



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΠΕΡΙΓΡΑΦΙΚΗ και ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Εισήγηση 4Α: Έλεγχοι Υποθέσεων και Διαστήματα Εμπιστοσύνης

Διδάσκων: Δαφέρμος Βασίλειος
ΤΜΗΜΑ ΠΟΛΙΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
ΣΧΟΛΗΣ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται στην άδεια χρήσης **Creative Commons** και ειδικότερα ***Αναφορά – Μη εμπορική Χρήση – Όχι Παράγωγο Έργο 3.0 Ελλάδα*** (***Attribution – Non Commercial – Non-derivatives 3.0 Greece***)



[ή επιλογή ενός άλλου από τους έξι συνδυασμούς]

[και αντικατάσταση λογότυπου άδειας όπου αυτό έχει μπει (σελ. 1, σελ. 2 και τελευταία)]

- Εξαιρείται από την ως άνω άδεια υλικό που περιλαμβάνεται στις διαφάνειες του μαθήματος, και υπόκειται σε άλλου τύπου άδεια χρήσης. Η άδεια χρήσης στην οποία υπόκειται το υλικό αυτό αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Δευτέρα 28-4-14 α

ΠΕΡΙΓΡΑΦΙΚΗ ΚΑΙ ΕΠΑΓΩΓΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Εισηγητής: Βασίλης Δαφέρμος, Αναπληρωτής Καθηγητής

ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΚΑΙ ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ



Τι είναι η Μηδενική και τι η Εναλλακτική υπόθεση

Οι στατιστικές υποθέσεις οι οποίες εμπλέκονται σε ένα οποιοδήποτε στατιστικό έλεγχο είναι πάντα δύο:

- Η πρώτη από αυτές ονομάζεται ***Μηδενική Υπόθεση*** και τη συμβολίζουμε με ***H₀*** , ενώ η δεύτερη ονομάζεται
- ***Εναλλακτική υπόθεση*** και τη συμβολίζουμε με ***H₁***.
- Δεν υπάρχει γενικός ορισμός ούτε για τη μηδενική, ούτε για την εναλλακτική υπόθεση. Πως ακριβώς διατυπώνεται τόσο η μία, όσο και η άλλη, εξαρτάται από το στατιστικό έλεγχο ή αλλιώς από το στατιστικό κριτήριο που θα χρησιμοποιήσουμε.
- Και βέβαια είναι ορθό να πούμε ότι η μηδενική υπόθεση τίθεται πάντα για ένα σκοπό και μόνον: Να ελεγχθεί, να αμφισβητηθεί ή για να το πούμε αλλιώς, **να κριθεί**.
- Ποιος όμως θα την κρίνει;
- Με άλλα λόγια σε τίνος την κρίση θα τεθεί;

Η απάντηση είναι ότι θα τεθεί στην κρίση ενός συγκεκριμένου (κάθε φορά) **στατιστικού** κριτηρίου.

$\neq \mu_2$.

Η μηδενική υπόθεση, είναι μια **υπόθεση μηδενικής διαφοράς**, όπως λέει και το όνομά της. Και προφανώς, ή θα γίνει δεκτή ή θα απορριφθεί.

• Αν γίνει δεκτή, αυτό θα σημαίνει όχι κατ' ανάγκη ότι είναι αληθινή, αλλά ότι **δεν υπάρχουν αρκετά στοιχεία ή πληροφορίες για να την απορρίψουμε**.

• Αν απορριφτεί, αυτό θα σημαίνει ότι οι πληροφορίες, **η μαρτυρία που διαθέτουμε είναι ικανή να κλονίσει την ισχύ της**.

Παραδείγματα μηδενικών και εναλλακτικών υποθέσεων

Παράδειγμα 1.

Ας υποθέσουμε ότι έχουμε να συγκρίνουμε δύο μέσους όρους μ_1 και μ_2 δύο ανεξάρτητων πληθυσμών. Τότε είναι γνωστό ότι ένα κατάλληλο στατιστικό κριτήριο, αρμόδιο για να ελέγχει την ύπαρξη πιθανής διαφοράς $\mu_2 - \mu_1$ είναι το λεγόμενο t-test για δύο ανεξάρτητα δείγματα (two samples independent t-test). Στην περίπτωση αυτού του στατιστικού ελέγχου, η μηδενική και η εναλλακτική υπόθεση διατυπώνονται ως εξής:

H_0 : ΔΕΝ υπάρχει διαφορά ανάμεσα στους πληθυσμιακούς μέσους όρους μ_1 και μ_2 , δηλ. $\mu_1 = \mu_2$.

H_1 : Υπάρχει διαφορά ανάμεσα στους πληθυσμιακούς μέσους όρους μ_1 και μ_2 , δηλ. $\mu_1 \neq \mu_2$.

Παραδείγματα μηδενικών και εναλλακτικών υποθέσεων

Παράδειγμα2.

Ας υποθέσουμε τώρα ότι έχουμε να συγκρίνουμε αντί για δύο, τρεις πληθυσμιακούς μέσους όρους μ_1 και μ_2 και μ_3 με τη βοήθεια της μονοπαραγοντικής Ανάλυσης Διασποράς (One Way ANOVA). Τότε η μηδενική και εναλλακτική υπόθεση διατυπώνονται ως εξής:

H_0 : $\mu_1 = \mu_2 = \mu_3$

H_1 : Υπάρχει τουλάχιστον μία διαφορά ανάμεσα στους τρεις πληθυσμιακούς μέσους όρους μ_1 και μ_2 και μ_3

Παράδειγμα3.

Ας υποθέσουμε ότι θέλουμε να ελέγξουμε την κανονικότητα ενός δείγματος, να ελέγξουμε δηλ. αν αυτό προέρχεται από ένα κανονικό πληθυσμό.

Τότε η μηδενική και εναλλακτική υπόθεση διατυπώνονται ως εξής:

H_0 : Η κατανομή του δείγματός μας **ΔΕΝ απέχει και πολύ** από την κανονική.

H_1 : Η κατανομή του δείγματός μας **απέχει πολύ** από την κανονική.

Για όλα, και πάντοτε, υπάρχει, τίθεται ένα όριο...

Για παράδειγμα, να δούμε μέχρι πού μ' αγαπάς....

Σ' όλα τα παραπάνω παραδείγματα, η υπόθεση που κρίνεται, είναι η μηδενική υπόθεση.

■ Αλλά για να κριθεί αυτή η υπόθεση χρειάζεται να λάβουμε και ένα **πιθανοθεωρητικό όριο**, να καθορίσουμε δηλ. **ένα επίπεδο στατιστικής σημαντικότητας**, πάνω από το οποίο θα δεχθούμε ως αληθινή τη μηδενική μας υπόθεση, ενώ κάτω από αυτό το όριο θα την απορρίψουμε ως ψευδή.

Ορισμός. Θα λέμε **επίπεδο στατιστικής σημαντικότητας** (significance level), και θα το συμβολίζουμε με **α** , την πιθανότητα, την τιμή της οποίας εμείς, ως ερευνητές ορίζουμε, και η οποία αντιστοιχεί στο ενδεχόμενο απόρριψης της μηδενικής μας υπόθεσης, ενώ αυτή (η μηδενική μας υπόθεση), είναι στην πραγματικότητα αληθής.

► Στο χώρο της Κοινωνικής Έρευνας, ως τιμές του επιπέδου στατιστικής σημαντικότητας επιλέγονται οι τιμές 5%, 3 %, 1% ή τέλος 1%. Η πιο συνηθισμένη τιμή επιπέδου σ.σ. είναι 5 %. Αυτή την τιμή θα λαμβάνουμε κι εμείς στις έρευνές μας, ως πιθανοθεωρητικό όριο, αν βέβαια δεν μας πουν κάτι διαφορετικό.

► Η **ελάχιστη τιμή** του επιπέδου στατιστικής σημαντικότητας, που αντιστοιχεί στην απόρριψη της μηδενικής υπόθεσης, είναι γνωστή στην αγγλική βιβλιογραφία, με τον όρο ***p-value***.

Τι κάνουμε στην πράξη;

► Στην πράξη, δηλ. στο πλαίσιο του SPSS, και του STATA, με τα οποία εμείς συνεχώς εργαζόμαστε, θα κάνουμε λόγο για δύο επίπεδα στατιστικής σημαντικότητας:

Το πρώτο αφορά το **παρατηρούμενο επίπεδο σ.σ.**, το οποίο πάντα θα μας το προσφέρει το SPSS (observed significance level), όταν εφαρμόζουμε ένα οποιοδήποτε στατιστικό κριτήριο.

Το δεύτερο αφορά το **θεωρητικό** επίπεδο σ.σ., αυτό δηλ. που εμείς σαν ερευνητές καθορίζουμε και με βάση το οποίο θα κριθεί η μηδενική μας υπόθεση.

Οι συλλογισμοί-πρακτικοί Κανόνες για να καταλήξουμε σε συμπέρασμα

■ *Αν το παρατηρούμενο επίπεδο σ.σ. είναι μεγαλύτερο του θεωρητικού
Ηο ισχύει.*

Αντίθετα...

=====

■ *Αν το παρατηρούμενο επίπεδο σ.σ. είναι μικτότερο του θεωρητικού
Ηο απορρίπτεται.*

Τελικά αυτό που πάντα μας ενδιαφέρει είναι η τύχη της Μηδενικής Υπόθεσης...

Μονόπλευρος και αμφίπλευρος έλεγχος υποθέσεων

Παράδειγμα 4.

Ας υποθέσουμε ότι μας έχει δοθεί δείγμα φοιτητών, οι οποίοι έχουν υποβληθεί σε κάποιο τεστ γνώσεων και μας τίθεται το ερώτημα: Ο μέσος όρος μ του πληθυσμού από τον οποίο προέρχεται το δείγμα αυτών των φοιτητών διαφέρει από την τιμή 80;

Η μηδενική και η εναλλακτική μας υπόθεση είναι τότε αντίστοιχα:

H_0 : Ο μέσος όρος μ του πληθυσμού από το οποίο προέρχεται το δείγμα των φοιτητών, **δεν διαφέρει** από την τιμή 80. Συμβολικά: $\mu=80$.

H_1 : Ο μέσος όρος του πληθυσμού από το οποίο προέρχεται το δείγμα των φοιτητών, **διαφέρει** από την τιμή 80. Συμβολικά: $\mu \neq 80$.

Το παραπάνω παράδειγμα 4, είναι ένα παράδειγμα ελέγχου υποθέσεων **διπλής κατεύθυνσης**. Διότι, όταν λέμε ότι $\mu = 80$, ίσως αυτό το μ να είναι μικρότερο του 80, οπότε έχουμε έλεγχο προς τα αριστερά ($\mu < 80$), ή αυτό το μ να είναι μεγαλύτερο του 80, οπότε έχουμε έλεγχο προς τα δεξιά ($\mu > 80$). Με άλλα λόγια, η λέξη 'διαφέρει', επειδή μας παραπέμπει σε έλεγχο και προς τα αριστερά (αρνητική κατεύθυνση) και προς τα δεξιά (θετική κατεύθυνση), αντανακλά αμφίπλευρο έλεγχο, δηλ. έλεγχο σε διπλή κατεύθυνση (two way testing hypothesis).

Τι κάνουμε στην πράξη όταν βρεθούμε μπροστά σε ένα ερευνητικό πρόβλημα;

► Με βάση όλα τα παραπάνω, και πάντα στο πλαίσιο του SPSS, όταν βρεθούμε μπροστά σε κάποιο στατιστικό πρόβλημα τα βήματα είναι τα εξής:

→ Αποφασίζουμε ποιο είναι το κατάλληλο στατιστικό κριτήριο που θα πρέπει να επιστρατεύσουμε για να λύσουμε το πρόβλημα.

→ Διατυπώνουμε με σαφήνεια τόσο τη μηδενική, όσο και την εναλλακτική μας υπόθεση.

→ Καθορίζουμε, μελετώντας τη βιβλιογραφία, την ιστορία του πράγματος, τις προηγούμενες σχετικές έρευνες, το θεωρητικό επίπεδο στατιστικής σημαντικότητας, με βάση το οποίο θα κρίνουμε τη μηδενική μας υπόθεση.

→ Τότε, αν το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, εκείνο δηλ. που μας δίνει το SPSS, είναι μικρότερο από το θεωρητικό τότε απορρίπτουμε τη μηδενική μας υπόθεση ως μη αληθή, θεωρούμε ότι αυτή δεν ισχύει και δεχόμαστε ως αληθή και επομένως ως αληθή την εναλλακτική μας υπόθεση.

Αντίθετα, αν το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας, εκείνο δηλ. που μας δίνει το SPSS, είναι μεγαλύτερο από το θεωρητικό, τότε δεχόμαστε τη μηδενική μας υπόθεση ως αληθή, δηλ. λέμε ότι αυτή ισχύει.

Σφάλματα στους ελέγχους υποθέσεων

◆ **Σφάλμα τύπου I.** Σε ένα έλεγχο υποθέσεων, αν απορρίψουμε ως εσφαλμένη τη μηδενική μας υπόθεση, ενώ αυτή είναι στην πραγματικότητα αληθής, τότε διαπράττουμε **σφάλμα τύπου I**. Η πιθανότητα να διαπράξουμε σφάλμα τύπου I, ονομάζεται συντελεστής α .

Συμβολικά: $P(I)=\alpha$.

◆ **Σφάλμα τύπου II.** Σε ένα έλεγχο υποθέσεων, αν δεχθούμε ως αληθή τη μηδενική μας υπόθεση, ενώ αυτή στην πραγματικότητα είναι εσφαλμένη, τότε διαπράττουμε **σφάλμα τύπου II**. Η πιθανότητα να διαπράξουμε σφάλμα τύπου II, ονομάζεται συντελεστής β .

Συμβολικά: $P(II)=\beta$.

Οι μεταβολές ή αλλιώς τα παιχνίδια του συντελεστή β

Οπωσδήποτε συντελεστής β γίνεται μεγαλύτερος:

1. Όσο κινούμαστε από την ποσοτική προς την ποιοτική ανάλυση, όσο δηλ. στις στατιστικές μας αναλύσεις κυριαρχούν τα ποιοτικά δεδομένα, σε βάρος των ποσοτικών.
2. Όσο ο συντελεστής α γίνεται μικρότερος.
3. Όσο το μέγεθος του δείγματος είναι μικρότερο.
4. Όσο μεγαλύτερη είναι η διασπορά του πληθυσμού.
5. Όσο πιο αναξιόπιστες είναι οι μετρήσεις.

Ας συζητήσουμε...

► Τα σφάλματα τύπου α και β , **δεν είναι σφάλματα σταθερού αθροίσματος**, παρά το γεγονός ότι αυξανόμενου του ενός μειώνεται το άλλο.

Με άλλα λόγια, **δεν ισχύει** $\alpha + \beta = \text{σταθ}$.

Για παράδειγμα στα μαθηματικά μπορεί να έχουμε τη σχέση $\alpha + \beta = \text{σταθ} = 10$
Και όσο μειώνουμε το α τόσο θα πρέπει να αυξάνουμε το β για να είναι σταθερό το άθροισμά τους στο 10.

Κάτι τέτοιο όμως **δεν ισχύει** στη Στατιστική, αν τα α και β είναι σφάλματα.

Ωστόσο εμείς στις Έρευνές μας πάντα θέλουμε :

μικρό σφάλμα τύπου I

και μικρό σφάλμα τύπου II

▶ Είναι σαφές ότι, αν θέλουμε να περιορίσουμε την πιθανότητα σφάλματος τύπου I, θα πρέπει να πάρουμε, όσο το δυνατόν μικρότερο επίπεδο στατιστικής σημαντικότητας α .

▶ Αλλά, όσο μικρότερο είναι το επίπεδο στατιστικής σημαντικότητας α , τόσο μεγαλύτερη είναι η πιθανότητα να διαπράξουμε σφάλμα τύπου II: Να δεχθούμε, εσφαλμένα, τη μηδενική μας υπόθεση, ως αληθή.

Τελικά, τι πρέπει να κάνει ο Κοινωνικός Ερευνητής αφού αυξανόμενου του ενός μειώνεται το άλλο και αντίστροφα;

▶ Να πούμε δηλ. εδώ στη στατιστική ανάλυση ότι ισχύει το γνωστό 'εμπρός γκρεμός και πίσω ρέμα'; Η απάντηση είναι πως ο Κοινωνικός ερευνητής θα πρέπει να καταφέρει το φαινομενικά ακατόρθωτο: την ταυτόχρονη μείωση και των δύο σφαλμάτων. Και κάτι τέτοιο είναι απολύτως εφικτό με την αύξηση του μεγέθους του δείγματος, στο μέτρο βέβαια του δυνατού.

Η διαφοροποιητική δύναμη ενός στατιστικού κριτηρίου (Power)

Ορισμός. Θα λέμε διαφοροποιητική δύναμη ενός στατιστικού κριτηρίου, και θα τη συμβολίζουμε με P , την πιθανότητα να μην διαπράξουμε σφάλμα τύπου II. Συμβολικά ισχύει:
 $P=1-\beta$

Με άλλα λόγια, θα λέμε διαφοροποιητική δύναμη ενός στατιστικού κριτηρίου, την πιθανότητα να κηρύξουμε αληθινή την εναλλακτική μας υπόθεση H_1 , όταν αυτή είναι πράγματι αληθινή.

Αυτό το τελευταίο πώς το λέμε με πολιτικούς όρους;

Αν θέλαμε να το πούμε αυτό με πολιτικούς όρους, θα λέγαμε πως είναι η πιθανότητα να κηρύξουμε ένοχο τον κατηγορούμενο, όταν πράγματι αυτός είναι ένοχος.

Η διαφοροποιητική δύναμη ενός στατιστικού κριτηρίου μας δείχνει πόσο βέβαιοι είμαστε ότι δεν διαπράξαμε σφάλμα τύπου II και επομένως είναι λογικό να είναι ίση με το αποτέλεσμα της διαφοράς $1-\beta$.

Γι αυτό ακριβώς ισχύει η σχέση που γράψαμε παραπάνω για την Power:

$$P=1-\beta$$

Από τη σχέση αυτή είναι προφανές ότι όσο μικρότερος είναι ο συντελεστής β , τόσο μεγαλύτερη είναι η P.

Έτσι,

Η διαφοροποιητική δύναμη ενός στατιστικού κριτηρίου, αφού είναι συνάρτηση του συντελεστή β , είναι λογικό να εξαρτάται όπως και εκείνος, από μια σειρά παραγόντων της ερευνητικής διαδικασίας.

Το περιεχόμενο των όρων

‘στατιστικώς σημαντική διαφορά’,

Και στατιστικώς ‘ασήμαντη διαφορά’

Ας υποθέσουμε και πάλι ότι έχουμε ένα τυχαίο δείγμα, μεγέθους n και μέσης τιμής \bar{x} , το οποίο θέλουμε να ξέρουμε αν προέρχεται από ένα πληθυσμό με μέση τιμή μ . Εάν σχηματίσουμε τη διαφορά $\bar{x} - \mu$, τότε αυτή η διαφορά μπορεί να είναι μια μικρή, συνήθης διαφορά ή να είναι μια μεγάλη, ασυνήθης διαφορά.

Γενικά, αν μια διαφορά μεταξύ ενός **στατιστικού δείκτη** ενός δείγματος και της αντίστοιχης **παραμέτρου του πληθυσμού** είναι τόσο μεγάλη, ώστε να μην είναι δυνατόν να αποδοθεί στη φυσική διακύμανση των τυχαίων δειγμάτων που λαμβάνουμε από αυτόν τον πληθυσμό, τότε λέμε ότι μπορεί να γίνεται λόγος για **στατιστικώς σημαντική διαφορά**. Η εμφάνιση στατιστικώς σημαντικής διαφοράς, οφείλεται στη δράση ενός συστηματικού εξωτερικού παράγοντα και όχι στην τυχαία δειγματοληψία.

Αν όμως η διαφορά ανάμεσα σε ένα στατιστικό δείκτη ενός δείγματος και στην αντίστοιχη παράμετρο του πληθυσμού είναι τόσο μικρή ώστε να τη θεωρούμε συνήθη διαφορά, τότε είναι δυνατόν να γίνει λόγος για **στατιστικώς ασήμαντη διαφορά**.

Τέλος Β' Εισήγησης

Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

