Μη-παραμετρικοί έλεγχοι

Ζιντζαράς Ηλίας, M.Sc., Ph.D.

Καθηγητής Βιομαθηματικών-Βιομετρίας
Εργαστήριο Βιομαθηματικών
Τμήμα Ιατρικής
Πανεπιστήμιο Θεσσαλίας

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

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

Τι σημαίνει «παραμετρικό»;

- Οι περισσότερες στατιστικές μέθοδοι (t-test, ANOVA, απλή γραμμική παλινδρόμηση κ.λπ.) κάνουν υποθέσεις σχετικά με την κατανομή πιθανοτήτων του υπό ανάλυση πληθυσμού (κανονική κατανομή)
- Από αυτήν την υπόθεση μπορούμε να δημιουργήσουμε δειγματικές κατανομές
- Και από τις δειγματικές κατανομές μπορούμε να αντλήσουμε στατιστικά δείγματος (μέση τιμή, τυπική απόκλιση, κ.λπ.)

Τι σημαίνει «μη-παραμετρικό»;

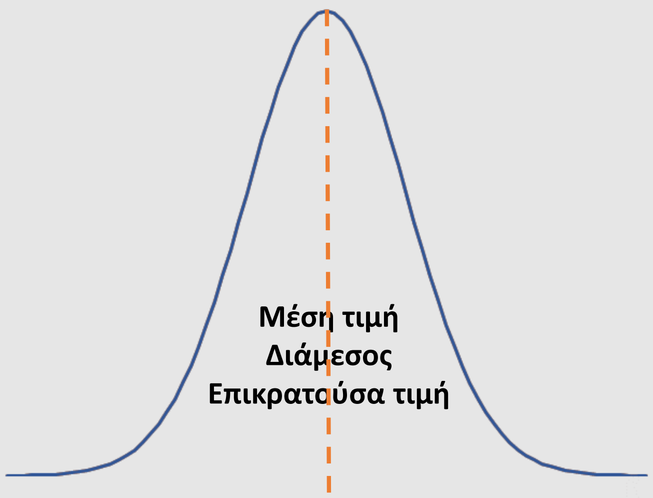
- Οι μη-παραμετρικές μέθοδοι δεν κάνουν αυτές τις υποθέσεις
- Ενώ οι παραμετρικές μέθοδοι προϋποθέτουν συνήθως ποσοτικά δεδομένα, οι μη παραμετρικές μέθοδοι μας επιτρέπουν να δουλεύουμε και με ποιοτικά δεδομένα (nominal / ordinal)
- Τις περισσότερες φορές, ακόμη και ποσοτικά δεδομένα μετατρέπονται σε ποιοτικά για χρήση σε μη-παραμετρικές μεθόδους· ο πιο κοινός τύπος είναι οι τάξεις παρατηρήσεων (ranked observations)

Πότε χρησιμοποιούνται;

- Όταν το μέγεθος των δειγμάτων είναι μικρό
- Όταν τα δεδομένα δεν κατανέμονται κανονικά
- Όταν τα δεδομένα δεν έχουν παρόμοια διακύμανση
- Όταν τα δεδομένα είναι διατάξιμα (ordinal)
- Όταν υπάρχουν πολλές ακραίες τιμές και κάποιος μετασχηματισμός δεν διορθώνει το πρόβλημα

Παραμετρικά test: Κανονικότητα

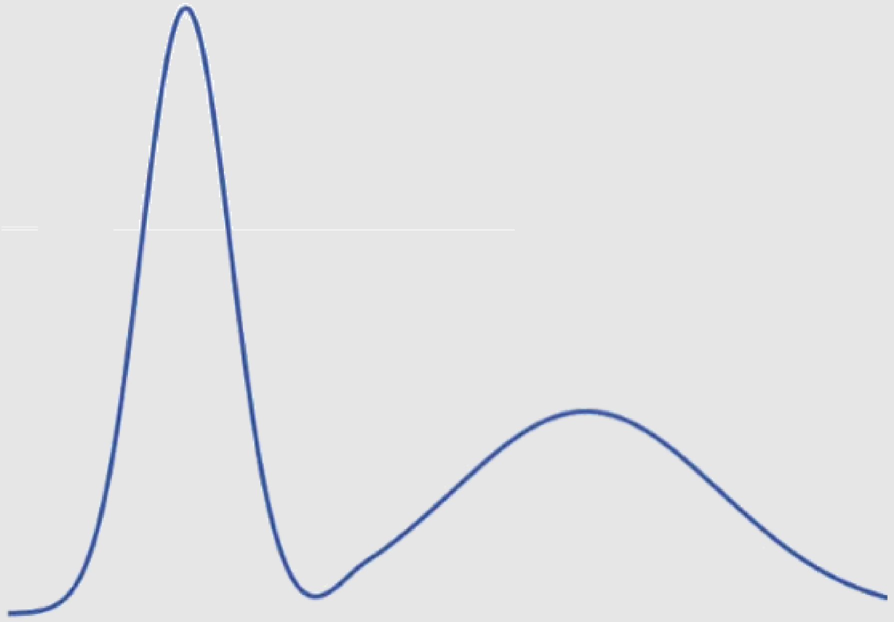
Συμμετρία



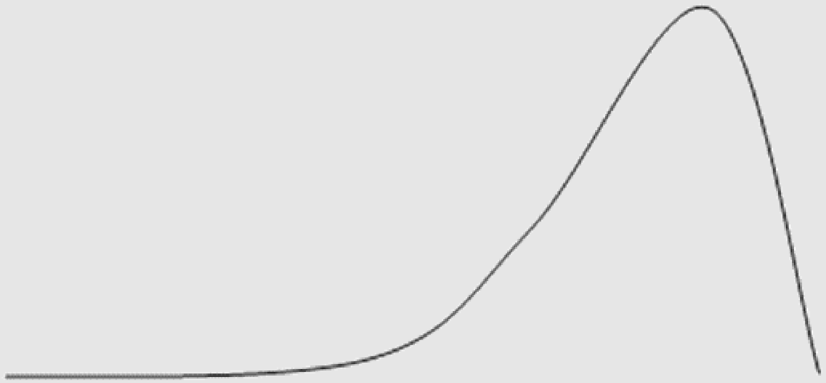
Συμμετρία

Τα παραμετρικά test προϋποθέτουν την κανονικότητα των δεδομένων

Μη Παραμετρικά test: Κανονικότητα



Bimodal distribution



Skewed Distribution

Τα μη παραμετρικά test ΔΕΝ προϋποθέτουν την κανονικότητα των δεδομένων

Πλεονεκτήματα

- Εφαρμόζονται σε **ποσοτικά χαρακτηριστικά των οποίων οι κατανομές είναι μη κανονικές**, ανεξάρτητα από τον αριθμό των παρατηρήσεων.
- Είναι δυνατό να εφαρμοστούν σε **ποσοτικά χαρακτηριστικά**, όταν ο αριθμός των παρατηρήσεων είναι περιορισμένος και οι κατανομές άγνωστες ή φαίνεται να εκτρέπονται από την μορφή της κανονικής κατανομής.
- Είναι δυνατό να εφαρμοστούν σε ι**εραρχικώς διατάξιμα (ordinal) χαρακτηριστικά**.
- Είναι πολύ απλούστερες όσον αφορά τους απαιτούμενους αριθμητικούς υπολογισμούς.

Μειονεκτήματα

- Όταν οι **παραμετρικές δοκιμασίες** είναι δυνατό να εφαρμοστούν, τότε αυτές διαθέτουν **μεγαλύτερη ισχύ** από εκείνη των αντίστοιχων μη παραμετρικών δοκιμασιών
- Οι μη παραμετρικές μέθοδοι δεν είναι δυνατό να εφαρμοστούν σε **σύνθετες στατιστικές αναλύσεις** (όπως η ανάλυση μεταβλητότητας και ο έλεγχος των αλληλεπιδράσεων - analysis of variance and test for interactions)
- Επιπλέον ο υπολογισμός των ορίων αξιοπιστίας είναι δύσκολος

Κυριότερες μη παραμετρικές δοκιμασίες

- Η δοκιμασία του **Wilcoxon για παρατηρήσεις κατά ζεύγη** η οποία αντιστοιχεί στην παραμετρική δοκιμασία κατά ζεύγη t – test (paired sample t – test)
- Η δοκιμασία **Mann-Whitney-Wilcoxon test** για τη σύγκριση δύο ανεξάρτητων δειγμάτων η οποία αντιστοιχεί στην παραμετρική δοκιμασία t – test για ανεξάρτητα δείγματα (independent samples t – test)
- Η δοκιμασία **Kruskal-Wallis** για την σύγκριση περισσότερων από δύο ανεξάρτητων δειγμάτων η οποία αντιστοιχεί στην παραμετρική δοκιμασία One Way ANOVA
- Ο βαθμός συσχέτισης μεταξύ ποσοτικών χαρακτηριστικών που δεν κατανέμονται κανονικά ή μεταξύ χαρακτηριστικών ιεραρχικά διατάξιμων, μπορεί να ελεγχθεί με το μη παραμετρικό συντελεστή συσχέτισης του **Spearman**, ο οποίος αντιστοιχεί στον παραμετρικό συντελεστή συσχέτισης του **Pearson**

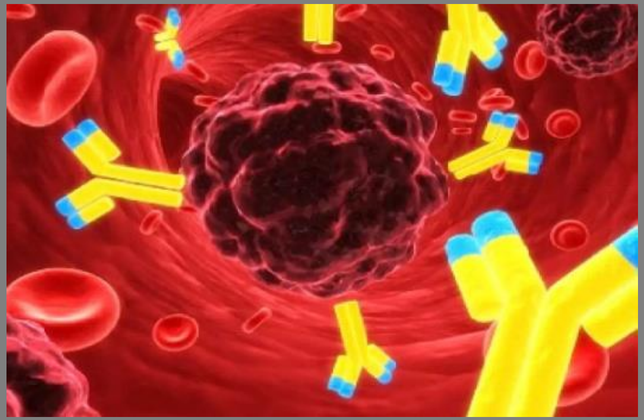
Κυριότερες μη παραμετρικές δοκιμασίες

Παραμετρική δοκιμασία	Μη παραμετρική δοκιμασία
Paired samples t - test	**Wilcoxon για παρατηρήσεις κατά ζεύγη**
Independent samples t - test	**Mann-Whitney-Wilcoxon test**
One way anova	**Kruskal-Wallis**
Συντελεστή συσχέτισης του Pearson	**Συντελεστή συσχέτισης του Spearman**

Σύγκριση των επιπέδων ανοσοσφαιρίνης Μ ορού σε ενήλικες πριν και μετά τον εμβολιασμό κατά της χολέρας

Δοκιμασία Wilcoxon για παρατηρήσεις κατά ζεύγη

(Paired sample t-test)



Ιστορικό

- Για τη σύγκριση παρατηρήσεων κατά ζεύγη κάνοντας χρήση ενός παραμετρικού τεστ θα κάναμε χρήση του paired t-test· πριν/μετά
- Στην περίπτωση αυτή θα ελέγχαμε αν η μέση διαφορά μεταξύ των ζευγών είναι μηδέν, $\bar{d} = 0$
- Η υπόθεση που κάνουμε είναι πως ο κάθε πληθυσμός είναι κανονικά κατανεμημένος
- Το Wilcoxon Signed-Rank test μας επιτρέπει να συγκρίνουμε δύο πληθυσμούς των οποίων οι κατανομές δεν είναι κανονικά κατανεμημένες
- Η μόνη απαίτηση είναι πως η κατανομή των διαφορών είναι συμμετρική (median), όχι απαραίτητα κανονικά κατανεμημένη

Στάδια εφαρμογής δύο ομάδων Α και Β

a) Υπολογίζονται οι διαφορές (Α – Β) για κάθε ζεύγος παρατηρήσεων.
Οι διαφορές ενδέχεται να είναι θετικές, αρνητικές ή ίσες με το μηδέν. Τα ζεύγη παρατηρήσεων για τα οποία η διαφορά είναι ίση με μηδέν **ΕΞΑΙΡΟΥΝΤΑΙ** από την παραπέρα ανάλυση,

b) Καταχωρούνται οι απόλυτες τιμές των διαφορών Α – Β,

c) Οι απόλυτες τιμές διατάσσονται με σειρά απόλυτου μεγέθους, έτσι ώστε η μικρότερη απόλυτη τιμή να παίρνει την $1^η$ θέση, η αμέσως μεγαλύτερη τιμή τη $2^η$ θέση κ.ο.κ. Αν δύο η περισσότερες απόλυτες τιμές διαφορών έχουν το ίδιο μέγεθος, τότε παίρνουν την ίδια θέση, και μάλιστα τη μέση τιμή των θέσεων, τις οποίες θα έπαιρναν αν δεν σημειώνονταν οι ισοβαθμίες.

d) Το σημείο (+ ή -) κάθε διαφοράς (Α - Β) επαναφέρεται στη «θέση» που αντιστοιχεί στη διαφορά αυτή (δηλαδή στη θέση που παίρνει η απόλυτη τιμής της).

e) Αθροίζονται χωριστά οι θέσεις με θετικό σημείο (το απόλυτο άθροισμά τους συμβολίζεται με T_+) και οι θέσεις με αρνητικό σημείο (το απόλυτο άθροισμά τους συμβολίζεται με T_-)

Στάδια εφαρμογής δύο ομάδων Α και Β

f) Όταν οι απόλυτες τιμές των T_+ και T_- βρίσκονται πολύ κοντά τότε **δεν** υπάρχει στατιστικά σημαντική διαφορά μεταξύ των ομάδων Α και Β. Όσο περισσότερο διαφέρουν οι (απόλυτες) τιμές των T_+ και T_-, τόσο μεγαλύτερη είναι η πιθανότητα να υπάρχει πραγματική διαφορά μεταξύ των ομάδων Α και Β (όταν $T_+ > T_-$ τότε Α > Β, και όταν $T_+ < T_-$ τότε Α < Β),

g) Ο έλεγχος της στατιστικής σημαντικότητας γίνεται με τον διπλανό πίνακα. Αν η τιμή του T_+ (ή T_-) είναι ίση ή μικρότερη από την κατώτερη τιμή του πίνακα ή ίση ή μεγαλύτερη από την ανώτερη τιμή του πίνακα, τότε η διαφορά μεταξύ των δύο ομάδων είναι στατιστικά σημαντική στο αντίστοιχο επίπεδο.

Αριθμός ζευγών χωρίς ισοβαθμίες	**Στατιστικά σημαντική στο επίπεδο 0.05**	
	Κατώτερο όριο	Ανώτερο όριο
5	–	–
6	0	21
7	2	26
8	3	33
9	5	40
10	8	47
11	10	56
…		

Δειγματική κατανομή

$$Mean:\ \mu_{T^+} = \frac{n(n+1)}{4}$$

$$Standart\ deviation:\ \sigma_{T^+} = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$n =$ αριθμός ζευγών εκτός των ισοβαθμιών

Κανονική (ή κατά προσέγγιση) κανονική κατανομή για $n \geq 20$

Test στατιστική και πιθανότητα

Τι είναι το μ_{T^+};

Είναι το T^+ στο δεξί άκρο ή αριστερό άκρο της δειγματικής κατανομής σε σχέση με το μ_{T^+};

Αν το T^+ είναι στο δεξί άκρο:

$$P\left(Z \geq \frac{\mathrm{T}^+ - \mu_{\mathrm{T}^+}}{\sigma_{\mathrm{T}^+}}\right)$$

Αν το T^+ είναι στο αριστερό άκρο:

$$P\left(Z \leq \frac{\mathrm{T}^+ - \mu_{\mathrm{T}^+}}{\sigma_{\mathrm{T}^+}}\right)$$

Εφαρμογή διόρθωσης στο T^+

Μετά διπλασιάζουμε την πιθανότητα στα άκρα για να γίνει διπλής ουράς (two-tailed)

Σύγκριση των επιπέδων ανοσοσφαιρίνης M ορού σε ενήλικες πριν και μετά τον εμβολιασμό κατά της χολέρας

Μετρήθηκαν οι ανοσοσφαιρίνες M του ορού 9 ενήλικων ατόμων πριν και 15 μέρες μετά τον αντιχολερικό εμβολιασμό. Οι τιμές που βρέθηκαν δίνονται στις στήλες 2 και 3 του επόμενου πίνακα.

Να αξιολογηθεί η διαφορά των τιμών των ανοσοσφαιρινών αυτών πριν και μετά τον εμβολιασμό.

Α/Α	Ανοσοσφαιρίνες Μ		Διαφορά (2) – (3)	Απόλυτες τιμές διαφοράς	Ανοδική σειρά (θέση) απόλυτων διαφορών	Οι θέσεις της (6) με τα σημεία της (4)
	Εμβολιασμός					
	Πριν	Μετά				
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	158	168	−10	10	3 - 3η	-3
2	140	168	−28	28	8 - 8η θέση	-8
3	165	161	+4	4	1 - 1η θέση, το 4 η μικρότερη τιμή	+1
4	100	120	−20	20	7 - 7η θέση	−7
5	110	122	−12	12	4.5 - ((4η+5η/2=4.5)	−4.5
6	142	134	+ 8	8	2 - 2η θέση, το 8 η αμέσως μεγαλύτερη	+2
7	104	116	−12	12	4.5 - ((4η θέση+5η θέση)/2=4.5)	−4.5
8	254	270	−16	16	6 - 6η θέση	−6
9	180	180	0	(εξαιρείται από την παραπέρα ανάλυση)		

T_+ έναντι T_-

Α/Α	Ανοσοσφαιρίνες Μ		Θέσεις
	Εμβολιασμός		
	Πριν	Μετά	
1	158	168	-3
2	140	168	-8
3	165	161	$+1$
4	100	120	-7
5	110	122	-4.5
6	142	134	$+2$
7	104	116	-4.5
8	254	270	-6
9	180	180	0

Αθροίζονται χωριστά οι θέσεις με θετικό σημείο και οι θέσεις με αρνητικό σημείο, οπότε προκύπτει:

$$T_+ = 1 + 2 = 3$$

$$T_- = (-3) + (-8) + (-7) + (-4.5) + (-4.5) + (-6) = -33$$

Απόλυτη τιμή του T_-

$$|T_-| = |-33| = 33$$

Στατιστική σημαντικότητα

Αριθμός ζευγών χωρίς ισοβαθμίες	Στατιστικά σημαντική στο επίπεδο 0.05	
	Κατώτερο όριο	Ανώτερο όριο
5	–	–
6	0	21
7	2	26
8	**3**	**33**
9	5	40
10	8	47
11	10	56
…		

Το $T_+ = 3 \leq$ από το κατώτερο όριο (3) για δείγμα μεγέθους 8 (δεν υπολογίζονται οι ισοβαθμίες).

Επομένως η διαφορά είναι στατιστικά σημαντική σε στάθμη σημαντικότητας 5%.

Οπότε μπορεί να υποστηριχτεί ότι 15 μέρες μετά τον εμβολιασμό υπάρχει αύξηση των ανοσοσφαιρινών M του ορού.

Σημείωση: Όταν η τιμή T_+ είναι ίση η μικρότερη από την κατώτερη τιμή του πίνακα, η τιμή T_- είναι υποχρεωτικά ίση ή μεγαλύτερη από την ανώτερη τιμή του πίνακα και αντίστροφα

Αποτελέσματα δειγματικής κατανομής

$$Mean:\ \mu_{T^+} = \frac{8(8+1)}{4} = \frac{72}{4} = 18$$

Θυμηθείτε πως είχαμε μία διαφορά που ήταν μηδέν, οπότε το $n = 9$ μειώθηκε σε $n = 8$.

$$Standart\ deviation:\ \sigma_{T^+} = \sqrt{\frac{8(8+1)(2\cdot 8+1)}{24}} = 7.14$$

$n =$ αριθμός ζευγών εκτός των ισοβαθμιών

Κανονική (ή κατά προσέγγιση) κανονική κατανομή για $n \geq 20$

Στατιστική, αποτελέσματα υπόθεσης

$$Z = \frac{\mathrm{T}^{+} - \mu_{\mathrm{T}^{+}}}{\sigma_{\mathrm{T}^{+}}} \qquad Z = \frac{3 - 18}{7.14} \qquad Z = -2.101$$

$$p - value = \mathrm{NORM.S.DIST}(-2.101, \mathrm{TRUE})$$

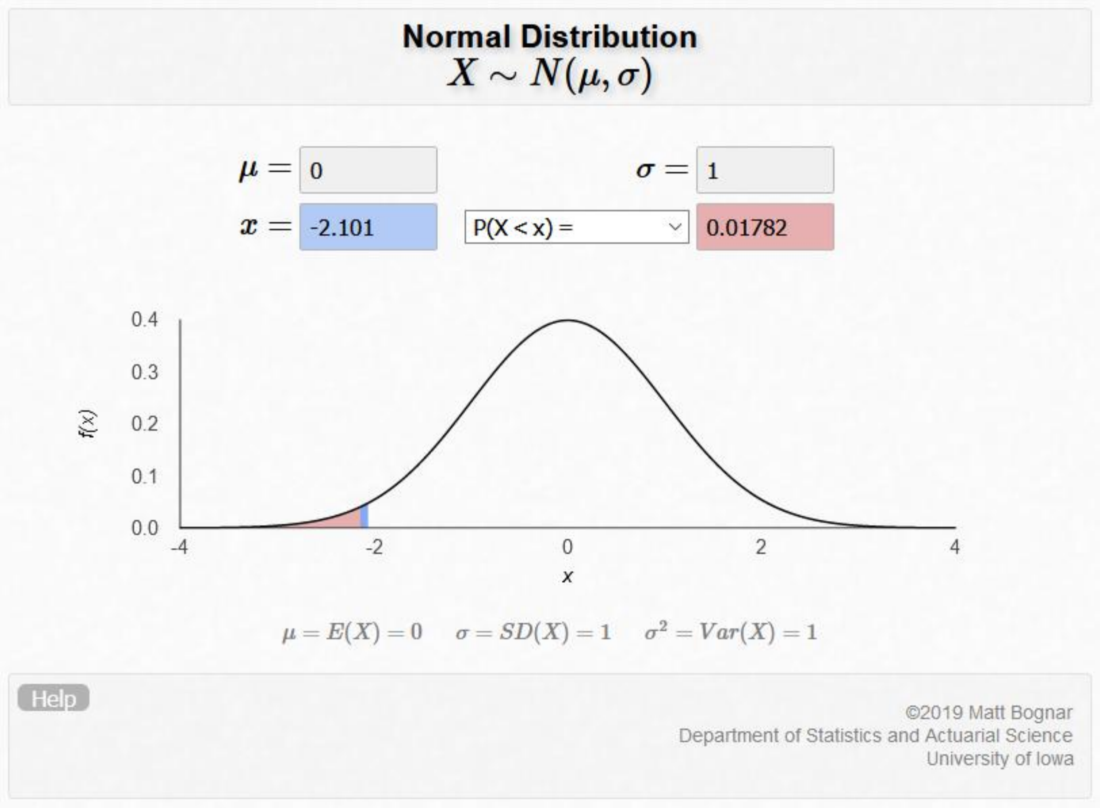
$$p - value\ (\mathrm{lower} - \mathrm{tailed}) = 0.01784595$$

Διπλασιάζουμε την πιθανότητα στο άκρο για να γίνει διπλής ουράς (two-tailed)

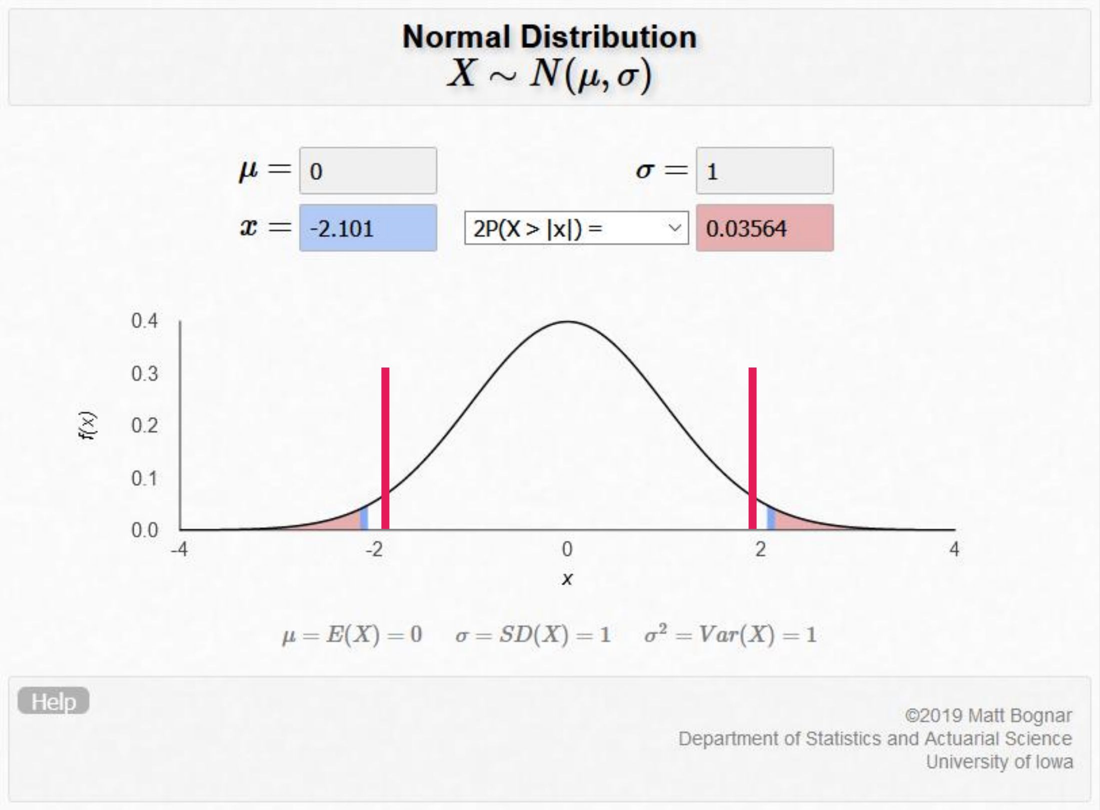
$$p - value\ (\mathrm{two} - \mathrm{tailed}) = 0.0356919 \qquad \alpha = 0.05$$

ΑΠΟΡΡΙΠΤΟΥΜΕ ΤΗ H_0: Τα αποτελέσματα προσφέρουν επαρκή στοιχεία για να συμπεράνουμε ότι 15 μέρες μετά τον εμβολιασμό υπάρχει αύξηση των ανοσοσφαιρινών M του ορού, στατιστικά σημαντική

Κανονική κατανομή, οπτικά



$$p - value \text{ (lower} - \text{tailed)} = 0.01784595$$



$$p - value \text{ (two} - \text{tailed)} = 0.0356919$$

$$z = \pm 1.96, 95\% \text{ interval}$$

https://homepage.divms.uiowa.edu/~mbognar/applets/normal.html

Κατανάλωση ενέργειας μεταξύ μίας ομάδας λεπτών και μίας ομάδας παχύσαρκων γυναικών Mann-Whitney Wilcoxon test

(Independent sample t-test)

Ιστορικό

- Για τη σύγκριση δύο δειγμάτων κάνοντας χρήση ενός παραμετρικού τεστ θα κάναμε χρήση του t-test για ανεξάρτητα δείγματα
- Στην περίπτωση αυτή θα ελέγχαμε αν οι μέσες τιμές των δύο δειγμάτων θα ήταν ίσες: $\bar{x}_1 - \bar{x}_1 = 0$
- Η υπόθεση που κάνουμε είναι πως ο κάθε πληθυσμός είναι κανονικά κατανεμημένος
- Το Wilcoxon Rank Sum / Mann-Whitney test μας επιτρέπει να συγκρίνουμε πληθυσμούς των οποίων οι κατανομές δεν είναι κανονικά κατανεμημένες
- Γι’ αυτό κάνουμε χρήση των ΤΑΞΕΩΝ των παρατηρήσεων

Προϋποθέσεις

Οι μόνες προϋποθέσεις για την εκτέλεση της δοκιμής Mann-Whitney είναι

- οι δύο ομάδες να είναι ανεξάρτητες και
- η εξαρτημένη μεταβλητή να είναι διατάξιμη ή συνεχής

Σημείωση: Ωστόσο, για να αναφερθεί η διαφορά μεταξύ των ομάδων με βάση τις διάμεσους, το σχήμα των κατανομών της εξαρτημένης μεταβλητής σε κάθε ομάδα πρέπει να είναι παρόμοιο (κυρτότητα)

Κατανάλωση ενέργειας μεταξύ μίας ομάδας λεπτών και μίας ομάδας παχύσαρκων γυναικών

Η κατανάλωση ενέργειας μέσα σε 24 ώρες μίας ομάδας λεπτών γυναικών και μίας ομάδας παχύσαρκων γυναικών δίνεται παρακάτω.

Υπάρχει διαφορά στην κατανάλωση ενέργειας μεταξύ των δύο ομάδων;

$$median_{λεπτών} = 7.9$$

$$median_{παχύσαρκων} = 9.69$$

Α/Α	Λεπτές	Παχύσαρκες
1	6.13	8.79
2	7.05	9.19
3	7.48	9.21
4	7.48	9.68
5	7.53	9.69
6	7.58	9.97
7	7.90	11.51
8	8.08	11.85
9	8.09	12.79
10	8.11	
11	8.40	
12	10.15	
13	10.88	

Βήμα 1: Υπολογισμός των τάξεων

Βρίσκουμε τις τάξεις (ranks) όλων των δεδομένων θεωρώντας ότι είναι ΕΝΑ δείγμα.
Κατόπιν βρίσκουμε το άθροισμα T των τάξεων της πιο μικρής ομάδας (παχύσαρκες).

Λεπτές		Παχύσαρκες	
Rank	Ενέργεια	Ενέργεια	Rank
1	6.13		
2	7.05		
3.5	7.48		
3.5	7.48		
5	7.53		
6	7.58		
7	7.90		
8	8.08		
9	8.09		
10	8.11		
11	8.40		

Λεπτές		Παχύσαρκες	
Rank	Ενέργεια	Ενέργεια	Rank
		8.79	12
		9.19	13
		9.21	14
		9.68	15
		9.69	16
		9.97	17
18	10.15		
19	10.88		
		11.51	20
		11.85	21
		12.79	22
$\boldsymbol{R = 103}$			$\boldsymbol{R = 150}$

A/A	Λεπτές	Παχύσαρκες
1	6.13	8.79
2	7.05	9.19
3	7.48	9.21
4	7.48	9.68
5	7.53	9.69
6	7.58	9.97
7	7.90	11.51
8	8.08	11.85
9	8.09	12.79
10	8.11	
11	8.40	
12	10.15	
13	10.88	

Βήμα 2: Υπολογισμός των τιμών U

$$U_{\lambda\varepsilon\pi\tau\varepsilon\varsigma} = R_{\lambda\varepsilon\pi\tau\varepsilon\varsigma} - \frac{n_{\lambda\varepsilon\pi\tau\varepsilon\varsigma}(n_{\lambda\varepsilon\pi\tau\varepsilon\varsigma} + 1)}{2}$$

$$U_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma} = R_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma} - \frac{n_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma}(n_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma} + 1)}{2}$$

$n_{\lambda\varepsilon\pi\tau\varepsilon\varsigma}$ = μέγεθος δείγματος ομάδας λεπτών γυναικών
$n_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma}$ = μέγεθος δείγματος ομάδας παχύσαρκων γυναικών

Βήμα 2: Υπολογισμός τιμών U

$$U_{\lambda\varepsilon\pi\tau\varepsilon\varsigma} = 103 - \frac{13(13+1)}{2} = 103 - 91 = 12$$

$$U_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma} = 150 - \frac{9(9+1)}{2} = 150 - 45 = 105$$

$n_{\lambda\varepsilon\pi\tau\varepsilon\varsigma}$ = μέγεθος δείγματος ομάδας λεπτών γυναικών
$n_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma}$ = μέγεθος δείγματος ομάδας παχύσαρκων γυναικών

Βήμα 3: Προσδιορισμός μικρότερης τιμής U

$$U = min(12, 105) = 12$$

Η μικρότερη τιμή U χρησιμοποιείται για την ερμηνεία, οπότε $U = 12$

Βήμα 4: Προσδιορισμός μικρότερης τιμής U

$n_1 \Rightarrow$ $n_2 \Downarrow$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3																			
4			0																
5		0	1	2															
6		1	2	3	5														
7		1	3	5	6	8													
8	0	2	3	6	8	10	13												
9	0	2	4	7	9	12	15	17											
10	0	3	5	8	11	14	17	20	23										
11	0	3	6	9	13	16	19	23	26	30									
12	1	3	7	11	14	18	22	26	29	33	37								
13	1	4	8	12	16	20	24	28	33	37	41	45							
14	1	5	9	13	17	22	26	31	36	40	45	50	54						
15	1	5	10	14	19	24	29	34	38	44	49	54	59	64					
16	1	6	11	16	21	26	31	36	42	48	53	59	64	69	74				
17	2	6	12	17	22	28	34	39	45	51	57	63	69	75	80	86			
18	2	6	13	18	24	30	36	42	49	55	61	67	73	80	86	93	99		
19	2	7	13	19	26	31	39	45	52	58	65	72	78	85	92	99	106	112	
20	2	8	14	20	27	34	41	48	55	62	68	76	83	90	98	104	112	119	127

Για έναν έλεγχο Mann-Whitney U με $n_{\lambda\varepsilon\pi\tau\varepsilon\varsigma} = 13, n_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma} = 9$, ελέγχουμε τον πίνακα κατανομής U ή χρησιμοποιούμε μία εφαρμογή για τον υπολογισμό του κρίσιμου σημείου

Εδώ, το κρίσιμο σημείο είναι **28** για έναν έλεγχο διπλής κατεύθυνσης

https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470776124.app1 (page 8 of 14)

Βήμα 5: Σύγκριση U με το κρίσιμο σημείο

Καθώς το $U = 12$ είναι μικρότερο από το 28, **απορρίπτουμε την μηδενική υπόθεση**. Αυτό υποδεικνύει πως υπάρχει στατιστικά σημαντική διαφορά μεταξύ των λεπτών και παχύσαρκων γυναικών

Wilcoxon W

- Η τιμή **Wilcoxon W** είναι απλά η μικρότερη τιμή από τα αθροίσματα των τάξεων, στην περίπτωση αυτή 103, αλλά
- το **SPSS** χρησιμοποιεί μία προσέγγιση της κανονικής κατανομής για να υπολογίσει την τιμή Z και το p-value (Asymptotic Sig).

Test Statistics[a]

	energy
Mann-Whitney U	12.000
Wilcoxon W	103.000
Z	-3.106
Asymp. Sig. (2-tailed)	.002
Exact Sig. [2*(1-tailed Sig.)]	.001[b]

a. Grouping Variable: group

b. Not corrected for ties.

Ερμηνεία

Η δοκιμή **Mann-Whitney U** έδειξε ότι υπάρχει **στατιστικά σημαντική διαφορά** στην κατανάλωσης ενέργειας μεταξύ της ομάδας λεπτών και της ομάδας των παχύσαρκων γυναικών **($U = 12$, $p < 0.05$)**. Η διάμεση κατανάλωση ενέργειας ήταν 7.9 για την ομάδα των λεπτών γυναικών σε σύγκριση με 9.69 για την ομάδα των παχύσαρκων γυναικών, υποδηλώνοντας ότι η παχύσαρκη ομάδα καταναλώνει περισσότερη ενέργεια

Δειγματική κατανομή

Μέση τιμή: $\mu_U = \frac{n_{λεπτες} n_{παχύσαρκες}}{2}$

Τυπική απόκλιση: $\sigma_U = \sqrt{\frac{n_{λεπτες} n_{παχύσαρκες}(n_{λεπτες} + n_{παχύσαρκες} + 1)}{12}}$

$n_{λεπτες}$ = μέγεθος δείγματος ομάδας λεπτών γυναικών
$n_{παχύσαρκες}$ = μέγεθος δείγματος ομάδας παχύσαρκων γυναικών

Προσέγγιση κανονικής κατανομής για $n \geq 10$

Δειγματική κατανομή

$$\text{Μέση τιμή: } \mu_U = \frac{n_{\lambda\varepsilon\pi\tau\varepsilon\varsigma} n_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma}}{2} = \frac{13 \cdot 9}{2} = 58.5$$

$$\text{Τυπική απόκλιση: } \sigma_U = \sqrt{\frac{13 \cdot 9(13 + 9 + 1)}{12}} = 14.97$$

$n_{\lambda\varepsilon\pi\tau\varepsilon\varsigma}$ = μέγεθος δείγματος ομάδας λεπτών γυναικών
$n_{\pi\alpha\chi\acute{\upsilon}\sigma\alpha\rho\kappa\varepsilon\varsigma}$ = μέγεθος δείγματος ομάδας παχύσαρκων γυναικών

Προσέγγιση κανονικής κατανομής για $n \geq 10$

Στατιστική δοκιμή

$$z = \frac{U - \mu_U}{\sigma_U}$$

Στατιστική δοκιμή και αποτέλεσμα

$$z = \frac{U - \mu_U}{\sigma_U} = \frac{12 - 58.5}{14.97} = \frac{-46.5}{14.97} = -3.106$$

$$p - value = \mathrm{NORM.S.DIST}(-3.106, \mathrm{TRUE})$$

$$p - value\ (\mathrm{lower - tailed}) = 0.0009475$$

Διπλασιάζουμε την πιθανότητα στο άκρο για να γίνει διπλής ουράς (two-tailed)

$$p - value\ (\mathrm{two - tailed}) = 0.01895 \qquad \alpha = 0.05$$

ΑΠΟΡΡΙΠΤΟΥΜΕ ΤΗ H_0: Τα αποτελέσματα προσφέρουν επαρκή στοιχεία για να συμπεράνουμε ότι υπάρχει διαφορά στην κατανάλωση ενέργειας μεταξύ των λεπτών και παχύσαρκων γυναικών

Μείωση της έντασης της κεφαλαλγίας στα παιδιά: Σύγκριση μεθόδων θεραπείας Kruskal-Wallis δοκιμή

(one-way ANOVA)

Ιστορικό

- Η σύγκριση περισσότερων από δύο ανεξάρτητων δειγμάτων με τη χρήση μιας **παραμετρικής μεθόδου** θα μας οδηγούσε στη στατιστική διαδικασία **one-way ANOVA**
- Στην περίπτωση αυτή θα ελέγχαμε αν οι μέσες τιμές των δειγμάτων είναι ίσες: $\boldsymbol{\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = \cdots \bar{x}_n}$
- Ωστόσο, προϋπόθεση για την εκτέλεση της one-way ANOVA είναι ότι ο κάθε πληθυσμός πρέπει είναι κανονικά κατανεμημένος
- Η δοκιμή **Kruskal-Wallis** μας επιτρέπει να συγκρίνουμε περισσότερους από δύο πληθυσμούς που παραβιάζουν την υπόθεση της κανονικότητας
- Χρησιμοποιούμε το άθροισμα των τάξεων των παρατηρήσεων

Δειγματική κατανομή

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} n_i \left(\frac{R_i}{n_i} - \frac{(n+1)}{2} \right)^2$$

Iterate through each sample/group

n_i = μέγεθος δείγματος στο δείγμα/ομάδα i
n = συνολικό μέγεθος δείγματος
R_i = άθροισμα τάξεων για δείγμα/ομάδα i

Στατιστική δοκιμή και p-value

Στατιστική δοκιμή: $H > {\chi^2}_{\alpha,k-1}$ χ^2 δοκιμή

Είναι το άθροισμα των τάξεων για κάθε ομάδα το ίδιο;

p-value: $P(H < \chi^2)$

Οι κρίσιμες τιμές της χ^2 δοκιμής μπορούν να βρεθούν στην διεύθυνση http://www.itl.nist.gov/div898/handbook/eda/section3/eda3674.htm

Μείωση της έντασης της κεφαλαλγίας στα παιδιά: Σύγκριση μεθόδων θεραπείας

Δεκαοκτώ παιδιά με ημικρανίες υποβλήθηκαν σε τρεις διαφορετικές θεραπείες. Η πρώτη ομάδα έλαβε **θεραπεία χαλάρωσης και βιοανάδρασης** (Θεραπεία Α), η δεύτερη ομάδα έλαβε μόνο **θεραπεία χαλάρωσης** (Θεραπεία Β), και η τρίτη ομάδα **δεν έλαβε καμία θεραπεία** (Θεραπεία Γ). Η μεταβλητή απόκρισης ήταν η **"μείωση της δραστηριότητας κεφαλαλγίας"**, όπου μια αρνητική τιμή υποδηλώνει αύξηση της δραστηριότητας κεφαλαλγίας και η τιμή 100 υποδηλώνει την απουσία πονοκεφάλων. Τα δεδομένα είναι (Fenttress 1986).

Θεραπεία Α	Θεραπεία Β	Θεραπεία Γ
62	69	50
74	43	−120
86	100	100
74	94	−288
91	100	4
37	98	−76

https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1469-8749.1986.tb03847.x

Μείωση της έντασης της κεφαλαλγίας στα παιδιά: Σύγκριση μεθόδων θεραπείας

Οι 18 δοκιμασίες αναμιγνύονται και υπολογίζονται οι τάξεις (από μικρότερο προς μεγαλύτερο)

Θεραπεία	Ανταπόκριση	Τάξη
B	100	17
B	100	17
C	100	17
B	98	15
B	94	14
A	91	13
A	86	12
A	74	10.5
A	74	10.5
B	69	9
A	62	8
C	50	7
B	43	6
A	37	5
C	4	4
C	-76	3
C	-120	2
C	-288	1

Άθροισμα τάξεων

Οι 18 δοκιμασίες αναμιγνύονται και υπολογίζονται οι τάξεις (από μικρότερο προς μεγαλύτερο)

Θεραπεία	Ανταπόκριση	Τάξη
B	100	17
B	100	17
C	100	17
B	98	15
B	94	14
A	91	13
A	86	12
A	74	10.5
A	74	10.5
B	69	9
A	62	8
C	50	7
B	43	6
A	37	5
C	4	4
C	-76	3
C	-120	2
C	-288	1

$\sum(\mathrm{A}) = 59$

$\sum(\mathrm{B}) = 78$

$\sum(\Gamma) = 34$

$\sum(\text{Τάξεις}) = 171$

Στατιστική δοκιμή και p-value

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} n_i \left(\frac{R_i}{n_i} - \frac{(n+1)}{2} \right)^2$$

$$H = \frac{12}{18(18+1)} \left[6\left(\frac{59}{6} - \frac{(18+1)}{2} \right)^2 + 6\left(\frac{78}{6} - \frac{(18+1)}{2} \right)^2 + 6\left(\frac{34}{6} - \frac{(18+1)}{2} \right)^2 \right]$$

$$H = 0.0351(0.6666 + 73.5 + 162.3333)$$

$$H = 0.0351(162.3333) = 5.6959$$

n_i = μέγεθος δείγματος στο δείγμα/ομάδα i
n = συνολικό μέγεθος δείγματος
R_i = άθροισμα τάξεων για δείγμα/ομάδα i

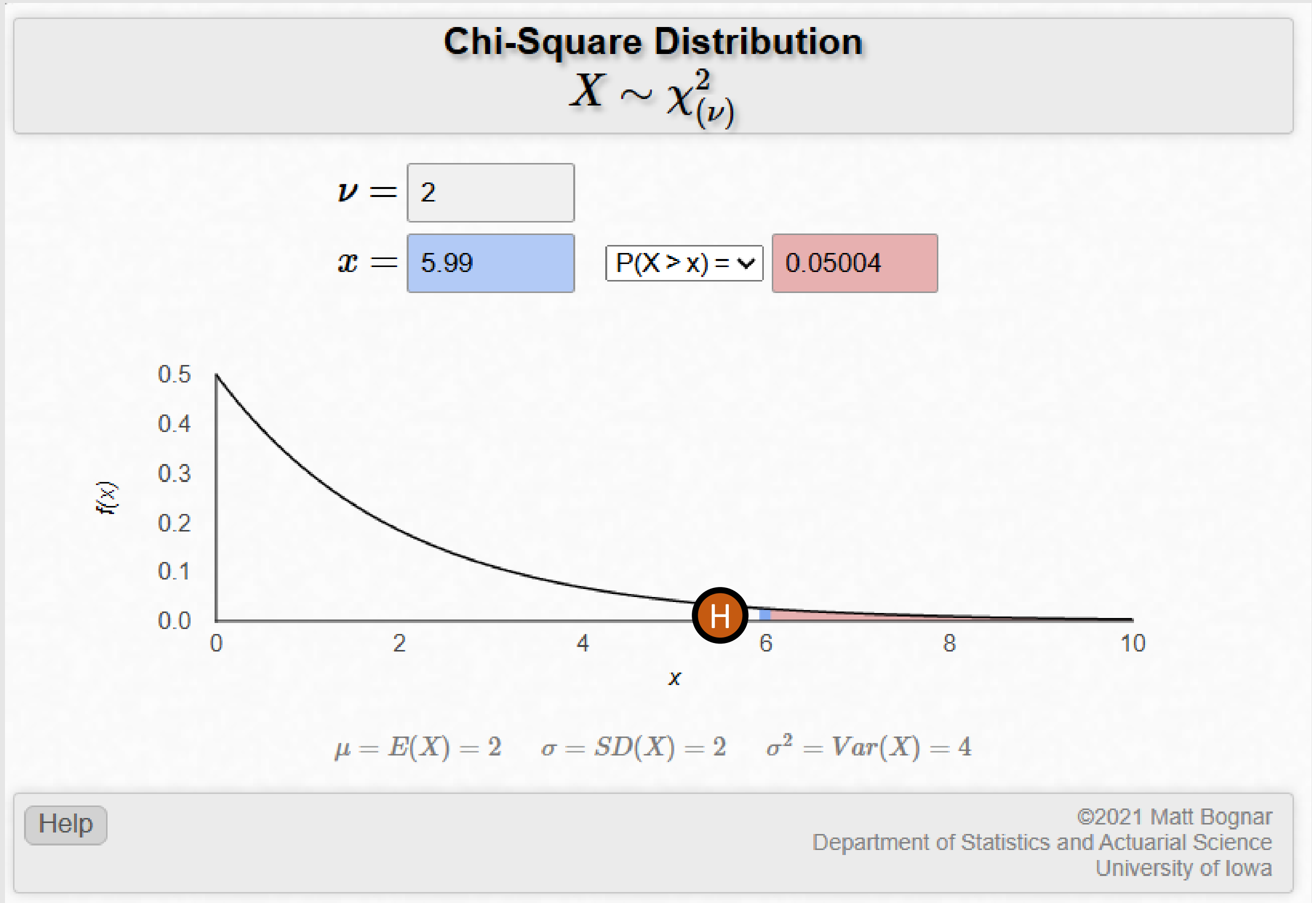
$\sum(\mathrm{A}) = 59$ $\sum(\mathrm{B}) = 78$ $\sum(\Gamma) = 34$

Στατιστική δοκιμή και p-value

Στατιστική δοκιμή:

$5.6959 < 5.99$

p-value: 0.06



https://homepage.divms.uiowa.edu/~mbognar/applets/chisq.html

http://www.itl.nist.gov/div898/handbook/eda/section3/eda3674.htm

Ερμηνεία

Σε επίπεδο σημαντικότητας 5%, προκύπτει ότι οι τρεις μέθοδοι δεν διαφέρουν στατιστικά σημαντικά, καθώς η τιμή H = 5.69 είναι μικρότερη από την κρίσιμη τιμή της χ^2 κατανομής με 2 βαθμούς ελευθερίας, που είναι 5.99 ($P \geq 0.05$). Η τιμή p που υπολογίζεται με χρήση στατιστικού πακέτου είναι 0.06

Εάν παρατηρούνταν ένα στατιστικά σημαντικό αποτέλεσμα, θα πραγματοποιούσαμε πολλαπλές συγκρίσεις χρησιμοποιώντας τη δοκιμή Mann-Whitney με μια προσαρμογή, όπως η διόρθωση Bonferroni