## Έλεγχοι για ένα δείγμα (One-sample problem)

* Έχουμε παρατηρήσεις  που προέρχονται από ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με αθροιστική συνάρτηση κατανομής .
* Θέλουμε να κάνουμε συμπερασματολογία για κάποια συγκεκριμένη πλευρά της , η οποία συνήθως δίνεται από κάποια παράμετρο που περιγράφει την .
* Συχνά ενδιαφερόμαστε για τη ***«θέση» (location) της κατανομής***, η οποία μπορεί να περιγράφεται με παραμέτρους όπως ο *μέσος*ή η *διάμεσος* . Τέτοιες παράμετροι χαρακτηρίζονται ως ***παράμετροι κεντρικής τάσης (location parameters)***.
* Οι δύο αυτές παράμετροι συμπίπτουν για συμμετρικές κατανομές. Η μη παραμετρική στατιστική εστιάζει περισσότερο στη διάμεσο, η οποία μπορεί να εκτιμηθεί με πιο ανθεκτικά (robust) στατιστικά.
* Αν λοιπόν μας ενδιαφέρει να κάνουμε ελέγχους που αφορούν τη "θέση" της κατανομής ***εξετάζουμε υποθέσεις της μορφής***

 έναντι της εναλλακτικής , ή έναντι  ή , για σταθερό  και όπου  παράμετροι που αφορούν τη θέση της κατανομής (π.χ. ο μέσος ή η διάμεσος)

* Οι έλεγχοι αυτοί αφορούν και τη περίπτωση όπου παρατηρώ ζεύγη για  (π.χ. τιμές *πριν και μετά μια θεραπεία*) άτομα και ενδιαφέρομαι να ελέγξω ***αν η θεραπεία είχε αποτέλεσμα***, δηλαδή αν ή (ισοδύναμα). Παίρνοντας ως  το πρόβλημα ανάγεται στα προηγούμενα: έλεγχος για τη θέση της κατανομής των .

### Ο κλασσικός παραμετρικός έλεγχος: t-test

* Στηρίζεται στο στατιστικό  , όπου  ο δειγματικός μέσος και  η ρίζα της δειγματικής διακύμανσης
* και στην μηδενική ότι υπό  η κατανομή της είναι η .
* Απορρίπτει την  σε επίπεδο 
	+ έναντι της εναλλακτικής  όταν , όπου  το -ποσοστιαίο σημείο της . Άρα λάθος τύπου Ι: 
	+ έναντι της εναλλακτικής  όταν , όπου  το -ποσοστιαίο σημείο της . Άρα λάθος τύπου Ι: 
	+ έναντι της εναλλακτικής  όταν , όπου  το -ποσοστιαίο σημείο της . Άρα λάθος τύπου Ι: 
* Οι υπολογισμοί του λάθους τύπου Ι ισχύουν ***με ακρίβεια αν τα  είναι κανονικά***, ενώ ισχύει προσεγγιστικά αν τα  είναι τέτοια που να ισχύει το ΚΟΘ. Αν η κατανομή των  είναι τέτοια που το ΚΟΘ αργεί να "τραβήξει" τότε ***το t-test μπορεί να είναι πολύ προβληματικό για μικρά n***.

### Προσημειακός έλεγχος (sign test)

* Ίσως το πιο πρώιμο παράδειγμα μη-παραμετρικού ελέγχου (Laplace, τη δεκαετία του 1700).
* Ελέγχει αν  (όπου  η διάμεσος), έναντι π.χ. της εναλλακτικής , , ή . Ως κλασσική εφαρμογή θυμηθείτε το παράδειγμα όπου ελέγχουμε αν μια θεραπεία είχε αποτέλεσμα.
* Το στατιστικό που χρησιμοποιείται είναι το . Για  απορρίπτουμε την  για μεγάλες τιμές του .
* Οι κρίσιμες τιμές (και τα p-values) υπολογίζονται από τη κατανομή της  υπό την . Η  είναι διωνυμική με παραμέτρους  και : αν  τότε .
* Άρα  και . Αυτό σημαίνει ότι η κατανομή το στατιστικού μας υπό τη μηδενική ***είναι ακριβώς γνωστή για πεπερασμένο  χωρίς καμία περαιτέρω υπόθεση για την κατανομή των ***.
* Χάριν ευκολίας μπορεί να χρησιμοποιηθεί και η κανονική προσέγγιση στη διωνυμική με  και . Έτσι η  υπό . Αυτό ισχύει για μεγάλα ******, ***αλλά ομοιόμορφα για όλες τις κατανομές των ***.

#### Ιδιότητες προσημειακού ελέγχου

##### Παρένθεση: σύγκριση δύο ελέγχων με την ασυμπτωτική σχετική αποδοτικότητα (Asymptotic relative efficiency, ARE)

* Έστω ένας έλεγχος για την μηδενική  που στηρίζεται σε ένα στατιστικό : απορρίπτουμε π.χ. για .
* Έστω  τα λάθη τύπου Ι και η ισχύς του ελέγχου (= η πιθανότητα να απορρίψω υπό  = 1- λάθος τύπου ΙΙ), και  ένα σημείο της εναλλακτικής.

Συμβολίζουμε με  το μικρότερο **μέγεθος δείγματος** που χρειάζεται ώστε να έχουμε

, 

* Τότε για δύο τέτοιους ελέγχους που στηρίζονται στα  και  αντίστοιχα, μπορούν **να συγκριθούν στη βάση του λόγου**



(H  είναι προτιμότερη αν ).

* Η ποσότητα αυτή είναι δύσκολο να υπολογιστεί, εκτός και αν υποθέσουμε π.χ. ότι , όπου  ανήκει στο όριο της μηδενικής. Τότε το



ονομάζεται **ασυμπτωτική σχετική αποδοτικότητα (**Asymptotic relative efficiency, **ARE ) κατά Pitman.**

* **Θεώρημα.** Αν για  και  και για  έχουμε , με  και  συνεχή στο , τότε: 
* **Θεώρημα**. Αν η αθροιστική συνάρτηση κατανομής  των *** είναι συμμετρική γύρω από το 0 και συνεχής ενώ η πυκνότητά τους στο 0 είναι θετική , τότε η ARE κατά Pitman*** *του προσημειακού ελέγχου έναντι του t-ελέγχου δίνεται από ,* όπου 
* Οι τιμές που παίρνει κανείς για διάφορα  είναι στο παρακάτω πίνακα:

F: Normal Uniform Logistic DE Cauchy *t*3 *t*5

** 0.637 0.333 0.822 2.000 *∞* 1.620 0.961

Δηλαδή: εκεί που το t-τεστ είναι βέλτιστο, δηλαδή

* **στη κανονικότητα**, ο προσημειακός έλεγχος είναι 37% χειρότερος από το t-τεστ
* **στη ομοιόμορφη** είναι 67% χειρότερος από το t-τεστ, και αυτό είναι και το κάτω όριο: αποδεικνύεται, δηλαδή ποτέ δεν είναι χειρότερος περισσότερο από αυτό.
* **Σε κάποιες κατανομές (με βαριές ουρές ή με άπειρη διακύμανση**) το t-τεστ μπορεί να είναι πολύ (άπειρα) χειρότερο από τον προσημειακό.

#### Εκτιμήτρια της διαμέσου (θέσης) βασισμένη στο προσημειακό έλεγχο

* Αν  (π.χ.  η επίδραση της θεραπείας,  καμιά επίδραση) τότε το  είναι συμμετρικό γύρω από το μέσο του 
* Μία **εκτιμήτρια του ** θα μπορούσε να κατασκευαστεί ως εκείνη η ποσότητα , που αν την αφαιρέσουμε από τα , η τιμή του **** που θα προκύπτει για το δείγμα  είναι όσο το δυνατό πλησιέστερα στο .
* Τότε αποδεικνύεται ότι το ** πρέπει να είναι η δειγματική διάμεσος** των . Ακριβέστερα αν  το διατεταγμένο δείγμα, τότε

.

* Αποδεικνύεται ότι ασυμπτωτικά

.

Και άρα **(ασυμπτωτικά)** **1-α διαστήματα εμπιστοσύνης** μπορούν να κατασκευαστούν ως .

* Ωστόσο μπορεί να κατασκευάσει κανείς και δ.ε. πεπερασμένου δείγματος για τη διάμεσο με βάση το διατεταγμένο δείγμα: Έστω και  μέγιστος θετικός ακέραιος με  ( με  Τότε το  είναι ένα **1-α διάστημα εμπιστοσύνης για τη διάμεσο, που είναι ελεύθερο κατανομής.**

### Έλεγχος Wilcoxon (Signed rank test)

* Ο προσημειακός έλεγχος χρησιμοποιεί μόνο τα πρόσημα των  και «αδιαφορεί» για τις τιμές των . Ο έλεγχος Wilcoxon ενσωματώνει και μέρος αυτή της πληροφορίας: όχι τις τιμές των , αλλά τουλάχιστο τη σειρά (τάξη) των  μεταξύ τους.
* **Ορισμός**: σε ένα σύνολο τιμών  η τάξη ενός συγκεκριμένου  ορίζεται ως .
* **Ορισμός:** Μία αθροιστική συνάρτηση κατανομής  λέγεται συμμετρική γύρω από το θ εάν .
* **Ορισμός:** Έστω παρατηρούμε  από ** συμμετρική** γύρω από κάποιο θ και θέλουμε να ελέγξουμε ***τη μηδενική ότι η διάμεσος της  είναι το 0***. Πρώτα παίρνουμε τα απόλυτα  και βρίσκουμε τις τάξεις  αυτών των απολύτων μεταξύ τους. Τότε το στατιστικό του Wilocoxon ορίζεται ως το άθροισμα των τάξεων που αντιστοιχούν σε θετικές παρατηρήσεις.

****

* Εάν η εναλλακτική είναι π.χ. η διάμεσος να είναι θετική τότε --υπό την εναλλακτική-- θα τείνω να έχω πιο πολλές θετικές παρατηρήσεις με μεγαλύτερες (απόλυτες) τιμές. Θα απέρριπτα λοιπόν τη μηδενική για μεγάλες τιμές της .
* Εάν θέσουμε  τότε αποδεικνύεται ότι
	+ **.**
	+ Υπό τη μηδενική (διάμεσος = 0) οι  είναι ανεξάρτητες Bernoulli με .
* Αυτό σημαίνει άμεσα ότι η ** είναι ελεύθερη κατανομής** και επιτρέπει τον **ακριβή υπολογισμό της κατανομής της** υπό τη μηδενική.
* Μπορούμε όμως για ευκολία να προσεγγίσουμε την κατανομή της για μεγάλα :

**Θεώρημα**. Εάν  συνεχής και συμμετρική γύρω από το 0 (άρα: υπό τη μηδενική),

.

Άρα θα μπορούσαμε να απορρίπτουμε τη μηδενική σε όφελος της εναλλακτικής μιας θετικής διαμέσου, όταν



με λάθος τύπου Ι να τείνει στο  για μεγάλα .

#### Σύγκριση Wilcoxon και t-test: ARE

* **Θεώρημα.** Αν η αθροιστική συνάρτηση κατανομής  των *** είναι συμμετρική γύρω από το 0 και συνεχής, τότε η ARE κατά Pitman*** *του ελέγχου Wilcoxon έναντι του t-ελέγχου δίνεται από ,* όπου . Η τιμή αυτή είναι πάντα μεγαλύτερη του .
* Οι τιμές που παίρνει κανείς για διάφορα  είναι στο παρακάτω πίνακα:

F: Normal Uniform Logistic DE Cauchy *t*3 *t*5

** 0.955 1.000 1.097 1.500 *∞* 1.900 1.240

Δηλαδή: εκεί που το t-τεστ είναι βέλτιστο, δηλαδή

* **στη κανονικότητα**, ο έλεγχος Wilcoxon είναι 5% χειρότερος από το t-τεστ
* στη **ομοιόμορφη** είναι εξίσου καλός.
* Σε **κάποιες κατανομές (με βαριές ουρές ή με άπειρη διακύμανση)** το t-τεστ μπορεί να είναι πολύ (άπειρα) χειρότερο και σχεδόν παντα είναι χειρότερο από τον Wilcoxon.
* Σε καμία κατανομή ο Wilcoxon δεν είναι χειρότερος περισσότερο από 14%!

#### Εκτιμήτρια της διαμέσου (θέσης) βασισμένη στον έλεγχο Wilcoxon

* **Σημείωση**. Αποδεικνύεται ότι , όπου .
* Αν  (π.χ.  η επίδραση της θεραπείας,  καμιά επίδραση) τότε το  είναι συμμετρικό γύρω από το μέσο του 
* Μία **εκτιμήτρια του ** θα μπορούσε να κατασκευαστεί ως εκείνη η ποσότητα , που αν την αφαιρέσουμε από τα , η τιμή του **** που θα προκύπτει για το δείγμα  είναι όσο το δυνατό πλησιέστερα στο .
* Τότε αποδεικνύεται ότι να πάρουμε

, όπου .

* Αποδεικνύεται ότι ασυμπτωτικά αν η  είναι συμμετρική γύρω από το 0 και συνεχής*,* και αν  τότε

.

* Έστω  μέγιστος θετικός ακέραιος με  και , με  το διατεταγμένο δείγμα των . Τότε το  είναι ένα **1-α διάστημα εμπιστοσύνης για τη διάμεσο, που είναι ελεύθερο κατανομής** (μέθοδος του Tukey).
* Προσεγγιστικά μπορεί να πάρει κανείς

.