



Μεθοδολογία
Κοινωνικών
Επιστημών

Δειγματοληψία



ΠΜΣ Πολιτικής Επιστήμης και
Κοινωνιολογίας

ΕΚΠΑ/Τμήμα Πολιτικής Επιστήμης &
Δημόσιας Διοίκησης

- Εξάμηνο: Χειμερινό 2024-25
- Μάνος Τσατσάνης (Επίκουρος Καθηγητής)
- etsats@pspa.uoa.gr

Δειγματοληψία

Η δειγματοληψία αποτελεί το μέσο για να μάθουμε κάτι για μία μεγάλη ομάδα ανθρώπων (ή άλλων μονάδων) χωρίς να χρειαστεί να μελετήσουμε κάθε μέλος αυτής της ομάδας

Ένα δείγμα το οποίο αντιπροσωπεύει τον πληθυσμό έρευνας σε παραμέτρους που μας ενδιαφέρουν αποκαλείται αντιπροσωπευτικό δείγμα.

Χρονικό και χρηματικό κόστος

Η μελέτη όλων ανεξαιρέτως των περιπτώσεων είναι συχνά ανέφικτη ή εξαιρετικά κοστοβόρα.

Γιατί επιλέγουμε δείγματα;

- Σε ποιες περιπτώσεις **δεν είναι απαραίτητη** η δειγματοληψία;

1. Όταν ο αριθμός αυτών που θέλουμε να εξετάσουμε είναι μικρός.

- Π.χ. Ενδιαφερόμαστε για την διερεύνηση του προγραμματικού λόγου (σημαντικών) ακροδεξιών κομμάτων στη Νότια Ευρώπη κατά την περίοδο της κρίσης. Επιλέγουμε προεκλογικά μανιφέστα από το σύνολο των κομμάτων (πληθυσμό έρευνας)

2. Όταν τα δεδομένα είναι εύκολα προσβάσιμα (π.χ. υφιστάμενες βάσεις δεδομένων πλήρους κάλυψης).

Ορισμός πληθυσμού έρευνας

- **Πληθυσμός Έρευνας** είναι το σύνολο των ατόμων/μονάδων που αποτελούν το τμήμα εκείνο του πληθυσμού που θέλουμε να διερευνήσουμε. Είναι ο πληθυσμός, στον οποίο θέλουμε να γενικεύσουμε παίρνοντας μόνο ένα δείγμα του.
- **Πληθυσμός-στόχος** (target population): Σχετίζεται με τον πληθυσμό έρευνας αλλά δεν είναι συνώνυμοι όροι. Ο πληθυσμός-στόχος είναι ο πληθυσμός που παρουσιάζει θεωρητικό ενδιαφέρον για την έρευνα και αποτελεί υποσύνολο του πληθυσμού έρευνας
- Ο Ορισμός του Πληθυσμού μιας έρευνας είναι μια απόφαση που απαιτεί γνώση επίσημων στατιστικών στοιχείων, έλεγχο των διαθέσιμων πρωτογενών πηγών, ευθυγράμμιση με το ερευνητικό μας ερώτημα και συνείδηση του διαθέσιμου χρόνου και των διαθέσιμων πόρων. Με μια λέξη: Απαιτεί κρίση από την πλευρά του ερευνητή.
- Τα ερωτήματα που θέτουμε στον εαυτό μας όταν θέλουμε να ορίσουμε τον πληθυσμό έρευνας είναι:
 1. Ποιοι «δικαιούνται» ή μπορούν να μας απαντήσουν;
 2. Πού θα βρούμε αυτούς πού θα μπορούσαν να μας απαντήσουν; (Πού είναι ευκολότερο να τους βρούμε; Υπάρχουν;)
 3. Ποιοι θέλουμε να μας απαντήσουν, με βάση τον ερευνητικό μας σχεδιασμό;
 4. Σε ποιο/ποια χρονικό/ά σημεία; (π.χ. ανάλυση περιεχομένου)

Ορισμός πληθυσμού έρευνας (παράδειγμα)

ΕΡΕΥΝΑ ΠΡΟΘΕΣΗΣ ΨΗΦΟΥ
ΓΙΑ ΤΙΣ ΔΗΜΟΤΙΚΕΣ ΕΚΛΟΓΕΣ
ΣΤΟ ΔΗΜΟ ΑΘΗΝΑΙΩΝ



Ποιοι **«δικαιούνται»/θα μπορούσαν** να μας απαντήσουν;

Όσοι έχουν δικαίωμα ψήφου και ψηφίζουν στο Δήμο Αθηναίων.

Πού θα βρούμε αυτούς που θέλουμε και δικαιούνται να μας απαντήσουν;

Στην Αθήνα+Σε όλη την Ελλάδα (Ετεροδημότες)

Ποιοι **θέλουμε** να μας απαντήσουν, με βάση τον ερευνητικό μας σχεδιασμό (λαμβάνοντας πάντα υπόψη μας το χρόνο, το κόστος και την εφικτότητα);

Δημότες + Ψηφοφόροι + Κάτοικοι Λεκανοπεδίου

ΠΛΗΘΥΣΜΟΣ ΕΡΕΥΝΑΣ:

Άνδρες και γυναίκες ηλικίας 17 ετών και άνω που ψηφίζουν στο Δήμο Αθηναίων και κατοικούν στο Λεκανοπέδιο Αττικής

Το δειγματοληπτικό πλαίσιο (sampling frame)

- Το Δειγματοληπτικό πλαίσιο είναι ένας (όσο το δυνατόν) εξαντλητικός κατάλογος (ηλεκτρονικά αρχεία, τηλεφωνικοί κατάλογοι, αρχείο εκπομπών, δημοτικά μητρώα κλπ) που περιλαμβάνει το σύνολο των δειγματοληπτικών μονάδων (άτομα, ψηφοφόρους, νοικοκυριά, κτίρια, κείμενα κλπ) που θέλουμε να διερευνήσουμε.
- Το δειγματοληπτικό πλαίσιο μπορεί είτε να το κατασκευάσει ο ερευνητής είτε να υπάρχει έτοιμο και διαθέσιμο.
- Σε έρευνες μεγάλης κλίμακας δεν υπάρχει συνήθως η πρακτική δυνατότητα κατασκευής δειγματοληπτικού πλαισίου και, συνεπώς, χρησιμοποιούμε έτοιμους καταλόγους.

Μέθοδοι δειγματοληψίας

- Η δειγματοληψία διαχωρίζεται σε δύο γενικές κατηγορίες:
 1. **Τυχαία** (πιθανοτική) δειγματοληψία
 2. **Μη τυχαία** (μη πιθανοτική) δειγματοληψία
- Στην **Τυχαία δειγματοληψία** κάθε μονάδα ενός πληθυσμού έχει γνωστή εκ των προτέρων (συντά την ίδια) πιθανότητα να επιλεγεί στο δείγμα, δηλαδή δεν αποκλείονται με συστηματικό τρόπο συγκεκριμένα τμήματα του πληθυσμού αυτού (bias-μεροληψία).
- Στη **Μη τυχαία δειγματοληψία** δεν είναι γνωστή η πιθανότητα κάθε μονάδας να επιλεγεί στο δείγμα.
- Τα **πιθανοτικά** δείγματα είναι συνήθως πιο **αντιπροσωπευτικά** του πληθυσμού έρευνας, σε σχέση με τα μη πιθανοτικά δείγματα.

Τύποι πιθανοτικής δειγματοληψίας

ΑΠΛΗ ΤΥΧΑΙΑ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

- Σε κάθε μονάδα του πληθυσμού δίνουμε με τυχαίο τρόπο έναν μοναδικό αριθμό και στη συνέχεια επιλέγουμε τυχαία τους αριθμούς που αντιστοιχούν στις δειγματοληπτικές μονάδες (άτομα) που θα διερευνήσουμε.

Η απλή τυχαία δειγματοληψία συνίσταται κυρίως στις περιπτώσεις που:

- Διαθέτουμε ένα πλήρες δειγματοληπτικό πλαίσιο
- Ο πληθυσμός μας είναι γεωγραφικά συγκεντρωμένος ή έχουμε όλες τις μονάδες διαθέσιμες.

Διαφορετικά, αυξάνεται ιδιαίτερα το κόστος και ο απαιτούμενος χρόνος διεξαγωγής της έρευνας.

Πλεονεκτήματα:

- Περισσότερο αντιπροσωπευτικά δείγματα
- Δυνατότητα μέτρησης πιστότητας/αντιπροσωπευτικότητας δείγματος.

Τύποι πιθανοτικής δειγματοληψίας

ΣΥΣΤΗΜΑΤΙΚΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

- Η Συστηματική δειγματοληψία είναι παρόμοια με την Απλή Τυχαία.
- Η διαφορά είναι ότι για να επιλέξουμε τις μονάδες (π.χ. άτομα) που θα συμπεριλάβουμε στην έρευνά μας «κατασκευάζουμε» έναν κανόνα:
- Διαιρούμε τον πληθυσμό μας δια το επιθυμητό δείγμα και το πηλίκο αποτελεί το «βήμα», με βάση το οποίο επιλέγουμε τα άτομα που θα συμμετάσχουν στην έρευνα.
- Π.χ. Αν το δειγματοληπτικό μας πλαίσιο περιέχει 2.000 άτομα και θέλουμε ένα δείγμα 100 ατόμων, διαιρούμε $2000/100=20$ και παίρνουμε, ξεκινώντας από τυχαίο σημείο, 1 άτομο κάθε 20.
- Στη συστηματική δειγματοληψία ισχύουν οι ίδιοι περιορισμοί με την απλή τυχαία.

Τύποι πιθανοτικής δειγματοληψίας

ΣΤΡΩΜΑΤΟΠΟΙΗΜΕΝΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

Η Στρωματοποιημένη δειγματοληψία είναι περισσότερο σύνθετη από τις δύο προηγούμενες (απλή τυχαία και συστηματική), αλλά εξασφαλίζει μεγαλύτερη ακρίβεια και αντιπροσωπευτικότητα, συνδυάζοντάς τες.

- Ομαδοποιούμε τον πληθυσμό σε ομοιογενείς ομάδες (στρώματα) πριν από τη δειγματοληψία, συνήθως με βάση χαρακτηριστικά που αφορούν την έρευνά μας.
- Εντός κάθε στρώματος εφαρμόζουμε απλή τυχαία ή συστηματική δειγματοληψία.
- Πλεονεκτήματα: Κόστος, Χρόνος, ακρίβεια (εκπροσώπηση) σε ειδικές ομάδες που μας ενδιαφέρουν.

Τύποι πιθανοτικής δειγματοληψίας

ΠΟΛΥΣΤΑΔΙΑΚΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΚΑΤΑ ΣΥΣΤΑΔΕΣ (MULTISTAGE CLUSTER SAMPLING)

- Στη δειγματοληψία κατά συστάδες, η πρωτογενής μονάδα δειγματοληψίας (στο πρώτο στάδιο της δειγματοληπτικής διαδικασίας) δεν είναι οι μονάδες του πληθυσμού που προορίζονται για τη δειγματοληψία αλλά ομαδοποιήσεις αυτών των μονάδων
- Εάν, για παράδειγμα, ο πληθυσμός της μελέτης σας είναι οι μαθητές λυκείου μιας χώρας, μπορείτε να ξεκινήσετε με έναν κατάλογο των λυκείων, να επιλέξετε με τυχαίο τρόπο ορισμένα σχολεία από αυτόν τον κατάλογο, και στη συνέχεια να επιλεγεί δείγμα των μαθητών εντός αυτών των σχολείων
- Χρησιμοποιείται όταν δεν γίνεται (ή δεν βολεύει ή δεν συμφέρει), να κατασκευαστεί/χρησιμοποιηθεί εξαντλητικό δειγματοληπτικό πλαίσιο των δειγματοληπτικών μονάδων.
- Πρακτικό πλεονέκτημα: μεγαλύτερη γεωγραφική συγκέντρωση

Πηγές σφάλματος σχετιζόμενες με τη δειγματοληψία

Είδη ερευνητικών σφαλμάτων, τα οποία συνδέονται με τη δειγματοληψία:

- Σφάλμα κάλυψης
- Σφάλμα μη απόκρισης
- Δειγματοληπτικό σφάλμα

Σφάλματα κάλυψης

- Σφάλματα, τα οποία προκύπτουν από τις διαφορές μεταξύ του δειγματοληπτικού πλαισίου και του πληθυσμού-στόχος
- Κατά κανόνα, αποκλείεται η συμμετοχή ορισμένων προσώπων, εφόσον τα δείγματα επιλέγονται από τηλεφωνικούς καταλόγους, άδειες κυκλοφορίας οχημάτων κτλ.
- Μη καταχωρημένες τηλεφωνικές συνδέσεις - Μία από τις σημαντικότερες πηγές σφάλματος κάλυψης.
 - Οι δημοσκοπici χρησιμοποιούν τη μέθοδο των τυχαίων κλήσεων για να ξεπεράσουν το πρόβλημα των μη καταχωρημένων συνδέσεων.

Πραγματοποίηση τυχαίων κλήσεων:

- Μέθοδος επιλογής ερωτώμενων σε μια τηλεφωνική έρευνα, η οποία στηρίζεται στον σχηματισμό τηλεφωνικών αριθμών με τυχαίο τρόπο.

Σφάλματα μη απόκρισης

Σφάλματα, τα οποία προκύπτουν από τη διαφορά μεταξύ των μελών του πληθυσμού έρευνας που συμμετέχουν και αυτών που δεν δέχονται ή δεν έχουν τη δυνατότητα να συμμετάσχουν σε μια έρευνα.

Για διάφορους λόγους, τα άτομα που θα επιλέξουμε δειγματοληπτικά τελικά μπορεί να μην συμπεριληφθούν στο δείγμα μας.

ΛΟΓΟΙ:

1. Άρνηση συμμετοχής.
 2. Αδυναμία επαφής μαζί τους/εύρεσής τους
 3. Αδυναμία συμμετοχής.
- Η μη απόκριση προκαλεί δύο βασικά προβλήματα:
 1. Μη αποδεκτή μείωση του δείγματος και
 2. Μεροληψία (bias).

Η μη απόκριση

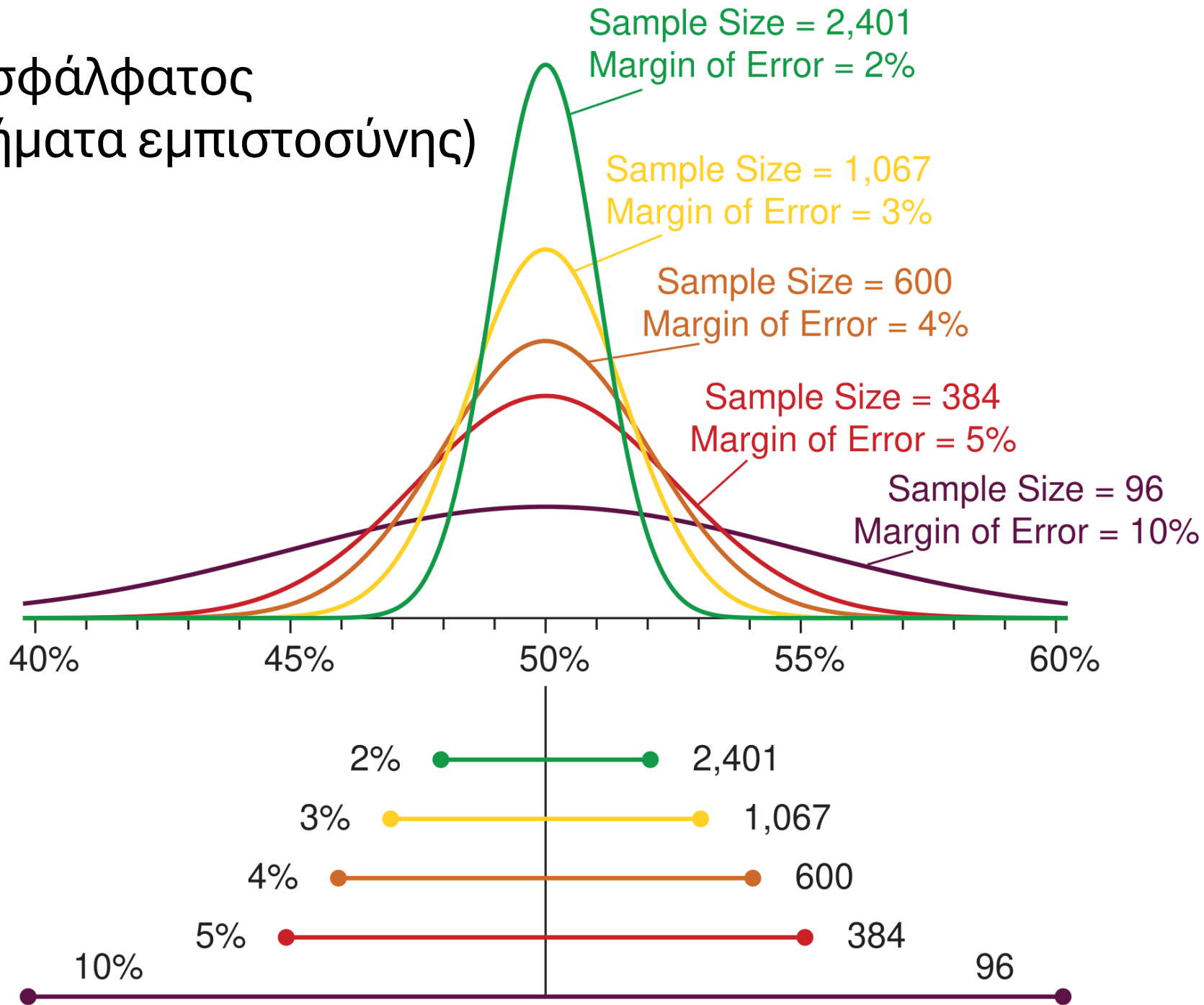
Η μείωση του δείγματος αντιμετωπίζεται με:

1. Καλή εκπαίδευση ερευνητών
2. Επανακλήσεις/επανεπισκέψεις/υπομνήσεις σε διαφορετικές ώρες και ημέρες
3. Επιλογή μεγαλύτερου δείγματος. Γνωρίζουμε ότι ακόμα και στην πιο προσεκτικά σχεδιασμένη έρευνα το ποσοστό μη απόκρισης θα σημαντικό (σίγουρα πάνω από 20%) οπότε επιλέγουμε αρχικά δείγμα σημαντικό μεγαλύτερο από αυτό που χρειαζόμαστε

Δειγματοληπτικά σφάλματα

- ▶ Αποκλίσεις μεταξύ των χαρακτηριστικών ενός δείγματος (**στατιστικά στοιχεία**) και των χαρακτηριστικών του πληθυσμού (**παράμετροι**) από τον οποίο επιλέχθηκε το δείγμα.
- ▶ Δηλαδή, μπορεί ένα δείγμα να μην αντιπροσωπεύει με ακρίβεια τον πληθυσμό από τον οποίο επιλέχτηκε, ακόμα και αν δεν υπάρχει σφάλμα κάλυψης ή σφάλμα μη-απόκρισης (να μην είναι **μεροληπτικό (biased) δείγμα**).
- ▶ Οφείλονται στη **δειγματοληπτική μεταβλητότητα**:
 - ▶ Μεταβλητότητα που ενδέχεται να παρατηρηθεί στα στατιστικά στοιχεία, όταν λαμβάνονται διαφορετικά δείγματα από τον ίδιο πληθυσμό
- ▶ Ο βασικός τρόπος να μειώσουμε το δειγματοληπτικό σφάλμα είναι να αυξήσουμε το μέγεθος του δείγματος

Περιθώριο σφάλματος (95% διαστήματα εμπιστοσύνης)



Η μη τυχαία δειγματοληψία

- ▶ Υπάρχουν περιπτώσεις όπου καμία μορφή τυχαίας δειγματοληψίας δεν είναι αποτελεσματική και χρειάζεται να εφαρμόσουμε **μη τυχαία δειγματοληψία**, κυρίως για λόγους κόστους και χρόνου.
- ▶ Π.χ. αν θέλουμε να κάνουμε έρευνα σε ανθρώπους που έχουν φυλακιστεί κάποια φορά στη ζωή τους, θα ήταν εξαιρετικά χρονοβόρο και κοστοβόρο να εφαρμόζαμε κάποια μορφή τυχαίας δειγματοληψίας στο γενικό πληθυσμό.
- ▶ Γενικά, μη τυχαία δειγματοληψία εφαρμόζουμε όταν **δεν γνωρίζουμε τα ακριβή ποσοτικά χαρακτηριστικά του πληθυσμού** που θέλουμε να διερευνήσουμε και όταν δεν υπάρχει ή δεν μπορεί να κατασκευαστεί δειγματοληπτικό πλαίσιο.
- ▶ Επίσης, η μη τυχαία δειγματοληψία ενδείκνυται για τα πρώτα, πιλοτικά στάδια μιας έρευνας.
- ▶ Το βασικό μειονέκτημα της μη τυχαίας δειγματοληψίας είναι ότι **δεν επιτρέπει εύκολα γενικεύσεις στο σύνολο του πληθυσμού**.

Η μη τυχαία δειγματοληψία

ΣΚΟΠΙΜΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ (PURPOSIVE SAMPLING)

- ▶ Βασίζεται στην ορθολογική κρίση του ερευνητή, ο οποίος επιλέγει τους ερωτώμενους (ή π.χ. τα κείμενα) με βάση την καταλληλότητα τους για τη διερεύνηση του ερευνητικού ερωτήματος
- ▶ Π.χ. Μια έρευνα για τους ηγέτες του ακτιβιστικού οικολογικού κινήματος/μια έρευνα για τον χουλιγκανισμό/ μια έρευνα σε δημόσιους λειτουργούς που εγκρίνουν ή απορρίπτουν αιτήσεις πολιτικού ασύλου.
- ▶ Πιο συχνά χρησιμοποιείται σε ποιοτικές έρευνες, έρευνες ελίτ, μέλη κοινωνικών κινημάτων, κλπ: Στοχευμένη επιλογή ατόμων με εξειδικευμένες γνώσεις, εμπειρίες κλπ.

Η μη τυχαία δειγματοληψία

ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΕΥΚΑΙΡΙΑΣ/ΕΥΚΟΛΙΑΣ (CONVENIENCE SAMPLING)

Επιλέγουμε με βάση και μόνο τη διαθεσιμότητα των υποκειμένων, αν δεν είναι εφικτή ή δεν χρειάζεται (π.χ. πιλοτική έρευνα) κάποια άλλη μέθοδος δειγματοληψίας.

ΔΕΙΓΜΑΤΟΛΗΨΙΑ «ΧΙΟΝΟΣΤΟΙΒΑΔΑΣ» (SNOWBALL SAMPLING)

Όταν αναζητούμε ένα πολύ ειδικό, μικρό και «κλειστό» κοινό (π.χ. πρώην φυλακισμένοι) ξεκινάμε βρίσκοντας ένα άτομο που ανήκει στο κοινό αυτό και στη συνέχεια του ζητάμε να μας συστήσει και άλλους. Με αυτόν τον τρόπο το δείγμα μας αυξάνεται όπως η χιονόμπαλα όταν κυλάει.

Ποιοτικές έρευνες – π.χ. Δύσκολα προσεγγίσιμες/περιθωριακές κοινωνικές ομάδες

Η μη τυχαία δειγματοληψία

ΠΟΣΟΣΤΙΚΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ (QUOTA SAMPLING)

- ▶ Φέρει ομοιότητες με την πιθανοτική δειγματοληψία. Γνωρίζοντας από πριν τα χαρακτηριστικά του πληθυσμού που μας ενδιαφέρει, σχεδιάζουμε εκ των προτέρων λεπτομερώς το δείγμα μας, χρησιμοποιώντας τις ποσοστώσεις που θέλουμε να επιτύχουμε.
- ▶ Μειονέκτημα: μπορεί να μεροληπτούμε χωρίς να το καταλαβαίνουμε.
- ▶ Συνηθισμένη σε διαδικτυακές έρευνες επισκόπησης (surveys)