την H_0 σε επίπεδο σημαντικότητας 10%.
- Δηλαδή, δεν έχουμε σαφείς ενδείξεις ότι ο τρόπος βαθμολόγησης των κριτών εξαρτάται από τους φιναλίστ.