

# Συγχρονικές μελέτες πληθυσμών Population Surveys

Σχεδιασμός, δειγματοληψία και μέθοδοι ανάλυσης

# Τί είναι

- Συγχρονικές μελέτες που αφορούν μεγάλες πληθυσμιακές ομάδες
  - Παράδειγμα: Γενικός πληθυσμός χωρών ή εθνοτήτων/μειονοτήτων που κατοικούν σε μια χώρα

# Πού χρησιμεύουν

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, ο σχεδιασμός εθνικής στρατηγικής πρόληψης των χρόνιων νοσημάτων προϋποθέτει τον **προσδιορισμό του επιπολασμού των παραγόντων κινδύνου** σε κάθε χώρα μέσω **επιδημιολογικών μελετών σε τυχαία επιλεγμένα δείγματα γενικού πληθυσμού**

# Ποιος είναι ο στόχος τους;

Στοχεύουν στην εκτίμηση της συχνότητας των χαρακτηριστικών υγείας

- Νοσημάτων
- Παραγόντων κινδύνου
- Δομών/Υπηρεσιών υγείας

# Παράγοντας κινδύνου

Οποιοδήποτε χαρακτηριστικό το οποίο μπορεί να αυξήσει την πιθανότητα εμφάνισης κάποιας ασθένειας.



# Κύριοι παράγοντες κινδύνου

- Αρτηριακή πίεση
- Σακχαρώδης διαβήτης
- Κάπνισμα
- Αλκοόλ
- Υψηλή χοληστερίνη
- Παχυσαρκία
- Κοινωνική διαβάθμιση
- Στρες
- Ανεργία
- Διατροφή
- Άσκηση
- Περιβαλλοντική ρύπανση
- Κοινωνικο-οικονομικοί

**ΠΟΥ: Παράγοντες κινδύνου εξηγούν το  
60% των θανάτων**

# National Center for Health Statistics

CDC > NCHS

[National Health and Nutrition Examination Survey](#)



## National Health and Nutrition Examination Survey

About NHANES +

What's New +

Webinar

Questionnaires, Datasets, and Related Documentation +

Survey Participants +

Biospecimen Program +

New Content and Proposal Guidelines

### Survey Participants



If you were selected, learn more about participating

### Survey Data and Documentation



Access data, documentation & response rates

### Publications and Products



View health and nutrition reports & CDC Growth Charts

### Data Analysis Tutorials



Review step-by-step guidance on using NHANES data

**CDC** Analytic Guidance and 2024 NHANES Survey Content Propo...  
**NHANES Webinar** Copy link

[What's New](#)

# NHANES

The National Health Survey Act of 1956 provided for a continuing survey and special studies “ . . . **to secure accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States . . .** ” and to provide methods and processes to secure this statistical information (1).

Vital and Health Statistics, Series 1, Number 56

National Health and Nutrition Examination Survey: Plan and Operations, 1999–2010

Program and Collection Procedures



# Εθνική Μελέτη Νοσηρότητας και Παραγόντων Κινδύνου (ΕΜΕΝΟ)

National Survey of Morbidity and Risk Factors (EMENO): Protocol for a Health Examination Survey Representative of the Adult Greek

Population, Touloumi et al, JMR Res Protoc 2019



Γιώτα Τουλούμη

Καθηγήτρια

Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής

Ιατρική Σχολή Πανεπιστημίου Αθηνών



Ιατρική Σχολή Πανεπιστημίου Αθηνών

Εργ. Υγιεινής, Επιδημιολογίας & Ιατρικής Στατιστικής





## Καλώς ήρθατε στην Ε.ΜΕ.ΝΟ.

Η Ε.ΜΕ.ΝΟ είναι η πρώτη Εθνική Μελέτη Νοσηρότητας και Παραγόντων Κινδύνου, που συνδυάζει συλλογή δεδομένων υγείας και ιατρικών εξετάσεων σε τυχαία επιλεγμένο δείγμα και αφορά στην παρακολούθηση, καταγραφή και αξιολόγηση της γενικότερης κατάστασης της υγείας του πληθυσμού στην Ελλάδα.

#  
Η ΕΛΛΑΔΑ  
ΠΡΩΤΑΘΛΗΤΡΙΑ  
ΤΗΣ ΠΑΧΥΣΑΡΚΙΑΣ



# ΕΜΕΝΟ: Κύριοι Στόχοι

- Χαρτογράφηση της κατάστασης υγείας πληθυσμού
- Εκτίμηση επιπολασμού κύριων παραγόντων κινδύνου χρόνιων νοσημάτων (ανά γεωγραφική περιοχή, ηλικία, κοινωνικο-οικονομική κατάσταση)
- Εκτίμηση επιπολασμού αναπνευστικών και καρδιαγγειακών νοσημάτων – παράγοντες κινδύνου
- Χρήση υπηρεσιών υγείας, φαρμάκων και μέτρων πρόληψης – σχετιζόμενοι παράγοντες
- Βαθμός προσκόλλησης στη μεσογειακή διαίτα και σωματική άσκηση
- Αξιολόγηση ψυχολογικής ευεξίας
- Σχέση ατμοσφαιρικής ρύπανσης αναπνευστικής λειτουργίας
- Εκτίμηση μελλοντικού νοσολογικού φορτίου και αναγκών συστήματος υγείας

# ΣΚΕΠΤΙΚΟ ΕΠΙΔΗΜΙΟΛΟΓΙΚΩΝ ΜΕΛΕΤΩΝ

- Είναι αδύνατον να συμπεριλάβουμε ολόκληρο τον πληθυσμό αναφοράς!
- Χρησιμοποιούμε (κατάλληλο) δείγμα ώστε...
  - Να έχουμε μικρότερο κόστος
  - Να χρειαστούμε λιγότερο χρόνο
  - Να επικεντρωθούμε στην ορθότερη/πιο λεπτομερή συλλογή δεδομένων

# Στοιχεία Σχεδιασμού

- Ορισμός πληθυσμού αναφοράς (reference/target population)
- Ορισμός πλυθησμού από τον οποίο θα προέλθει το δείγμα (sampling frame)
- Επιλογή μεθόδου δειγματοληψίας (sampling method)
- Μέγεθος δείγματος (sample size)
- Ορισμος δειγματοληπτικής μονάδας (sampling unit)

# Άλλα θέματα

- Προϋπολογισμός
- Ethics-Confidentiality
- Επιλογή μεταβλητών-δεδομένων που θα συλλεχθούν
  - Ερωτηματολόγιο-Εξετάσεις υγείας (SOPs)
- Στατιστική μεθοδολογία (λαμβάνει υπόψιν και στοιχεία της δειγματοληψίας)
- Οργάνωση έρευνας πεδίου
- Συλλογή δεδομένων
  - - Βάση δεδομένων και καταγραφή
- Ανάλυση δεδομένων

# Θέματα Δεοντολογίας

- Έγκριση από την Επιτροπή Δεοντολογίας Ιατρικής Σχολής ΕΚΠΑ
- Αρχή προστασίας προσωπικών δεδομένων
- Ενυπόγραφη συναίνεση
  - Συμπλήρωση ερωτηματολογίων
  - Εξετάσεις αίματος
  - Αποθήκευση υπολειπόμενου δείγματος

# Σχεδιασμός (1)

- Καθορισμός του πληθυσμού αναφοράς (target population)
  - Είναι ο πληθυσμός στον οποίο θέλουμε να γενικεύονται, δηλαδή να ισχύουν, τα ευρήματά μας
- Είναι σημαντικό να καθορισθεί με ακρίβεια, γιατί καθορίζει την επιλογή του δείγματος
  - Παράδειγμα: Μόνιμος πληθυσμός χώρας- Άτομα ηλικίας >18 ετών, που διαμένουν στην χώρα >5έτη



# Σχεδιασμός (2)

- Δειγματοληπτικό πλαίσιο (sampling frame)
  - Είναι το τμήμα του πληθυσμού από το οποίο θα αντλήσουμε το δείγμα
- Αποφασίζεται βάσει:
  - Ευκολίας πρόσβασης, εγγύτητας, διαθεσιμότητας και άλλων πρακτικών λόγων
  - Καταλληλότητας- Θα πρέπει να οδηγεί σε γενικεύσιμα αποτελέσματα για την μελέτη (external validity)

# Σχεδιασμός (3)

- Καθορισμός δειγματοληπτικής μονάδας (sampling unit)
  - Είναι η μονάδα που επιλέγεται από το δειγματοληπτικό σχέδιο και μπορεί να είναι άτομο, νοικοκυριό, οικοδομικό τετράγωνο κλπ
- Υπόκειται σε κριτήρια επιλεξιμότητας/εξαίρεσης (inclusion/exclusion criteria)

# Σχεδιασμός (4)

- Επιλογή δείγματος με συγκεκριμένη μέθοδο ώστε να είναι κατάλληλο, δηλαδή τα αποτελέσματά που θα προκύψουν να μπορούν να γενικευθούν στον πληθυσμό αναφοράς (sampling method)
- Συγκεκριμένου μεγέθους ώστε να μπορούμε να απαντήσουμε στο ερευνητικό ερώτημα με ακρίβεια (sample size)

# Αντιπροσωπευτικό δείγμα

- Αντιπροσωπευτικό δείγμα ενός πληθυσμού αναφοράς είναι ένα δείγμα το οποίο αντανακλά επαρκώς τα χαρακτηριστικά του πληθυσμού αναφοράς
- Μεγαλύτερα δείγματα τείνουν να είναι αντιπροσωπευτικά (εφόσον δεν υπάρχει μεροληψία επιλογής)

# Τυχαίο δείγμα

Τυχαίο δείγμα είναι μια ομάδα ατόμων του πληθυσμού αναφοράς επιλεγμένη με τυχαίο (μή συστηματικό) τρόπο.

- Αυτό σημαίνει ότι κάθε άτομο του πληθυσμού έχει την ίδια πιθανότητα να επιλεγεί
- Τα τυχαία δείγματα δεν είναι απαραίτητα και αντιπροσωπευτικά

# Είδη Δειγματοληπτικών Μεθόδων

- Δειγματοληψία μή βασιζόμενη σε πιθανότητα (Non-probability sampling)
- Δειγματοληψία βάσει πιθανότητας (Probability sampling)

# Δειγματοληψία μή βασιζόμενη σε πιθανότητα

- Είδη δειγματοληψίας μή βασιζόμενης σε πιθανότητα:
  - Δείγμα ευκολίας (Convenience sampling): Επιλογή ατόμων/δειγματολείπτικων μονάδων που είναι εύκολα προσβάσιμα (Ρομά στην μελέτη EMENO)
  - Δειγματοληψία οδηγούμενη από τους συμμετέχοντες (Respondent driven sampling): επιλογή των πλέον προσβάσιμων ατόμων ενός πληθυσμού και των ατόμων του κοινωνικού κύκλου τους κ.ο.κ (XEN στην μελέτη Αριστοτέλης)
  - Επιλογή ατόμων από τον ερευνητή, συμμετοχή όσων θεωρεί ότι πρέπει να συμμετέχουν (purposeful sampling). Συνήθως επιλέγονται άτομα που θα συνεισφέρουν περισσότερες πληροφορίες.
  - Επιλογή ατόμων με συγκεκριμένα χαρακτηριστικά ή ιδότητες (Quota sampling), είναι μέθοδος αντίστοιχη της στρωματοποιημένης δειγματοληψίας, αλλά εδώ τα “στρώματα” είναι φτιαγμένα από τον ερευνητή
  -

# Δειγματοληψία μή βασιζόμενη σε πιθανότητα

- Τα είδη δειγματοληψίας μή βασιζόμενης σε πιθανότητα θεωρούνται γενικά μεροληπτικές μέθοδοι επιλογής δείγματος,
- Σε κάποιες περιπτώσεις δεν είναι δυνατή άλλου τύπου δειγματοληψία
- Τα quota και purposeful sampling χρησιμοποιούνται κυρίως στην έρευνα αγοράς και είναι μή κατάλληλα για επιδημιολογικές μελέτες
- Τα convenience sampling και RDS με κατάλληλες μεθόδους ανάλυσης μπορούν να οδηγήσουν σε αμερόληπτα αποτελέσματα





**ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**

**ΙΑΤΡΙΚΗ ΣΧΟΛΗ**

**ΕΡΓΑΣΤΗΡΙΟ ΥΓΙΕΙΝΗΣ ΕΠΙΔΗΜΙΟΛΟΓΙΑΣ ΚΑΙ ΙΑΤΡΙΚΗΣ ΣΤΑΤΙΣΤΙΚΗΣ**

**ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ**

**ΡΟΜΑ ΚΑΙ ΥΓΕΙΑ ΣΤΗΝ ΕΛΛΑΔΑ:**

**ΧΡΗΣΗ ΚΑΙ ΠΡΟΣΒΑΣΗ ΣΤΙΣ ΥΠΗΡΕΣΙΕΣ ΥΓΕΙΑΣ, ΧΡΟΝΙΑ ΚΑΙ ΑΙΜΑΤΟΓΕΝΩΣ  
ΜΕΤΑΔΙΔΟΜΕΝΑ ΝΟΣΗΜΑΤΑ (ΙΟΓΕΝΕΙΣ ΗΠΑΤΙΤΙΔΕΣ Β, C ΚΑΙ HIV ΛΟΙΜΩΞΗ)**

**Roma Health in Greece:**

**Use of and access to health services, chronic and blood transmitted diseases (viral Hepatitis B, C and  
HIV infection)**

**ΙΩΑΝΝΑ ΠΕΤΡΑΚΗ  
ΔΙΑΙΤΟΛΟΓΟΣ - ΔΙΑΤΡΟΦΟΛΟΓΟΣ**

# Δείγμα Ευκολίας

- Πληθυσμός Ρομά: Μή καταγεγραμμένος και δύσκολα προσβάσιμος
- Μή δυνατή η εφαρμογή probability sampling
- Χρησιμοποιήθηκε δείγμα ευκολίας
- Κριτήρια:

Ικανοποιητική γεωγραφική διασπορά,

Από διαφορετικές κατηγορίες οικισμού

να είναι αριθμητικά επαρκές σε κάθε οικισμό

# Δειγματοληψία

**1ο στάδιο:** Συλλογή πληροφοριών σχετικά με τον αριθμό των οικισμών και τον αριθμό των ατόμων που διαμένουν σε αυτούς, καθώς και τη γενική κατάσταση των οικισμών.

**2ο στάδιο** Επιλογή αντιπροσωπευτικών οικισμών από τέσσερις περιφέρειες της Ελλάδας

**3ο στάδιο** Δείγμα των 540 ατόμων μη αναλογικό του πληθυσμού των περιφερειών, ώστε υπάρχει ικανός αριθμός για ανά περιοχή συγκρίσεις, να επιλεγούν ισόποσα δείγματα (135 άτομα) σε κάθε μια από τις τέσσερις περιφέρειες

# Δειγματοληψία βάσει πιθανότητας (Probability sampling)

Κάθε δειγματοληπτική μονάδα έχει συγκεκριμένη μή μηδενική πιθανότητα να επιλεγεί στο δείγμα της μελέτης

- Είδη:
  - Απλή τυχαία δειγματοληψία (Simple random sampling)
  - Συστηματική τυχαία δειγματοληψία (Systematic random sampling)
  - Στρωματοποιημένη τυχαία δειγματοληψία (Stratified random sampling)
  - Κατά συστάδες δειγματοληψία (Cluster sampling)

# Δειγματοληψία βάσει πιθανότητας (Probability sampling)

- Επίσης μπορεί να γίνει:
  - Πολυσταδιακή τυχαία δειγματοληψία (Multi-stage random sampling)
  - Ή δειγματοληψία σε πολλαπλές φάσεις (Multi-phase sampling)

# Απλή Τυχαία Δειγματοληψία

- Κάθε άτομο/μονάδα του πληθυσμού έχει την ίδια πιθανότητα να επιλεγεί στο δείγμα
- Μπορεί να γίνει με χρήση λίστας τυχαίων αριθμών και αντιστοίχισης με τα άτομα του πληθυσμού ή με κάποιο αυτοματοποιημένο τρόπο (για παράδειγμα επιλογή από αρχείο registry με χρήση αλγορίθμου)
- Προϋποθέτει καταγραφή των ατόμων του πληθυσμού ή τουλάχιστον γνώση του αριθμού των ατόμων του πληθυσμού

# Συστηματική Τυχαία Δειγματοληψία

- Κατά τη **συστηματική** δειγματοληψία επιλέγεται τυχαία το πρώτο άτομο (για παράδειγμα σε λίστα του πληθυσμού) και κατόπιν κάθε  $n$ -οστό άτομο, μέχρι να συμπληρωθεί το δείγμα. Ιδιαίτερα χρήσιμη μέθοδος για μεγάλους πληθυσμούς.
- Παράδειγμα: Αν επιλέγουμε άτομα από τον τηλεφωνικό κατάλογο, επιλέγεται τυχαία ένας αριθμός από το 1 έως το 20 και έκτοτε επιλέγονται τα άτομα ανά 50 θέσεις (Δηλαδή ο 70κοστός, 120οστός κ.ο.κ).
- Δειγματοληπτικό κλάσμα (sampling fraction) είναι το πηλίκο του μεγέθους δείγματος προς το μέγεθος του πληθυσμού

# Στρωματοποιημένη Δειγματληψία

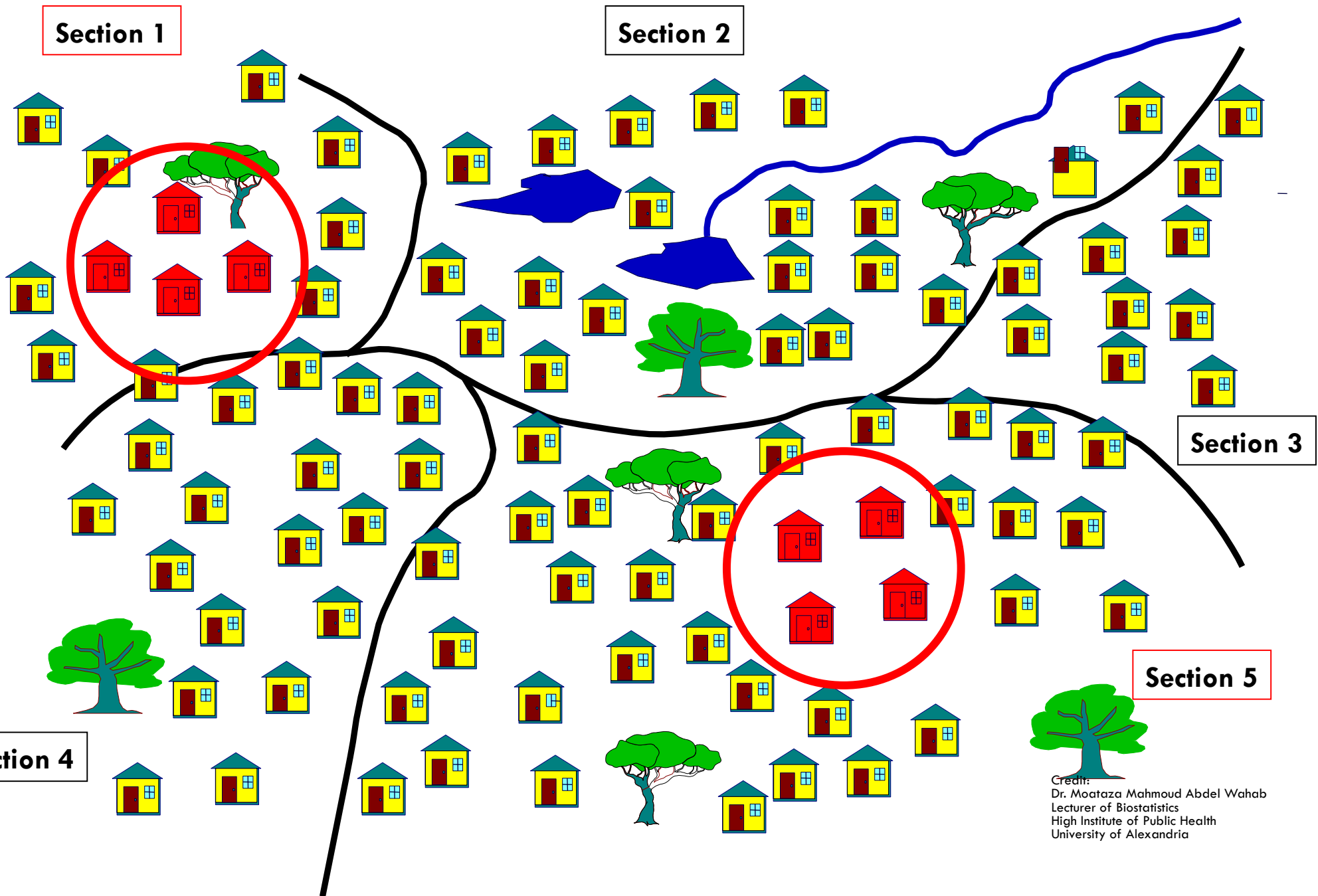
- Η δειγματοληψία πραγματοποιείται ανά στρώματα ή υποπληθυσμούς, είτε ανάλογα με το μέγεθός τους, είτε σε ίσους αριθμούς. Ο σκοπός είναι να εκπροσωπούνται επαρκώς αυτές οι υποομάδες του πληθυσμού στο δείγμα
- Για παράδειγμα, αν τα στρώματα αφορούν σε πόλεις της Ελλάδας, τότε η στρωματοποιημένη δειγματοληψία εξασφαλίζει ότι θα έχουμε άτομα από όλες τις πόλεις.



# Κατά συστάδες δειγματοληψία

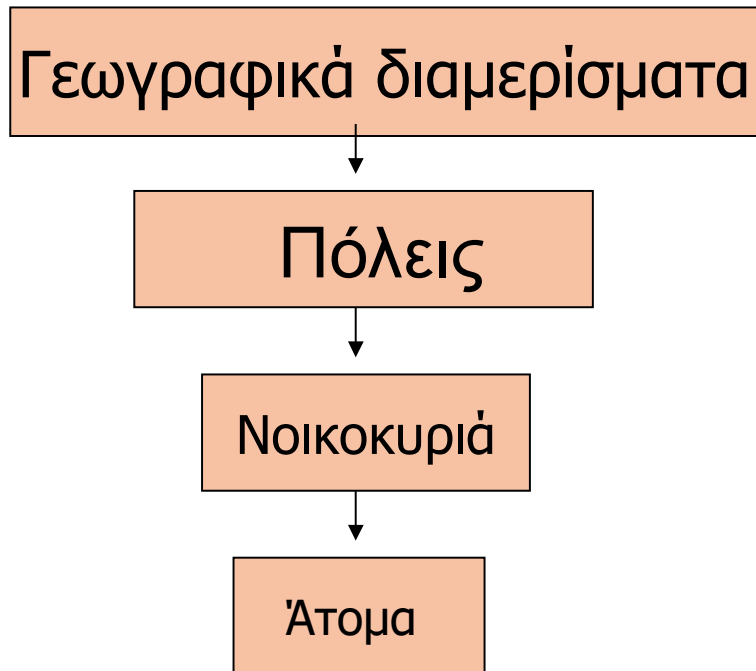
- Συστάδα είναι μια ομάδα ατόμων/δειγματοληπτικών μονάδων που βρίσκονται κοντά, για παράδειγμα στην ίδια περιοχή ή γειτονιά
- Επιλέγουμε τυχαίο δείγμα στρωμάτων και κατόπιν συμπεριλαμβάνουμε στο δείγμα όλα τα άτομα/μονάδες των στρωμάτων αυτών
- Για παράδειγμα, αν τα στρώματα ήταν σχολεία μιας πόλης, θα επιλέγαμε τυχαία κάποια σχολεία και θα συμπεριλάβαμε στη μελέτη όλους τους μαθητές των σχολείων αυτών

# Cluster Samples of Households

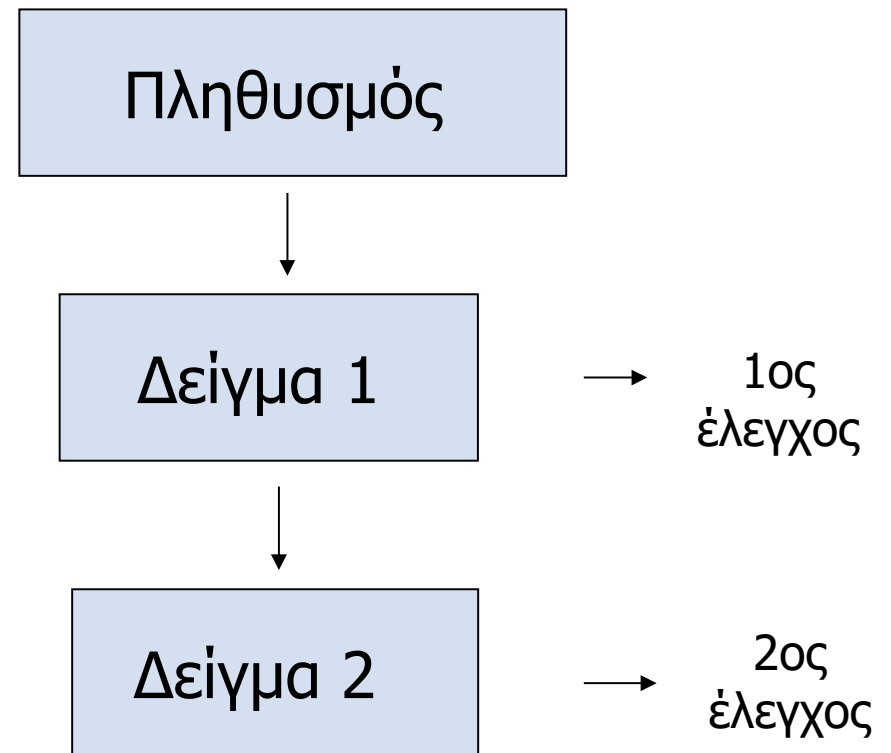


# Σύνθετες Δειγματοληπτικές μέθοδοι

Πολυσταδιακή  
δειγματοληψία



□ Δειγματοληψία σε  
πολλαπλές φάσεις



# Δειγματοληψία NHANES

National Center for Health Statistics

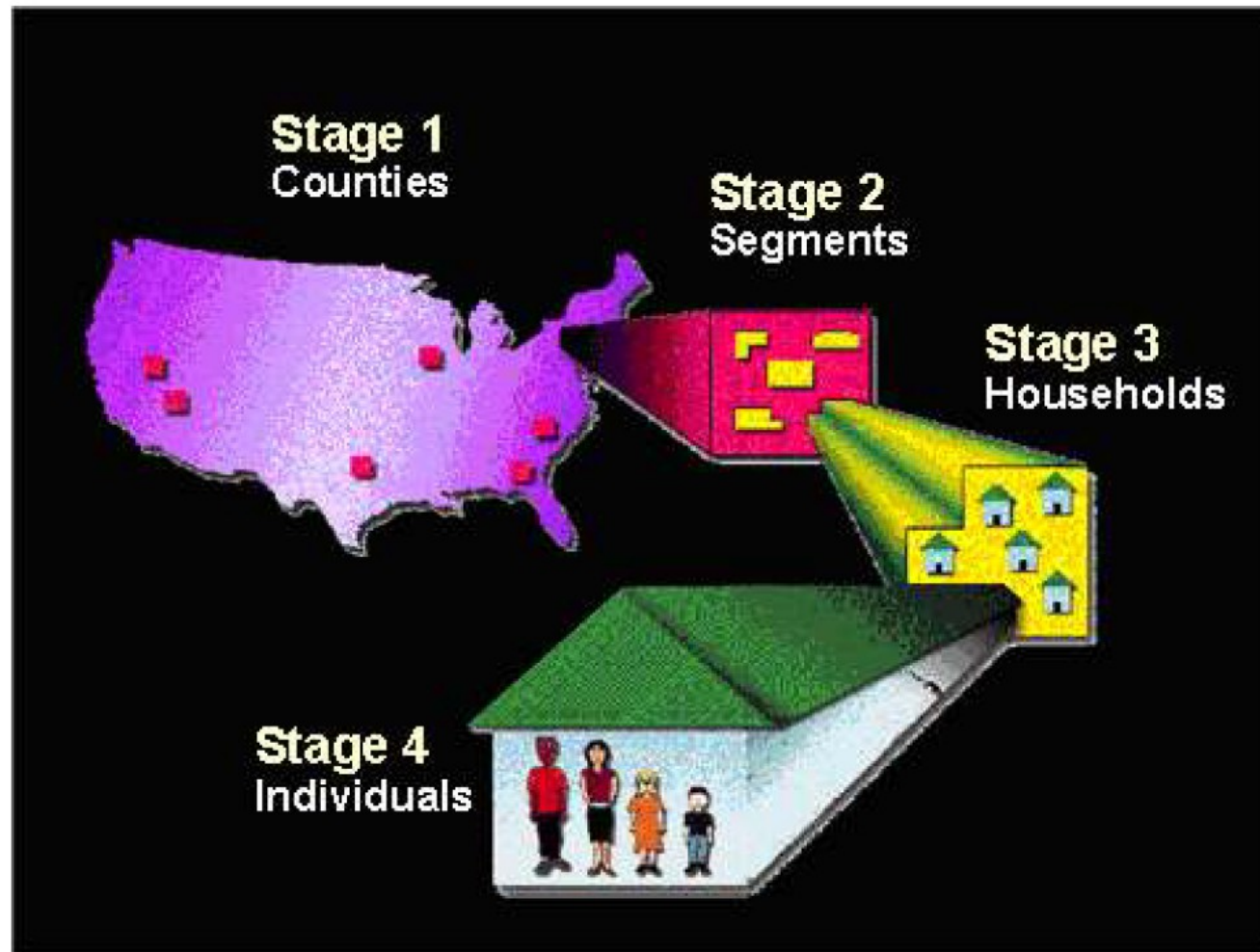


National Health and Nutrition Examination Survey

## Module 2: Sample Design

The NHANES samples are not simple random samples. Rather, a complex, multistage, probability sampling design is used to select participants representative of the civilian, non-institutionalized US population. Oversampling of certain population subgroups is also done to increase the reliability and precision of health status indicator estimates for these particular subgroups. Researchers need to take this into account in their analyses by appropriately specifying the sampling design parameters. This can be done using any statistical software that can analyze complex survey designs, such as SAS Survey procedures, SUDAAN, Stata, or R. This module provides an overview of the sample design parameters in NHANES.

# NHANES: Δειγματοληπτικό σχέδιο 4 σταδίων



*Four Stages of NHANES Sampling Procedure*

# Δειγματοληπτικό σχέδιο NHANES

## Στάδιο 1

Στάδιο 1: Επιλέχθηκαν οι πρωταρχικές δειγματοληπτικές μονάδες (Primary sampling units, PSUs). Αυτές ήταν πολιτείες ή σε ελάχιστες περιπτώσεις ομάδες πολιτειών.

Οι PSUs επιλέχθηκαν με πιθανότητα ανάλογη ενός μέτρου του μεγέθους τους (PPS). Αυτό σημαίνει ότι πολιτείες με μεγαλύτερο πληθυσμό ήταν πιθανότερο να επιλεγούν από άλλες με μικρότερο πλυθησμό.

Για τη NHANES χρησιμοποιήθηκε ένας σταθμισμένος μέσος του πληθυσμού, με βάρη που ήταν υπολογισμένα ώστε να δίνουν υψηλότερη πιθανότητα επιλογής σε πολιτείες με μεγαλύτερα ποσοστά ατόμων που βάσει σχεδιασμού είχαν επιλεγεί να υπερ-επιλεχθούν (oversampling)

# Δειγματοληπτικό σχέδιο NHANES

## Στάδια 2-4

2ο Στάδιο: Κάθε επιλεγμένη πολιτεία χωρίστηκε σε οικοδομικά τετράγωνα και εφαρμόστηκε δειγματοληψία με πιθανότητα ανάλογη μέτρου του μεγέθους τους

3ο Στάδιο: Σε κάθε ένα από τα επιλεγμένα οικοδομικά τετράγωνα επιλέχθηκαν νοικοκυριά από λίστα, επιτρέποντας και πάλι oversampling

4ο Στάδιο: Από κάθε επιλεγμένο νοικοκυριο, επιλέχθηκαν 2 άτομα με τυχαία δειγματοληψία

# Δειγματοληπτικά βάρη

- **Oversampling:** Η NHANES σχεδιάστηκε ώστε να παίρνει μεγαλύτερα (αναλογικά) δείγματα από κάποιες συγκεκριμένες υποομάδες που έχουν ιδιαίτερο ενδιαφέρον από πλευράς δημόσιας υγείας
- Στόχος του oversampling να έχουμε ένα αρκετά μεγάλο δείγμα, που θα επιτρέψει αξιόπιστες εκτιμήσεις σε αυτούς τους πληθυσμούς.



# Δειγματοληπτικά βάρη

- Τα δειγματοληπτικά βάρη αποδίδονται σε κάθε συμμετέχοντα
- Αντιστοιχούν στον αριθμό των ατόμων του πληθυσμού που αντιπροσωπεύει και αντανakλούν την άνιση πιθανότητα επιλογής
- Επιτρέπουν οι εκτιμήσεις που βασίζονται σε δείγματα μή ανάλογα του πραγματικού πληθυσμού να χρησιμοποιούνται και να δίνουν αμερόληπτες εκτιμήσεις για τον πληθυσμό.

# Δειγματοληπτικά βάρη

- Αν όλα τα άτομα ενός πληθυσμού 1000 ατόμων επιλέγονται με την ίδια πιθανότητα σε δείγμα 100 ατόμων, τότε κάθε ένας από τους συμμετέχοντες στο δείγμα αντιπροσωπεύει 10 άτομα του πληθυσμού με τα ίδια χαρακτηριστικά:
  - Αν ο πληθυσμός αποτελείται από γυναίκες 550 γυναίκες και 450 άνδρες και το δείγμα από 55 γυναίκες και 45 άνδρες, τότε κάθε γυναίκα εκπροσωπεί 10 γυναίκες και κάθε άνδρας 10 άνδρες
  - Σε αυτή την περίπτωση δεν υπάρχει λόγος να γίνει στάθμιση του δείγματος

# Δειγματοληπτικά βάρη

- Αν σε μία εθνική μελέτη μιας χώρα 5000000 κατοίκων θέλουμε να έχουμε ένα αρκετά μεγάλο δείγμα κατοίκων μιας απομακρυσμένης περιοχής με 1000 κατοίκους τότε θα επιλέξουμε συμμετέχοντες από αυτή την περιοχή με μεγαλύτερη πιθανότητα από αυτή θα αντιστοιχούσε σε τυχαίο δείγμα (π.χ. 0.5). Αν μια πόλη έχει πλυθυσμό 1000000 τότε μας αρκεί πιθανότητα επιλογής συμμετεχόντων από την πόλη ίση με 0.0005.

# Δειγματοληπτικά βάρη

- Αυτό σημαίνει ότι κάθε συμμετέχοντας από την απομακρυσμένη περιοχή αντιπροσωπεύει 2 άτομα ( $sw=2$ ) ενώ κάθε συμμετέχοντας από την πόλη 2000 άτομα ( $sw=2000$ )
- Επομένως για να αποκατασταθεί η δομή του αρχικού πληθυσμού, τα δεδομένα κάθε ατόμου της απομακρυσμενης περιοχής θα διπλασιαστούν, ενώ τα δεδομένα κάθε ατόμου της πόλης θα πολλαπλασιαστούν επί 2000

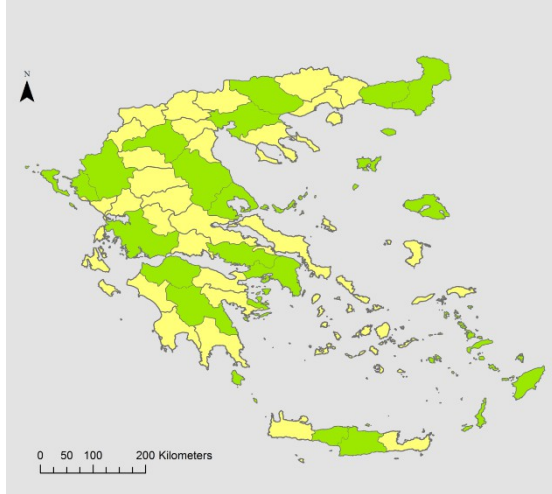
# Δειγματοληψία ΕΜΕΝΟ

Πολυσταδιακή Στρωματοποιημένη κατά Συστάδες

- Στρώματα
  - Γεωγραφική περιοχή / βαθμός αστικότητας
- Οικοδομικά Τετράγωνα (συστάδες)
  - proportional to size sampling
- Νοικοκυριά: Συστηματική Δειγματοληψία
- Άτομα
  - Απλή τυχαία δειγματοληψία
- Μεθοδολογία: Πόρτα-πόρτα

# Μέθοδος Δειγματοληψίας: Γενικός πληθυσμός (Πολυσταδιακή στρωματοποιημένη κατά συστάδες δειγματοληψία)

Στάδιο 1: Νομοί/βαθμός αστικ.



Στάδιο 2: οικοδομικά τετράγωνα

Στάδιο 3: Κατοικίες

Στάδιο 4: Άτομα  
(πιο πρόσφατη  
ημερομηνία γέννησης)

- Συνεντεύξεις: «πόρτα-πόρτα»

## Αντιπροσωπευτικότητα: Στρωματοποίηση

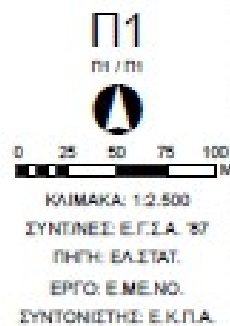
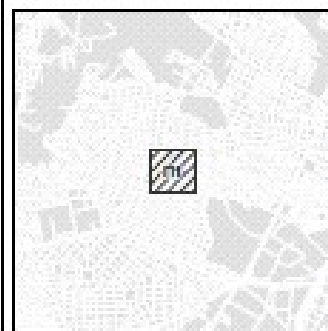
- Συνολικό Απαιτούμενο Δείγμα: 6.000

	Απαιτούμενο Δείγμα	Δειγμ. Σημεία	Απαιτούμενες Συνεντεύξεις	Μέγιστο Δείγμα	Γεωγρ. Κατανομή Δειγματοληπτικών Σημείων	
ΜΑΚΕΔΟΝΙΑ	666	257	22	12	264	10 ΣΕΡΡΕΣ
		141	12	12	144	12 ΚΟΖΑΝΗ
		268	34	8	272	7 ΣΕΡΡΕΣ
						5 ΚΟΖΑΝΗ
ΣΤΕΡΕΑ ΕΛΛΑΔΑ	479	147	13	12	156	21 ΣΕΡΡΕΣ
		139	12	12	144	13 ΚΟΖΑΝΗ
		193	25	8	200	8 ΑΙΤΩΛΟΑΚΑΡΝΑΝΙΑ
						5 ΒΟΙΩΤΙΑ
						7 ΑΙΤΩΛΟΑΚΑΡΝΑΝΙΑ
					5 ΒΟΙΩΤΙΑ	
					18 ΑΙΤΩΛΟΑΚΑΡΝΑΝΙΑ	
					7 ΒΟΙΩΤΙΑ	

- 22 Νομαρχίες
- 66 Στρώματα (22x3 βαθμοί αστικότητας)
- 577 δειγματοληπτικά σημεία
- 12 (ή 8) συνεντεύξεις/σημείο
- 6.156 μέγιστος αριθμός συνεντεύξεων

# ΑΓΙΑ ΒΑΡΒΑΡΑ - ΔΣ 365

ΔΗΜΟΣ ΑΓΙΑΣ ΒΑΡΒΑΡΑΣ





		Οικ. Τετρ	Μον. Πληθ	Νοικοκυριά	Μον. Πληθ %	Δείγμα	Αναλυτ. Δείγμα	Δειγμ. Κλάσμα	Νοικυριά/ Τετράγωνο	Ελάχιστα Τετράγωνα	Τετράγωνα Δείγματος	Νοικοκυριά Δείγματος
ΘΡΑΚΗ	ΑΣΤΙΚΑ	2.938	161.570	53.851	3,3%	299	133	1.214,8	18,3	8	16	330
	ΗΜΙΑΣΤΙΚΑ	1.551	48.811	15.742			40	1.220,3	10,1	4	8	92
	ΑΓΡΟΤΙΚΑ	7.346	153.098	49.914			126	1.215,1	6,8	19	37	232
ΜΑΚΕΔΟΝΙΑ	ΑΣΤΙΚΑ	8.685	476.436	156.043	12,2%	689	247	1.928,9	18,0	14	28	641
	ΗΜΙΑΣΤΙΚΑ	8.827	262.636	84.631			136	1.931,1	9,6	15	29	265
	ΑΓΡΟΤΙΚΑ	34.247	591.485	194.548			306	1.933,0	5,7	54	108	640
ΘΕΣΣΑΛΟΝΙΚΗ	ΑΣΤΙΚΑ	7.439	852.441	295.357	9,9%	570	448	1.902,8	39,7	12	24	725
	ΗΜΙΑΣΤΙΚΑ	4.215	163.604	50.889			86	1.902,4	12,1	8	16	368
	ΑΓΡΟΤΙΚΑ	3.742	67.956	21.755			36	1.887,7	5,8	7	14	57
ΘΕΣΣΑΛΙΑ	ΑΣΤΙΚΑ	7.166	356.059	114.720	6,8%	451	217	1.640,8	16,0	14	28	486
	ΗΜΙΑΣΤΙΚΑ	3.719	123.568	38.480			75	1.647,6	10,3	8	14	107
	ΑΓΡΟΤΙΚΑ	14.685	260.488	84.609			159	1.638,3	5,8	28	56	425
ΣΤΕΡΕΑ	ΑΣΤΙΚΑ	4.790	223.588	70.474	7,1%	480	138	1.620,2	14,7	10	20	267
	ΗΜΙΑΣΤΙΚΑ	7.438	214.135	67.578			132	1.622,2	9,1	15	29	337
	ΑΓΡΟΤΙΚΑ	18.558	339.513	112.618			210	1.616,7	6,1	35	69	543
ΗΠΕΙΡΟΣ	ΑΣΤΙΚΑ	1.416	110.398	38.052	3,1%	295	97	1.138,1	26,9	4	8	192
	ΗΜΙΑΣΤΙΚΑ	1.285	48.711	15.454			43	1.132,8	12,0	4	8	49
	ΑΓΡΟΤΙΚΑ	9.779	177.283	59.137			155	1.143,8	6,0	26	51	223
ΙΟΝΙΟ	ΑΣΤΙΚΑ	775	46.991	16.206	1,9%	241	54	870,2	20,9	3	6	137
	ΗΜΙΑΣΤΙΚΑ	966	29.461	9.682			34	866,5	10,0	4	8	207
	ΑΓΡΟΤΙΚΑ	5.310	133.156	44.996			153	870,3	8,5	19	38	612
ΑΙΓΑΙΟ	ΑΣΤΙΚΑ	3.546	157.741	52.835	4,6%	355	111	1.421,1	14,9	8	15	249
	ΗΜΙΑΣΤΙΚΑ	4.447	123.539	40.845			87	1.420,0	9,2	10	18	167
	ΑΓΡΟΤΙΚΑ	12.718	222.417	79.213			157	1.416,7	6,2	26	50	358
ΑΤΤΙΚΗ	ΑΣΤΙΚΑ	48.506	3.632.079	1.269.324	35,6%	1.620	1.511	2.403,8	26,2	58	116	3730
	ΗΜΙΑΣΤΙΚΑ	9.878	221.673	68.196			92	2.409,5	6,9	14	28	142
	ΑΓΡΟΤΙΚΑ	2.610	40.821	14.097			17	2.401,2	5,4	4	8	24
ΠΕΛΟΠΟΝΝΗΣΟΣ	ΑΣΤΙΚΑ	8.582	416.985	130.915	10,1%	607	230	1.813,0	15,3	16	32	647
	ΗΜΙΑΣΤΙΚΑ	5.626	187.234	56.036			103	1.817,8	10,0	11	22	153
	ΑΓΡΟΤΙΚΑ	23.037	495.852	157.896			274	1.809,7	6,9	40	80	588
ΚΡΗΤΗ	ΑΣΤΙΚΑ	4.298	261.832	86.811	5,4%	393	173	1.513,5	20,2	9	18	413
	ΗΜΙΑΣΤΙΚΑ	2.364	92.154	29.210			61	1.510,7	12,4	5	10	149
	ΑΓΡΟΤΙΚΑ	11.860	240.382	84.278			159	1.511,8	7,1	23	45	344
		<b>292.349</b>	<b>10.934.097</b>	<b>3.664.392</b>	<b>100,0%</b>	<b>6.000</b>	<b>6.000</b>	<b>1.822,3</b>		<b>535</b>	<b>1.057</b>	<b>13.899</b>

# Weights στην EMENO (1)

- Βάρη δειγματοληψίας *sampling weights* ή *base weights*:
  - Στοιχεία του σχεδιασμού, διορθώνουν για *oversampling* και γενικά πιθανότητα επιλογής ατόμων με διαφορετική πιθανότητα ανά στρώμα
  - Σε κάθε συμμετέχοντα αποδίδεται βάρος που αντιστοιχεί στον αντίστροφο της πιθανότητας επιλογής
  - Κάθε συμμετέχων “βαρύνει” στην ανάλυση ανάλογα με τον αριθμό ατόμων που εκπροσωπεί στον πληθυσμό αναφοράς

# Weights στην EMENO (2)

- Βάρη μετα τη στρωματοποίηση:
  - Παρά το ότι η επιλογή του ατόμου του νοικοκυριού έγινε τυχαία, το δείγμα δεν είχε τη δομή του πληθυσμού ως προς φύλο και ηλικία (γυναίκες και άτομα μεγαλύτερης ηλικίας τείνουν να συμμετέχουν περισσότερο από άνδρες και νεότερα άτομα)
  - Δημιουργήθηκαν βάρη ώστε η κατανομή φύλου και ηλικίας στη μελέτη να αντικατοπτρίζει σωστά τις αντίστοιχες κατανομές στον ελληνικό πληθυσμό

# Weights στην EMENO (3)

- Κάποιοι συμμετέχοντες αρνήθηκαν να συμμετέχουν στην υπομελέτη των εξετάσεων υγείας.
  - Εκτιμήθηκε η πιθανότητα συμμετοχής, δεδομένων των ατομικών χαρακτηριστικών μέσω μοντέλων παλινδρόμησης
  - Δημιουργηθηκαν βάρη αντίστροφα της εκτιμημένης πιθανότητας συμμετοχής.
  - Αποδίδοντας αυτά τα βάρη, κάθε συμμετέχων “βαρύνει” στη μελέτη ανάλογα με το πόσα άτομα με τα χαρακτηριστικά του αρνήθηκαν να συμμετέχουν

# Weights στην ΕΜΕΝΟ (4)

- Θα έπρεπε να γίνει κάτι ανάλογο και για την άρνηση συμμετοχής στη μελέτη
- Πώς θα γινόταν αυτό;
- Γιατί είναι δύσκολο;

**Table 2.** Age and sex distribution in the National Survey of Morbidity and Risk Factors (EMENO) sample and the corresponding distributions based on the 2011 census.

Age (years)	Population (Census 2011), n (%)		Sample (EMENO), n (%)	
	Male	Female	Male	Female
18-29	817,789 (9.16)	765,498 (8.57)	298 (4.97)	336 (5.61)
30-39	827,542 (9.27)	807,762 (9.05)	350 (5.84)	469 (7.82)
40-49	781,112 (8.75)	799,983 (8.96)	406 (6.77)	614 (10.24)
50-59	677,018 (7.58)	714,836 (8.01)	425 (7.09)	688 (11.48)
60-69	543,421 (6.09)	590,624 (6.62)	484 (8.08)	627 (10.46)
70-79	456,247.9 (5.11)	560,995 (6.28)	383 (6.39)	474 (7.87)
80+	231,746 (2.60)	351,588 (3.94)	200 (3.34)	241 (4.02)
Total	4,334,875 (48.56)	4,591,286 (51.44)	2546 (42.48)	3447 (57.52)

# Χαρακτηριστικά συμμετεχόντων EMENO πριν και μετά τη χρήση βαρών

	N (%) Αδρό	Σταθμισμένο
<b>Φύλο</b>		
Άνδρας	2043 (42.7)	48.64%
Γυναίκα	2743 (57.3)	51.36%
<b>Ηλικιακή ομάδα</b>		
18-29	487 (10.2)	17.54%
30-39	652 (13.6)	18.28%
40-49	812 (17.0)	17.76%
50-59	878 (18.4)	15.64%
60-69	892 (18.7)	12.73%
70-79	701 (14.7)	11.46%
80+	359 (7.5)	6.59%
<b>Median (IQR)</b>	55.0 (40.4, 68.0)	48 (34.4,64)

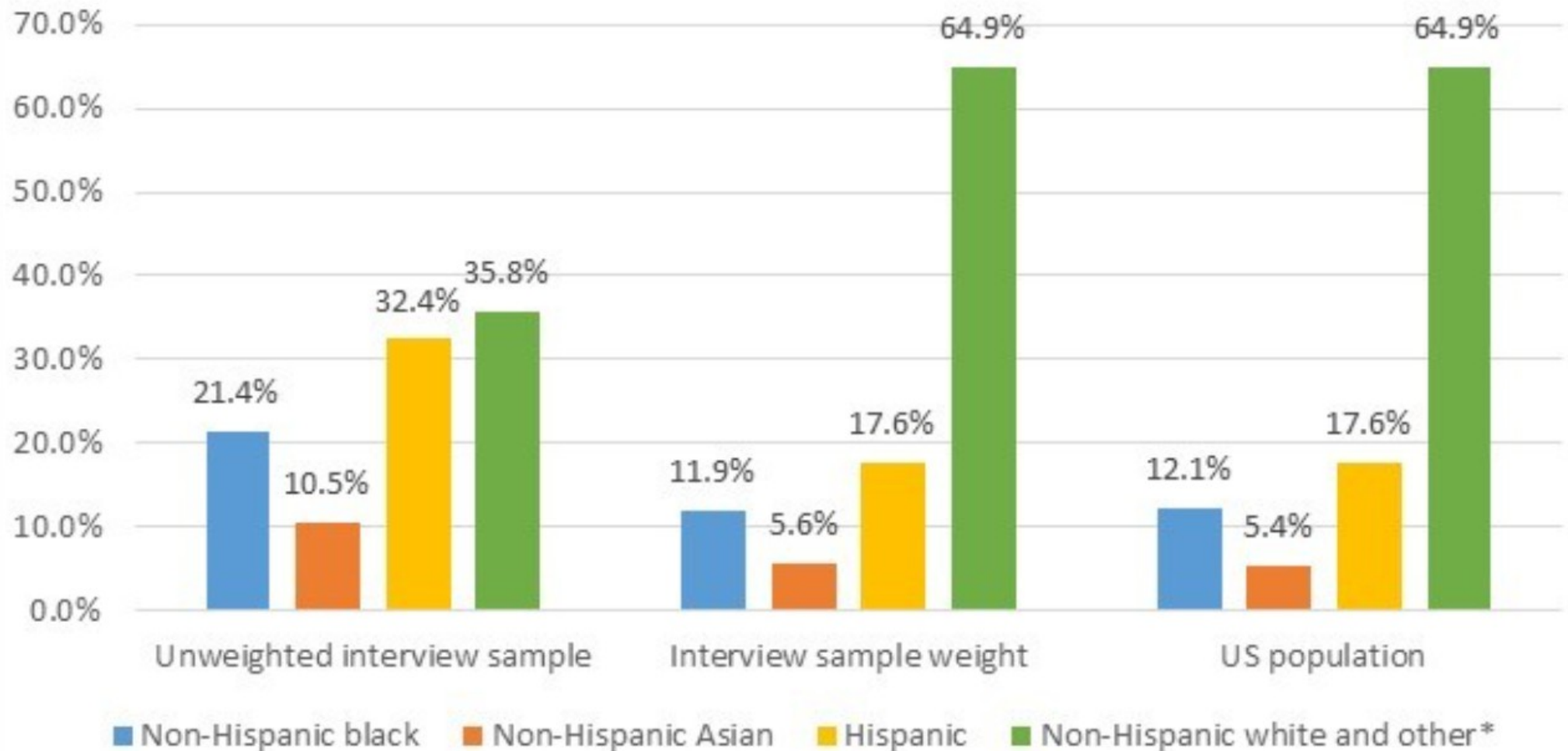
## • Adjusting for oversampling

“NHANES is designed to sample larger numbers of certain subgroups of particular public health interest in order to increase the reliability and precision of estimates of health status indicators for these population subgroups. The sample weights allow estimates from these subgroups to be combined to obtain national estimates that reflect the true relative proportions of these groups in the U.S. population as a whole.”



# Unweighted and weighted interview samples vs. the US population

Distribution of race and Hispanic origin, NHANES 2015-2016



# Μέγεθος δείγματος

- Ο αριθμός των ατόμων που πρέπει να συμπεριληφθούν στη μελέτη εξαρτάται από τις παρακάτω παραμέτρους:
  - I. Διαφορά που θέλουμε να εντοπίσουμε
  - II. Μεταβλητότητα τους μεγέθους που μελετάμε
  - III. Επίπεδο σημαντικότητας
  - IV. Ισχύ της μελέτης

# I. Διαφορά που θέλουμε να εντοπίσουμε

- Μια μεγάλη διαφορά εντοπίζεται και σε ένα μικρό δείγμα, ενώ όσο πιο μικρή διαφορά θέλουμε να εντοπίσουμε, τόσο μεγαλύτερο δείγμα θα χρειαστούμε.
- Αυτό σημαίνει πως το μέγεθος του δείγματος είναι αντιστρόφως ανάλογο με της διαφοράς που θέλουμε να εντοπίσουμε

# Μεταβλητότητα του χαρακτηριστικού που μελετάμε

- Η μεταβλητότητα των μετρήσεων αντανακλάται μέσω της τυπικής απόκλισης ή της διασποράς
- Όσο μεγαλύτερη είναι η μεταβλητότητα (τυπική απόκλιση και διασπορά) τόσο μεγαλύτερο δείγμα απαιτείται.
- Συνεπώς το μέγεθος του δείγματος είναι ανάλογο με την μεταβλητότητα

# Επίπεδο σημαντικότητας

- Συνδέεται με το σφάλμα τύπου I που είναι η πιθανότητα απόρριψης της μηδενικής υπόθεσης ενώ είναι αληθής και υνήθως επιλέγεται ίσο με 0.05.
- Αν το επίπεδο σημαντικότητας επιλεγεί να είναι μικρότερο, τότε το απαιτούμενο μέγεθος δείγματος (δεδομένων των υπολοίπων παραμέτρων) θα είναι μεγαλύτερο.
- Συνεπώς το μέγεθος του δείγματος συνδέεται αντίστροφα με το επίπεδο σημαντικότητας.

Προσοχή! Το επίπεδο σημαντικότητας καθορίζεται κατά το σχεδιασμό μιας μελέτης και είναι μας ενδιαφέρει μόνο εφόσον εντοπιστεί μια στατιστικώς σημαντική διαφορά

# Ισχύς της μελέτης

- Η ισχύς της μελέτης είναι η πιθανότητα να εντοπιστεί μια αληθής στατιστικώς σημαντική διαφορά.
- Η ισχύς είναι ίση με  $1 - \beta$ , όπου  $\beta$  το σφάλμα τύπου II, δηλαδή η πιθανότητα να μην απορρίψουμε τη μηδενική υπόθεση, ενώ η μηδενική υπόθεση δεν ισχύει.
- Χρειαζόμαστε μεγαλύτερο δείγμα για να ελζχιστοποιήσουμε το σφάλμα τύπου II
- Συνεπώς το μέγεθος του δείγματος είναι μεγαλύτερο για μικρότερο  $\beta$  ή αντίστοιχα για μεγαλύτερη ισχύ.

Προσοχή! Το  $\beta$  καθορίζεται κατά το σχεδιασμό και είναι σημαντικό μόνο στην περίπτωση που δεν απορριφθεί η μηδενική υπόθεση

Η ισχύς λαμβάνεται υπόψιν μόνο στις συγκριτικές μελέτες και όχι σε όσες στοχεύουν στην εκτίμηση ποσοστών (π.χ. Επιπολασμό). Σε αυτού του είδους τις μελέτες, εκτιμάται διάστημα εμπιστοσύνης με συγκεκριμένο περιθώριο λάθους.

# Μέγεθος δείγματος και είδος ερευνητικού ερωτήματος

## Κύριος στόχος μελετών:

- Διερεύνηση, περιγραφή και εκτίμηση δεικτών (Περιγραφικές μελέτες)
- Έλεγχο υπόθεσης, για παράδειγμα σύγκριση τιμών δείκτη σε διαφορετικές ομάδες (Αναλυτικές)
- Στατιστικά μοντέλα για τη σχέση έκβασης με περισσότερες από μία μεταβλητές (Πολυπαραγοντικές μελέτες)
- Μέγεθος δείγματος και είδος του ερευνητικού ερωτήματος:
  - Περιγραφικές μελέτες: Διάστημα εμπιστοσύνης,  $\alpha$
  - Συγκριτικές μελέτες: Ισχύ μελέτης,  $\beta$
  - Επιπλέον μεταβλητές απαιτούν μεγαλύτερο δείγμα, ώστε να διατηρείται η ισχύς και στις υποομάδες

# Υπολογισμός Μεγέθους δείγματος

- Ακρίβεια εκτίμησης:
- Παράδειγμα: Δεδομένα συγχρονικής μελέτης (Περιγραφικά)
- Ποσοτική μεταβλητή:
  - 1 δείγμα:  $n = (z\sigma/E)^2$
  - 2 δείγματα:  $n_i = 2(z\sigma^*/E)^2$
- Κατηγορική μεταβλητή:
  - 1 δείγμα  $n = p(1-p)(z/E)^2$
  - 2 δείγματα  $n = [p_1(1-p_1) + p_2(1-p_2)](z/E)^2$

$$*\sigma^{pooled} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$



# Υπολογισμός Μεγέθους δείγματος

□ Έλεγχος Υπόθεσης:

□ Παράδειγμα: Δεδομένα συγκριτικής μελέτης

□ Ποσοτική μεταβλητή:

- 1 δείγμα:  $n = \left( \frac{z_{1-\alpha/2} + z_{1-B}}{ES} \right)^2$        $ES = \frac{[\mu_1 - \mu_0]}{\sigma}$

- 2 δείγματα:  $n_i = 2 \left( \frac{z_{1-\alpha/2} + z_{1-B}}{ES} \right)^2$        $ES = \frac{[\mu_1 - \mu_2]}{\sigma^*}$

□ Κατηγορική μεταβλητή:

- 1 δείγμα  $n = \left( \frac{z_{1-\alpha/2} + z_{1-B}}{ES} \right)^2$        $ES = \frac{|p_1 - p_0|}{\sqrt{p_0(1-p_0)}}$

- 2 δείγματα  $n_i = 2 \left( \frac{z_{1-\alpha/2} + z_{1-B}}{ES} \right)^2$        $ES = \frac{|p_1 - p_2|}{\sqrt{p(1-p)}}$

# Παράδειγμα 1: Εκτίμηση επιπέδων σακχάρου προπτυχιακών φοιτητών

Εκτίμηση δείκτη/ποσοτική μεταβλητή/1 δείγμα

- Αν θέλουμε  $\alpha=0.05$ , δηλαδή 95% διάστημα εμπιστοσύνης το  $z= 1.96$
- Από προηγούμενη πιλοτική μελέτη η τυπική απόκλιση έχει εκτιμηθεί ίση με 25
- Θέλουμε το περιθώριο σφάλματος να είναι το πολύ 5mg/dl

$$n = (z\sigma/E)^2$$

$$n = (1.96 \times 25 / 5)^2$$

Δηλαδή το  $n$  είναι 96

## Παράδειγμα 2: Ποιο είναι το ποσοστό ανδρών και γυναικών καπνιστών στο ΕΚΠΑ;

### Εκτίμηση/κατηγορικά δεδομένα/2 δείγματα

- Για 95% διάστημα εμπιστοσύνης  $z = 1.96$
- Από πιλοτική μελέτη για τους άνδρες  $p = 0.25$  ενώ για τις γυναίκες  $p = 0.2$
- Θέλουμε περιθώριο σφάλματος το πολύ 10% (.1)

- Συνεπώς:  $n = [p_1(1 - p_1) + p_2(1 - p_2)](z / E)^2$

$$n = [.25(1 - .25) + .2(1 - .2)](1.96 / .1)^2$$

$$n = 95 \text{ (ανά ομάδα)}$$

# Παράδειγμα 3: Διαφέρουν τα ποσοστά καπνιστών μεταξύ των φοιτητών και του γενικού πληθυσμού?

## Έλεγχος υπόθεσης/κατηγορικά δεδομένα/1 δείγμα

- Επιλέγουμε  $\alpha=0.05$  (άρα  $z_{1-\alpha/2} = 1.96$ ); Ισχύ (1-B) 80% (άρα  $z_{1-B} = 0.84$ )
- Από πιλοτική μελέτη μεταξύ των φοιτητών είναι  $p = 0.22$  ενώ στο γενικό πληθυσμό είναι  $p = 0.3$

➤ Τότε:

$$n = \left( \frac{z_{1-\alpha/2} + z_{1-B}}{ES} \right)^2$$

$$ES = \frac{|p_1 - p_0|}{\sqrt{p_0(1-p_0)}}$$

$$ES = \frac{|.22 - .3|}{\sqrt{.3(1-.3)}}$$

$$n = \left( \frac{1.96 + .84}{.17} \right)^2$$

$$n = 272$$

# Παράδειγμα 4: Διαφέρουν τα επίπεδα σακχάρου σε άνδρες και γυναίκες?

## Έλεγχος υπόθεσης/συνεχή δεδομένα/2 δείγματα

- Επιλέγουμε  $\alpha=0.05$  (άρα  $z_{1-\alpha/2} = 1.96$ ); Ισχύ (1-B) 80% (άρα  $z_{1-B} = 0.84$ )

➤ Από πιλοτική μελέτη μεση τιμή σακχάρου στις γυναίκες: 95mg/dl με  $sd = 10$ ; στους άνδρες: 100mg/dl με  $sd = 10$

➤ Συνεπώς:

$$n_i = 2 \left( \frac{z_{1-\alpha/2} + z_{1-B}}{ES} \right)^2$$

$$ES = \frac{[\mu_1 - \mu_2]}{\sigma}$$

$$ES = \frac{|95 - 100|}{10}$$

$$n = 2 \left( \frac{1.96 + .84}{.5} \right)^2$$

$$n = 63 \text{ (ανά ομάδα)}$$

# Μέγεθος δείγματος και σχεδιασμός

- Ο υπολογισμός του μεγέθους του δείγματος πρέπει να λαμβάνει υπόψιν στοιχεία της δειγματοληπτικής μεθόδου
  - Για παράδειγμα τη στρωματοποίηση

# ΕΜΕΝΟ μέγεθος δείγματος

1 δείγμα με κύριο στόχο εκτίμηση επιπολασμού νοσημάτων

Βασίστηκε στην εκτίμηση του επιπολασμού της ιογενούς ηπατίτιδας Β- Κατηγορική έκβαση

Εκτίμηση επιπολασμού από προηγούμενες μελέτες: 1.5% δηλαδή  $p=0.015$

Σφάλμα τύπου I:  $\alpha=0.05$

Περιθώριο σφάλματος:  $E=0.003$

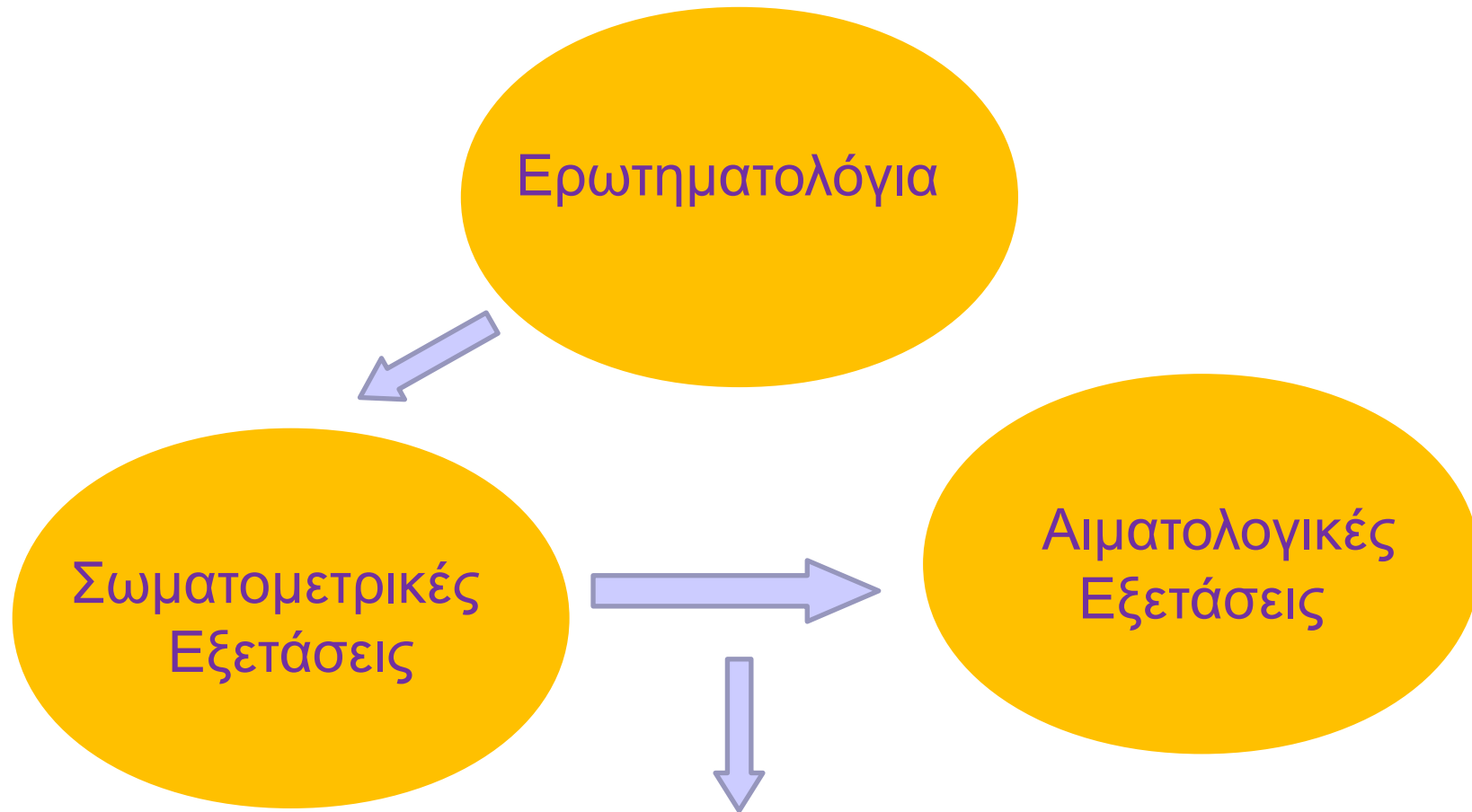
$$N=z^2 p(1-p)/E^2 \approx 6000$$

# EMENO μέγεθος δείγματος

On the basis of the ATTICA study results, the overall prevalence of hypercholesterolemia (total serum cholesterol levels  $>200$  mg/dL or use of lipid lowering agents), hypertension (average blood pressure levels  $>140/90$  mmHg or receiving antihypertensive medication), and of diabetes mellitus (fasting blood glucose  $>125$  mg/dL or receiving antidiabetic medication) was about 40%, 32%, and 7%, respectively. **Calculating sampling errors using the corresponding formula for stratified sampling and with a target sample size of 6000 individuals**, the above-mentioned prevalence could be estimated with 1.28%, 1.22%, and 0.66% precision, respectively.



# EMENO: Health Examination Survey



**Τυχαίο Δείγμα 6.000 ενηλίκων ( $\geq 18$  ετών)**

# Δομημένα Πρωτυποποιημένα Ερωτηματολόγια

- Βασικά δημογραφικά/ Κοινωνικο-οικονομική κατάσταση
- Κατάσταση υγείας
- Κλίμακα ικανοποίησης από τη ζωή
- Χρήση υπηρεσιών υγείας / ικανοποίηση/ φάρμακα
- Μέτρα πρόληψης
- Σωματική δραστηριότητα
- Προσθήλωση μεσογειακή διατροφή
- Κάπνισμα, αλκοόλ
- Περιβάλλον /μετακινήσεις
-

# Σωματομετρικές Παράμετροι

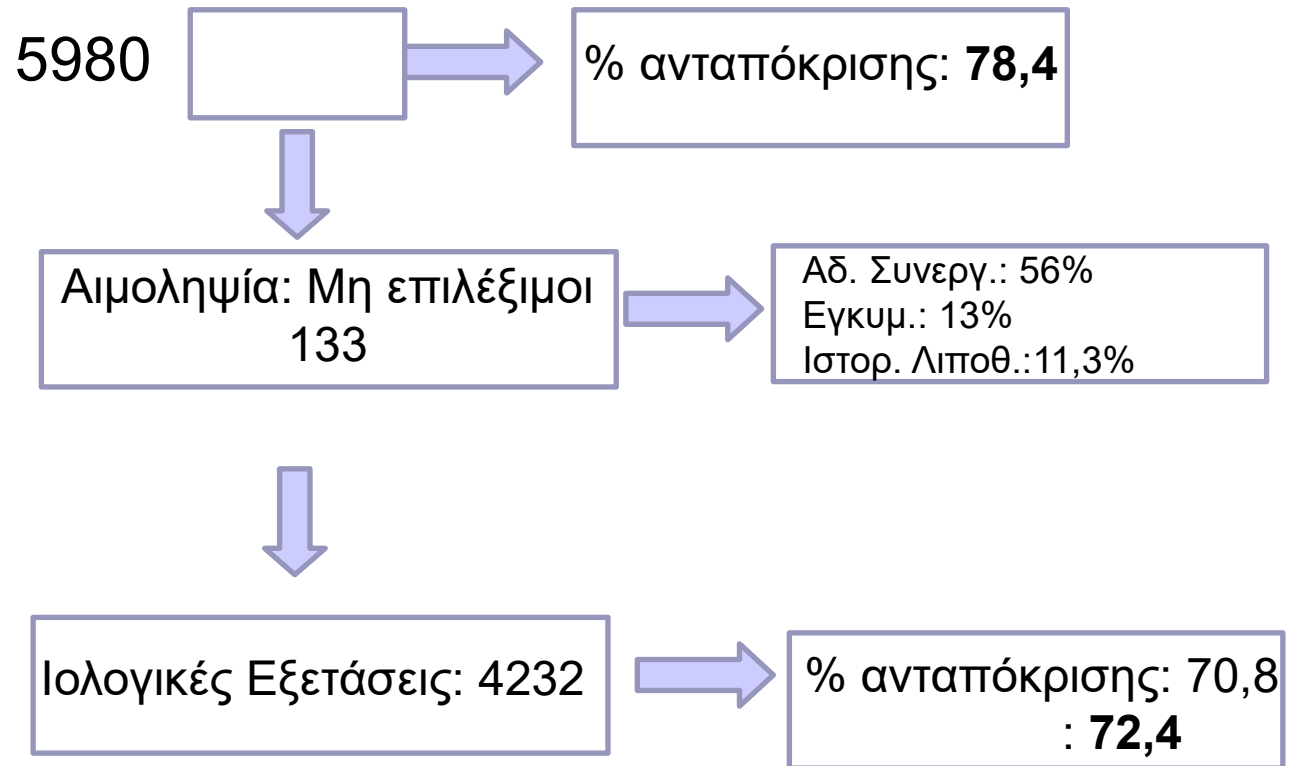
- Ύψος
- Βάρος
- Περίμετρος μέσης, βραχίονα
- Αρτηριακή πίεση (ένδειξη κοιλιακής μαρμαρυγής)
- Σπυρομέτρηση

# Αιματολογικές Εξετάσεις

- Ολική χοληστερόλη ορού
- HDL- χοληστερόλη
- Γλυκόζη
- Γλυκοζυλιωμένη αιμοσφαιρίνη

**Δεδομένα Ατμοσφαιρικής Ρύπανσης**

# Γενικός πληθυσμός: διάγραμμα ροής



# Ανάλυση Surveys Stata

[SVY] svyset -- Declare survey design for dataset

## Syntax

### Single-stage design

- `svyset [psu] [weight] [, design_options options]`

### Multiple-stage design

- `svyset psu [weight] [, design_options] [|| ssu , design_options] ... [options]`

## Example (NHANES)

- `. webuse nhanes2f`
- `. svyset psuid [pweight=finalwgt], strata(stratid)`

# Stata

- .  
Αφού δηλωθεί ότι τα δεδομένα προέρχονται από survey με συγκεκριμένο σχεδιασμό, ο σχεδιασμός λαμβάνεται υπόψη στις εκτιμήσεις.
- webuse nhanes2f
- . svyset psuid [pweight=finalwgt], strata(stratid)
- . svy: mean zinc
- . svy: regress zinc age age2 weight female black orace rural

# Δυνατότητα παρακολούθησης

- Συχνά ζητείται από τους συμμετέχοντες να συναινέσουν για μελλοντική επικοινωνία
- Για όσους συμμετέχοντες συναινέσουν, γίνεται νέος κύκλος συλλογής δεδομένων, με όσες πληροφορίες πρέπει να ανανεωθούν
- Η μελέτη σε αυτή την περίπτωση γίνεται από συγχρονική, μελέτη κοορτής



# Respondent Driven Sampling

- Χρήσιμο για μελέτες “κρυφών πληθυσμών”
- Εύκολα προσβάσιμα άτομα του πληθυσμού προσεγγίζονται
- Στη συνέχεια, συστήνουν μέλη του κύκλου τους για να συμμετέχουν στη μελέτη
- - Ποιο είναι το πρόβλημα με το δείγμα που έχει συλλεχθεί με αυτό τον τρόπο;
  - Μπορεί να διορθωθεί και πώς;