

# Lab 4: Solutions

13/01/2026

## Contents

<b>Εισαγωγή: Τυχαιοποίηση (Randomization)</b>	<b>3</b>
Γιατί τη χρησιμοποιούμε; . . . . .	3
<b>Άσκηση 1 – Απλή τυχαιοποίηση (Simple Randomization)</b>	<b>4</b>
1. Πιθανότητα ακριβούς ισορροπίας (Exact balance) . . . . .	4
2. Προσομοίωση (Simulation) μιας Απλής Τυχαιοποίησης . . . . .	5
3. Monte Carlo Simulation: Πολλαπλές επαναλήψεις της ίδιας προσομοίωσης ( $K = 1000$ replications) . . . . .	7
<b>Άσκηση 2 – Τυχαιοποίηση σε blocks (Permuted Block Randomization)</b>	<b>9</b>
1. Πλήθος δυνατών αναθέσεων εντός block . . . . .	9
2. Block randomization με τυχαία μεγέθη block (random block sizes) . . . . .	9
<b>Άσκηση 3 – Blocking μέσα σε strata (π.χ. φύλο)</b>	<b>13</b>
Λύση . . . . .	14
Σημαντική Διδακτική Παρατήρηση για τα Baseline Tables . . . . .	15
Θεωρητική Σημείωση: Όρια της Stratification . . . . .	16
Αριθμητικό Παράδειγμα: Early Stopping με Stratified Block Randomization . . . . .	17
<b>Άσκηση 4 – Urn scheme</b>	<b>19</b>
i. Υλοποίηση ως συνάρτηση . . . . .	19
ii. Προσομοίωση 2500 αναθέσεων και γράφημα πιθανότητας . . . . .	20

<b>Η έννοια της “Μνήμης”</b>	<b>22</b>
<b>Άσκηση 5 – Σύγκριση SR vs Block vs Urn σε μικρό N</b>	<b>22</b>

## Εισαγωγή: Τυχαιοποίηση (Randomization)

Η **τυχαιοποίηση (randomization)** είναι διαδικασία ανάθεσης (allocation) με βάση έναν **προκαθορισμένο στοχαστικό μηχανισμό (stochastic allocation mechanism)**, ώστε κάθε συμμετέχων να έχει γνωστή πιθανότητα να ανατεθεί σε κάθε σκέλος της μελέτης. Στην **απλή τυχαιοποίηση (simple randomization)**, οι αναθέσεις γίνονται ανεξάρτητα μεταξύ συμμετεχόντων.

Κεντρική ιδιότητα της τυχαιοποίησης είναι ότι καθιστά την ανάθεση θεραπείας **ανεξάρτητη (independent)** από προγνωστικούς παράγοντες (prognostic factors) και συγχυτικούς παράγοντες (confounders), **γνωστούς ή άγνωστους, κατά μέσο όρο (on average)**.

### Γιατί τη χρησιμοποιούμε;

1. **Περιορισμός μεροληψίας επιλογής (selection bias):** Η τυχαιοποίηση, σε συνδυασμό με κατάλληλη απόκρυψη της ακολουθίας τυχαιοποίησης (**allocation concealment**), μειώνει δραστικά τη δυνατότητα συστηματικής επιρροής της ανάθεσης.
2. **Ισορροπία χαρακτηριστικών κατά μέσο όρο (covariate balance on average):** Αν και δεν εγγυάται τέλεια ισορροπία σε κάθε δείγμα, οδηγεί σε ισορροπία *κατά μέσο όρο* (in expectation), συμπεριλαμβανομένων μη μετρήσιμων ή άγνωστων συγχυτών.
3. **Θεμέλιο για έγκυρη αιτιώδη σύγκριση:** Επειδή η τυχαιοποίηση δημιουργεί ομάδες που είναι συγκρίσιμες κατά μέσο όρο, οποιαδήποτε διαφορά παρατηρήσουμε στις εκβάσεις μπορεί να αποδοθεί στην επίδραση της θεραπείας και όχι σε άλλους παράγοντες—υπό την προϋπόθεση ότι δεν υπάρχουν άλλα προβλήματα στη διεξαγωγή της μελέτης (π.χ. αλληλεπίδραση μεταξύ ασθενών, προβλήματα στη μέτρηση κ.λπ.).

**Σημείωση:** Τα στατιστικά μοντέλα μπορούν να προσαρμόσουν για *γνωστούς και μετρήσιμους* παράγοντες (measured confounders), αλλά δεν εξασφαλίζουν προστασία από *άγνωστους/μη μετρήσιμους* συγχυτικούς παράγοντες. Η τυχαιοποίηση είναι ο μηχανισμός που στοχεύει σε αυτό *κατά μέσο όρο*.

## Άσκηση 1 – Απλή τυχαιοποίηση (Simple Randomization)

Πρόκειται να πραγματοποιηθεί **απλή τυχαιοποίηση (simple randomization)** για  $N = 100$  συμμετέχοντες σε δύο σκέλη (arms), A και B, με **αναλογία κατανομής (allocation ratio)** 1:1. Για κάθε συμμετέχοντα, η πιθανότητα ανάθεσης στο σκέλος A είναι  $p = 0.5$  και στο σκέλος B είναι  $1 - p = 0.5$ , ανεξάρτητα από τις αναθέσεις των άλλων συμμετεχόντων.

### 1. Πιθανότητα ακριβούς ισορροπίας (Exact balance)

i. Υπολογίστε την πιθανότητα η απλή τυχαιοποίηση να οδηγήσει σε **ακριβή ισορροπία (exact balance)**, δηλαδή  $N_A = 50$  στο σκέλος A και  $N_B = 50$  στο σκέλος B.

#### Λύση

Η μεταβλητή που περιγράφει τον αριθμό των ατόμων στο σκέλος A ακολουθεί τη Διωνυμική Κατανομή:  $X \sim \text{Binomial}(N = 100, p = 0.5)$ . Η πιθανότητα για  $k = 50$  υπολογίζεται από τον τύπο:

$$P(X = 50) = \binom{100}{50} \cdot 0.5^{50} \cdot 0.5^{50} = \binom{100}{50} \cdot 0.5^{100}$$

Υπόδειξη: Χρησιμοποιήστε τη συνάρτηση `dbinom()` στην R.

```
N <- 100
# Υπολογισμός P(N_A = 50)
dbinom(x = N/2, size = N, p = 0.5)
```

```
[1] 0.07958924
```

ii. Σχολιάστε το αποτέλεσμα: Θεωρείτε την πιθανότητα αυτή «μεγάλη» ή «μικρή»; Τι συνεπάγεται αυτό για τη χρήση της απλής τυχαιοποίησης σε μελέτες με μικρό ή μεσαίο δείγμα;

#### Λύση

Με βάση το αποτέλεσμα ( $P \approx 0.08$ ), προκύπτουν ορισμένα κρίσιμα συμπεράσματα για τον σχεδιασμό των κλινικών δοκιμών:

- **Αμερόληπτη Κατανομή (Unbiased Allocation):** Η απλή τυχαιοποίηση (simple randomization) είναι μια **αμερόληπτη** διαδικασία. Η **αναμενόμενη τιμή** (expected value) της κατανομής είναι η πλήρης ισορροπία 50-50 μεταξύ των δύο **σκελών** (arms).
- **Το Παράδοξο της Ισορροπίας:** Παρόλο που η διαδικασία δεν “ευνοεί” κανένα σκέλος, η **τέλεια ισορροπία** (perfect balance) είναι σπάνια στην πράξη.

**Αναλογία:** Το φαινόμενο είναι ταυτόσημο με τη ρίψη 100 νομισμάτων. Ενώ αναμένουμε στατιστικά 50 κορώνες και 50 γράμματα, η πιθανότητα να φέρουμε ακριβώς αυτή την αναλογία είναι μόλις  $\sim 8\%$ .

- **Πρακτικές Συνέπειες:** Στις κλινικές δοκιμές, η σημαντική απόκλιση από την επιθυμητή **αναλογία κατανομής** (allocation ratio) μπορεί να επηρεάσει τη στατιστική ισχύ (statistical power) της μελέτης, ειδικά σε αναλύσεις υποομάδων.
- **Ανάγκη για Εναλλακτικές Μεθόδους:** Για να εξασφαλιστεί η ισορροπία στον αριθμό των συμμετεχόντων, ιδιαίτερα όταν το δείγμα  $N$  είναι μικρό, χρησιμοποιούνται τεχνικές όπως:
  1. **Τυχαιοποίηση κατά ομάδες** (Permuted Block Randomization).
  2. **Τυχαιοποίηση τύπου κάλπης** (Urn Randomization).

## 2. Προσομοίωση (Simulation) μιας Απλής Τυχαιοποίησης

Προσομοιώστε μία φορά την απλή τυχαιοποίηση για  $N = 100$  συμμετέχοντες και παρουσιάστε τα πλήθη/ποσοστά ανά σκέλος.

### Λύση

Στην ενότητα αυτή, θα προσομοιώσουμε τη διαδικασία της απλής τυχαιοποίησης για  $N = 100$  συμμετέχοντες χρησιμοποιώντας τη γεννήτρια τυχαίων αριθμών της R.

### Βήματα:

1. Δημιουργούμε ένα σύνολο δεδομένων (dataset) με τις στήλες `id` (αναγνωριστικό) και `trt` (treatment group).
2. Χρησιμοποιούμε τη συνάρτηση `runif(N)`, η οποία παράγει τυχαίους αριθμούς από την ομοιόμορφη κατανομή στο διάστημα  $[0, 1]$ .
3. Ορίζουμε ένα **seed** (τυχαίο σπόρο) για να διασφαλίσουμε την **αναπαραγωγιμότητα** (reproducibility) των αποτελεσμάτων.
4. Αναθέτουμε τους συμμετέχοντες ως εξής:
  - **Σκέλος A** αν η τυχαία τιμή  $u < 0.5$
  - **Σκέλος B** αν η τυχαία τιμή  $u \geq 0.5$

```

# Παράμετροι
N <- 100
set.seed(125) # Για αναπαραγωγικότητα

# Δημιουργία data frame
trial_data <- data.frame(
  id = paste0("id", 1:N),
  trt = NA_character_
)

# Παραγωγή τυχαίων αριθμών
u <- runif(N)

# Ανάθεση στα σκέλη (Arms)
trial_data$trt[u < 0.5] <- "A"
trial_data$trt[u >= 0.5] <- "B"

# Προβολή των πρώτων 10 παρατηρήσεων
head(trial_data, 10)

```

	id	trt
1	id1	B
2	id2	A
3	id3	A
4	id4	A
5	id5	B
6	id6	B
7	id7	B
8	id8	A
9	id9	B
10	id10	B

```

# Έλεγχος κατανομής (Counts & Percentages)
table(trial_data$trt)

```

A	B
44	56

```

prop.table(table(trial_data$trt))

```

A	B
0.44	0.56

### 3. Monte Carlo Simulation: Πολλαπλές επαναλήψεις της ίδιας προσομοίωσης ( $K = 1000$ replications)

i. Επαναλάβετε την προσομοίωση  $K = 1000$  φορές ii. Υπολογίστε την εμπειρική μέση τιμή και διακύμανση του  $N_A$ . iii. Υπολογίστε το ποσοστό των επαναλήψεων στις οποίες παρατηρείται «σοβαρή» αριθμητική ανισορροπία, οριζόμενη ως:  $N_A > 60$  ή  $N_A < 40$

Λύση:

```
# i. Εκτέλεση προσομοίωσης
# Παράμετροι προσομοίωσης
N <- 100
K <- 1000
set.seed(2342)

# Αποθήκευση αποτελεσμάτων
results <- data.frame(
  replication = 1:K,
  NumberA = NA_integer_
)

# Εκτέλεση Loop
for (i in 1:K) {
  u <- runif(N)
  # Υπολογίζουμε απευθείας το πλήθος στο σκέλος A
  results$NumberA[i] <- sum(u < 0.5)
}

# ii. Μέση τιμή και Διακύμανση
empirical_mean <- mean(results$NumberA)
empirical_var <- var(results$NumberA)

# iii. Ποσοστό σοβαρής ανισορροπίας (>60 ή <40)
imbalance_rate <- mean(results$NumberA > 60 | results$NumberA < 40)

# Προβολή αποτελεσμάτων
cat("Εμπειρική Μέση Τιμή:", empirical_mean, "\n")
```

Εμπειρική Μέση Τιμή: 49.887

```
cat("Εμπειρική Διακύμανση:", empirical_var, "\n")
```

Εμπειρική Διακύμανση: 28.05228

```
cat("Ποσοστό Ανισορροπίας > 60/40:", imbalance_rate * 100, "%\n")
```

Ποσοστό Ανισορροπίας > 60/40: 4.2 %

### Σχολιασμός Θεωρίας: Η Σύνδεση με τη Διωνυμική Κατανομή

**Γιατί επιλέγουμε τα όρια 40 και 60 ως κριτήριο ανισορροπίας;**

Στην απλή τυχαιοποίηση (simple randomization), ο αριθμός των ασθενών στο **σκέλος (arm) A** ( $N_A$ ) ακολουθεί τη **Διωνυμική Κατανομή** με παραμέτρους  $N = 100$  και  $p = 0.5$ . Σύμφωνα με τη στατιστική θεωρία:

- **Αναμενόμενη Τιμή:**  $E(N_A) = N \cdot p = 100 \cdot 0.5 = 50$
- **Θεωρητική Διακύμανση:**  $\sigma^2 = N \cdot p \cdot (1 - p) = 100 \cdot 0.5 \cdot 0.5 = 25$
- **Τυπική Απόκλιση:**  $\sigma = \sqrt{25} = 5$

Σύμφωνα με την προσέγγιση της **Κανονικής Κατανομής**, το 95% των παρατηρήσεων αναμένεται να βρίσκεται εντός  $\pm 1.96$  τυπικών αποκλίσεων από τη μέση τιμή:

$$50 \pm 1.96 \cdot 5 \approx 50 \pm 10 \Rightarrow [40, 60]$$

**Συμπέρασμα:** Περιμένουμε θεωρητικά ότι περίπου το **5%** των κλινικών δοκιμών που χρησιμοποιούν απλή τυχαιοποίηση θα καταλήξουν σε “προβληματική” ανισορροπία (δηλαδή εκτός του εύρους 40-60).

Η προσομοίωση που εκτελέσαμε επιβεβαιώνει τη θεωρία, αναδεικνύοντας τον κίνδυνο της απλής τυχαιοποίησης σε μελέτες με μικρό ή μεσαίο δείγμα, όπου η **αναλογία κατανομής (allocation ratio)** μπορεί να αποκλίνει σημαντικά από το επιθυμητό 1:1 λόγω τύχης.



## Άσκηση 2 – Τυχαιοποίηση σε blocks (Permuted Block Randomization)

Θεωρούμε τυχαιοποίηση 1:1 σε δύο σκέλη (arms) A και B με **τυχαιοποίηση σε blocks** (*permuted block randomization*).

### 1. Πλήθος δυνατών αναθέσεων εντός block

Θεωρήστε block συνολικού μεγέθους  $b = 4$  (δηλαδή 4 διαδοχικές αναθέσεις), με **ακριβώς 2 αναθέσεις στο A και 2 στο B**.

Πόσες διαφορετικές ακολουθίες ανάθεσης (**permutations**) είναι δυνατές;

**Λύση:** Σε block  $b = 4$  με 2 αναθέσεις στο A και 2 στο B, το πλήθος διαφορετικών ακολουθιών ισούται με το πλήθος τρόπων να επιλέξουμε τις 2 θέσεις (από τις 4) που θα πάρουν το A:

$$\binom{4}{2} = \frac{4!}{2!2!} = 6.$$

```
choose(4, 2)
```

```
[1] 6
```

### 2. Block randomization με τυχαία μεγέθη block (random block sizes)

Χρησιμοποιήστε το πακέτο `blockrand` για να παράγετε λίστα τυχαιοποίησης για τουλάχιστον  $N = 100$  συμμετέχοντες, με δύο σκέλη (A/B), όπου το συνολικό μέγεθος του block επιλέγεται τυχαία από το σύνολο  $\{2, 4, 6, 8\}$ .

Στη `blockrand()` αυτό υλοποιείται με `block.sizes = 1:4` και `num.levels = 2`, επειδή τα τελικά block sizes είναι το γινόμενο `num.levels * block.sizes`, δηλαδή  $2 \times \{1, 2, 3, 4\} = \{2, 4, 6, 8\}$ .

- Εξηγήστε γιατί η `blockrand()` μπορεί να επιστρέψει **περισσότερες** από  $N$  γραμμές.
- Δώστε την τελική κατανομή A/B στα **πρώτα 100** άτομα.

**Λύση:**

Η `blockrand()` δημιουργεί τη λίστα τυχαιοποίησης σε **πλήρη (ολοκληρωμένα) blocks**.

Αν το ζητούμενο μέγεθος  $N$  (π.χ.  $N = 100$ ) **δεν συμπίπτει** με το τέλος του τελευταίου block (λόγω block sizes 2/4/6/8), η συνάρτηση **μπορεί να επιστρέψει περισσότερες από  $N$  αναθέσεις** ώστε να ολοκληρώσει το τελευταίο block.

```
library(blockrand)
set.seed(1322)

block_rand <- blockrand(
  n = 100,
  num.levels = 2,
  levels = c("A", "B"),
  block.sizes = 1:4
)

# (i) Πόσες γραμμές επέστρεψε συνολικά;
nrow(block_rand)
```

```
[1] 102
```

```
# Κρατάμε ακριβώς τους πρώτους 100
block_rand_100 <- block_rand[1:100, ]

# (ii) Κατανομή A/B στα πρώτα 100 άτομα
table(block_rand_100$treatment)
```

```
  A  B
50 50
```

```
prop.table(table(block_rand_100$treatment))
```

```
  A  B
0.5 0.5
```

```
# Προαιρετικά: δείτε τα πρώτα blocks
block_rand[block_rand$block.id %in% 1:3, ]
```

	id	block.id	block.size	treatment
1	1	1	4	A
2	2	1	4	B
3	3	1	4	B
4	4	1	4	A
5	5	2	6	B
6	6	2	6	A
7	7	2	6	A
8	8	2	6	B
9	9	2	6	A
10	10	2	6	B
11	11	3	8	B
12	12	3	8	B
13	13	3	8	A
14	14	3	8	B
15	15	3	8	B
16	16	3	8	A
17	17	3	8	A
18	18	3	8	A

### iii. Ερμηνεία: γιατί τυχαία μεγέθη block;

Εξηγήστε τη σημασία των **τυχαίων μεγεθών block** (*randomly varying block sizes*) ως προς:

- την **ισορροπία** (balance) κατά τη διάρκεια της ένταξης,
- την **προβλεψιμότητα** (predictability) της επόμενης ανάθεσης και τον κίνδυνο **μεροληψίας επιλογής** (*selection bias*), ιδίως όταν δεν υπάρχει πλήρης **απόκρυψη ακολουθίας ανάθεσης** (*allocation concealment*).

---

### Ενδεικτική Απάντηση

Η block randomization εξασφαλίζει καλή **αριθμητική ισορροπία** μεταξύ των σκελών, ακόμη και κατά τη διάρκεια της ένταξης. Σε κάθε ολοκληρωμένο block, η κατανομή είναι ακριβώς 1:1.

Η χρήση **τυχαίων μεγεθών block** διατηρεί αυτό το πλεονέκτημα, αλλά ταυτόχρονα μειώνει την **προβλεψιμότητα** της ακολουθίας.

Αν το μέγεθος του block ήταν σταθερό και γνωστό (π.χ. πάντα 4), τότε προς το τέλος κάθε block η επόμενη ανάθεση θα μπορούσε συχνά να προβλεφθεί. Αυτό δημιουργεί κίνδυνο **selection bias**, αν το προσωπικό μπορεί να επηρεάσει ποιος θα ενταχθεί και πότε.

Με τυχαία block sizes, ακόμη και αν κάποιος γνωρίζει ότι χρησιμοποιείται block randomization, δεν γνωρίζει το τρέχον μέγεθος του block και επομένως δεν μπορεί εύκολα να προβλέψει την επόμενη ανάθεση.

---

### Κλινική Σημείωση

Η χρήση τυχαίων μεγεθών block αφορά κυρίως την προστασία από **μεροληψία επιλογής** μέσω καλύτερης

**allocation concealment**. Δεν ταυτίζεται με την **τυφλότητα (blinding)**, η οποία αφορά το αν τα άτομα γνωρίζουν τη θεραπεία *μετά* την ανάθεση.

---

### Σχόλιο: Μέγιστη Προσωρινή Ανισορροπία (Maximum Temporary Imbalance)

Η block randomization περιορίζει αυστηρά την προσωρινή ανισορροπία κατά την ένταξη.

#### Γενικός κανόνας (2 σκέλη, 1:1 allocation):

Σε block συνολικού μεγέθους  $b$  (άρτιος), με  $b/2$  αναθέσεις σε κάθε σκέλος, η μέγιστη δυνατή διαφορά πριν ολοκληρωθεί το block είναι:

$$\max |N_A - N_B| = \frac{b}{2}.$$

Στο παράδειγμα  $b = 4$ , η ανισορροπία δεν μπορεί να ξεπεράσει τα 2 άτομα, ακόμη κι αν η μελέτη σταματήσει στη μέση ενός block.

Αντίθετα, στην απλή τυχαιοποίηση δεν υπάρχει τέτοιο θεωρητικό όριο και η ανισορροπία μπορεί να γίνει πολύ μεγαλύτερη.

### Άσκηση 3 – Blocking μέσα σε strata (π.χ. φύλο)

Σε κλινική δοκιμή με συνολικό στόχο  $N = 100$ , αναμένουμε να ενταχθούν **55 άνδρες** και **45 γυναίκες**.

Η τυχαιοποίηση γίνεται: - με **block randomization**, - **ξεχωριστά μέσα σε κάθε stratum** (στρωματοποιημένη τυχαιοποίηση ως προς το φύλο), - με σκοπό τελικά να κρατήσουμε: - τις **πρώτες 55** αναθέσεις για άνδρες, - τις **πρώτες 45** αναθέσεις για γυναίκες.

#### Ζητούνται:

1. Να γίνει block randomization για Male και Female χωριστά.
  2. Να κρατηθούν οι πρώτες 55 και 45 αναθέσεις αντίστοιχα.
  3. Να ελεγχθεί ότι:
    - συνολικά η κατανομή A/B είναι περίπου 50/50,
    - μέσα σε κάθε φύλο η κατανομή είναι παρόμοια.
  4. Να γίνει έλεγχος  $\chi^2$  ανεξαρτησίας μεταξύ θεραπείας και φύλου με Monte Carlo p-value.
-

## Λύση

### Δημιουργία στρωματοποιημένης block randomization

```
library(blockrand)

set.seed(2435)
male   <- blockrand(n = 100, id.prefix = "M", block.prefix = "M", stratum =
  ↪ "Male")
female <- blockrand(n = 100, id.prefix = "F", block.prefix = "F", stratum =
  ↪ "Female")

male   <- male[1:55, ]
female <- female[1:45, ]

mystudy <- rbind(male, female)
```

### Έλεγχος συνολικής κατανομής

```
table(mystudy$treatment)
```

A	B
49	51

Η συνολική κατανομή A/B είναι κοντά στο 50/50, όπως αναμένεται από στρωματοποιημένη block randomization.

### Έλεγχος κατανομής μέσα σε κάθε φύλο

```
x <- table(mystudy$treatment, mystudy$stratum)
x
```

	Female	Male
A	22	27
B	23	28

Η κατανομή A/B είναι παρόμοια μέσα σε άνδρες και γυναίκες, επειδή η τυχαιοποίηση έγινε **ανεξάρτητα σε κάθε stratum**.

## Έλεγχος ανεξαρτησίας θεραπείας-φύλου

```
chisq.test(x, simulate.p.value = TRUE, B = 10000)
```

Pearson's Chi-squared test with simulated p-value (based on 10000 replicates)

```
data: x  
X-squared = 0.0004042, df = NA, p-value = 1
```

### Ερμηνεία του αποτελέσματος

Το  $p$ -value είναι σχεδόν 1, άρα δεν υπάρχει στατιστική ένδειξη συσχέτισης θεραπείας και φύλου στο συγκεκριμένο δείγμα. Αυτό είναι αναμενόμενο, αφού η στρωματοποιημένη τυχαιοποίηση επιβάλλει ισορροπία ως προς το φύλο.

---

## Σημαντική Διδακτική Παρατήρηση για τα Baseline Tables

Στους πίνακες baseline χαρακτηριστικών των κλινικών δοκιμών:

- **Δεν πρέπει** να χρησιμοποιούνται  $p$ -values για να “ελεγχθεί” αν η τυχαιοποίηση πέτυχε.

Γιατί:

1. Η τυχαιοποίηση είναι σωστή **εκ του σχεδιασμού**, όχι επειδή το δείγμα έδωσε “καλό”  $p$ -value.
2. Το  $p$ -value δεν μετρά την ποιότητα της τυχαιοποίησης, αλλά τη συμβατότητα του συγκεκριμένου δείγματος με τη μηδενική υπόθεση.
3. Ακόμη και αν  $p < 0.05$ , αυτό δεν σημαίνει αποτυχία της randomization, αλλά τυχαία απόκλιση (αναμενόμενη στο 5% των μελετών).
4. Με πολλές μεταβλητές, κάποια θα δώσουν “σημαντικό”  $p$ -value από τύχη (Type I error).

## Σύγχρονη πρακτική (CONSORT 2010)

- Παρουσιάζουμε baseline characteristics **χωρίς p-values**.
- Δίνουμε μόνο περιγραφικά μέτρα (μέσους, SD, ποσοστά).
- Τυχόν διαφορές αντιμετωπίζονται με **covariate adjustment**, όχι ως “απόδειξη” αποτυχίας της randomization.

Το τεστ εδώ έγινε **μόνο για εκπαιδευτικούς λόγους**.

---

## Θεωρητική Σημείωση: Όρια της Stratification

### 1. Πολλά strata

Αν ορίσουμε πάρα πολλά strata (π.χ. φύλο  $\times$  ηλικία  $\times$  στάδιο  $\times$  κέντρο), ο αριθμός των strata μπορεί να γίνει τεράστιος και κάθε stratum να έχει ελάχιστους ασθενείς. Τότε η stratification γίνεται πρακτικά άχρηστη.

### 2. Early stopping

Η block randomization εγγυάται ισορροπία **στο τέλος κάθε block**.

Αν η μελέτη σταματήσει στη μέση block, μπορεί να υπάρξει προσωρινή ανισορροπία έως  $b/2$  μέσα σε κάθε stratum.

Με πολλά strata, κάθε stratum μπορεί να έχει τη δική του προσωρινή ανισορροπία, αλλά αυτές τείνουν να εξισορροπούνται στο συνολικό δείγμα. Το βασικό πρόβλημα είναι η **διαχειριστική πολυπλοκότητα**, όχι η συσσώρευση ανισορροπίας.

### 3. Κανόνας αντίχειρα

- Stratify μόνο για 1–3 πολύ σημαντικούς prognostic factors.
  - Αν δημιουργούνται  $>5-6$  strata, σκέψου εναλλακτικές (minimization, covariate-adjusted analysis).
-



## Αριθμητικό Παράδειγμα: Early Stopping με Stratified Block Randomization

### Εκφώνηση

Σε μελέτη με  $N = 100$  εφαρμόζεται stratified block randomization ως προς το φύλο (2 strata: άνδρες/γυναίκες) με block size  $b = 4$  σε κάθε stratum.

Η μελέτη σταματά πρόωρα όταν έχουν τυχαιοποιηθεί συνολικά  $N = 47$  άτομα: - 25 άνδρες, - 22 γυναίκες.

### Ερώτημα:

Γιατί η κατανομή A/B σε κάθε φύλο δεν είναι μοναδικά καθορισμένη στο σημείο της διακοπής; Δώσε παραδείγματα έγκυρων κατανομών.

---

### Λύση

Η τυχαιοποίηση γίνεται **ανεξάρτητα σε κάθε stratum**.

Αρα υπάρχουν δύο ξεχωριστές ακολουθίες blocks (μία για άνδρες, μία για γυναίκες).

Με  $b = 4$ , κάθε πλήρες block έχει 2A και 2B.

Αν η μελέτη σταματήσει στη μέση ενός block, η μέγιστη προσωρινή ανισορροπία μέσα σε κάθε stratum είναι:

$$\max |A - B| = \frac{b}{2} = 2.$$

Στο σημείο διακοπής:

- **Άνδρες (25):** επιτρεπτές κατανομές  
 $13A - 12B, 12A - 13B, 14A - 11B, 11A - 14B$
- **Γυναίκες (22):** επιτρεπτές κατανομές  
 $11A - 11B, 12A - 10B, 10A - 12B$

όλες με  $|A - B| \leq 2$ .

### Συμπέρασμα:

Η κατανομή A/B στο σημείο του early stopping **δεν είναι μοναδική**, επειδή η διακοπή μπορεί να συμβεί σε διαφορετικό σημείο του τελευταίου block *ξεχωριστά σε κάθε stratum*.

**Βιβλιογραφία:**

Senn, S. (1994). *Testing for baseline balance in clinical trials*.  
Statistics in Medicine, 13(17), 1715–1726.

## Άσκηση 4 – Urn scheme

Υλοποίησε ένα απλό urn scheme όπου:

- Ξεκινάμε με 1 μπάλα A και 1 μπάλα B.
- Σε κάθε draw, τραβάμε A με πιθανότητα  $\frac{n_A}{n_A + n_B}$ .
- **Μετά** προσθέτουμε 1 μπάλα από την *αντίθετη* θεραπεία (δηλ. αν βγει A, προσθέτουμε μία B).

Η πιθανότητα να επιλεγεί η θεραπεία A στο επόμενο βήμα είναι το ποσοστό των μπαλών A στην κληρωτίδα:

$$P(A) = \frac{n_A}{n_A + n_B}$$

Όπου  $n_A$  και  $n_B$  είναι ο αριθμός των μπαλών στην κληρωτίδα.

Η διαδικασία εξισορρόπησης λειτουργεί ως εξής: Αν επιλεγεί η θεραπεία A, προσθέτουμε μία μπάλα της θεραπείας B (αντίθετη). Έτσι, ο παρονομαστής μεγαλώνει και ο αριθμητής ( $n_A$ ) παραμένει σταθερός, με αποτέλεσμα η  $P(A)$  να μειώνεται στο επόμενο βήμα. Αυτό ονομάζεται “Negative Feedback” και εμποδίζει τη μία θεραπεία να κυριαρχήσει.

### i. Υλοποίηση ως συνάρτηση

Γράψε `urnModel(n)` που επιστρέφει:

- `treatment`: η ακολουθία αναθέσεων
- `probA`: η πιθανότητα A σε κάθε βήμα

```

urnModel <- function(n, na0 = 1, nb0 = 1) {
  na <- na0
  nb <- nb0

  probA <- rep(NA_real_, n)
  trt <- rep(NA_character_, n)
  u <- runif(n)

  for (i in 1:n) {
    probA[i] <- na / (na + nb)

    if (u[i] < probA[i]) {
      trt[i] <- "A"
      nb <- nb + 1 # Προσθέτουμε μπάλα στο B (negative feedback για
↪ εξισορρόπηση)
    } else {
      trt[i] <- "B"
      na <- na + 1 # Προσθέτουμε μπάλα στο A (negative feedback για
↪ εξισορρόπηση)
    }
  }

  data.frame(treatment = trt, probA = probA)
}

```

## ii. Προσομοίωση 2500 αναθέσεων και γράφημα πιθανότητας

1. Τρέξε `urnModel(2500)` και δες πόσοι καταλήγουν σε A/B.
2. Σχεδίασε `probA` ως προς το βήμα και πρόσθεσε οριζόντια γραμμή στο 0.5.

```

set.seed(12)
draws <- urnModel(n = 2500)

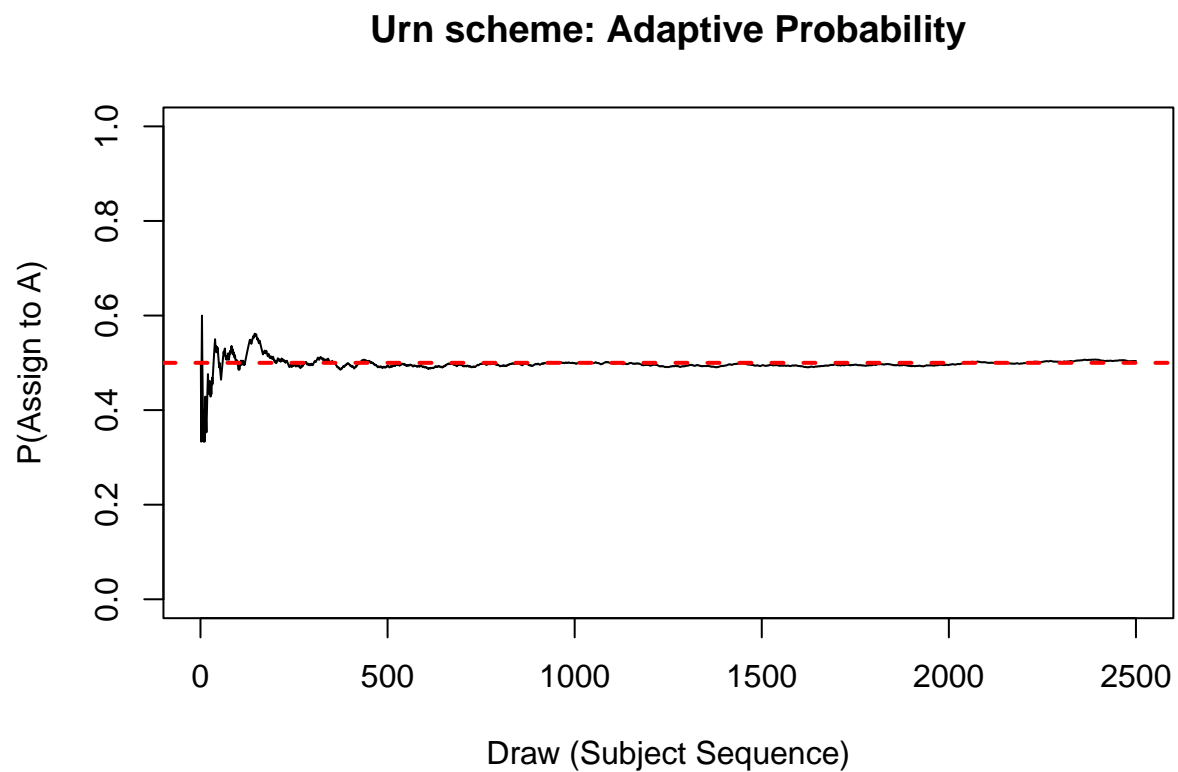
table(draws$treatment)

```

A	B
1241	1259

```
plot(draws$probA, type = "l",  
      xlab = "Draw (Subject Sequence)",  
      ylab = "P(Assign to A)",  
      main = "Urn scheme: Adaptive Probability",  
      ylim = c(0,1))
```

```
# Προσθήκη κόκκινης γραμμής στο 0.5 για έμφαση  
abline(h = 0.5, lty = 2, col = "red", lwd = 2)
```



## Η έννοια της “Μνήμης”

Η σύγκριση των μεθόδων αναδεικνύει μια θεμελιώδη διαφορά:

### Memoryless (Simple Randomization):

- Κάθε ανάθεση είναι ανεξάρτητη
- Αν τύχουν 10 ‘Α’ στη σειρά, το 11ο έχει πάλι  $P(A) = 0.5$
- Απλό, αμερόληπτο
- Μπορεί να δώσει ανισορροπία σε μικρά  $N$

### With Memory (Urn Scheme):

- “Θυμάται” την ιστορία
- Αν υπάρξει ανισορροπία, **αυτόματα διορθώνει** την πορεία
- Εγγυάται καλή ισορροπία
- Ελαφρώς πιο πολύπλοκο

### Perfect Memory (Block Randomization):

- Εξαναγκάζει ισορροπία κάθε  $b$  ασθενείς
- Μέγιστη εγγύηση
- Μπορεί να είναι προβλέψιμο (αν το block size είναι γνωστό)

### Πότε να χρησιμοποιήσεις ποιο;

- $N > 200$ : Simple Randomization (αρκετά μεγάλο δείγμα)
- $N = 50-200$ : Block ή Urn (ανάλογα με logistics)
- $N < 50$ : Block με random block sizes (εγγυημένη ισορροπία)

## Άσκηση 5 – Σύγκριση SR vs Block vs Urn σε μικρό $N$

Θεώρησε μικρή μελέτη  $N = 14$  και επανάλαβε 200 φορές:

- SR: απλή τυχαιοποίηση
- BR: block randomization (τυχαία block sizes)
- Urn: urn scheme

Σύγκρινε μέση τιμή και τυπική απόκλιση της αναλογίας που παίρνει A.

```

set.seed(5)
N <- 14
it <- 200

results <- data.frame(
  SR = rep(NA_real_, it),
  BR = rep(NA_real_, it),
  urn = rep(NA_real_, it)
)

for (i in 1:it) {
  # SR
  trt_sr <- ifelse(runif(N) < 0.5, "A", "B")
  results$SR[i] <- mean(trt_sr == "A")

  # BR
  br <- blockrand(n = N, num.levels = 2, levels = c("A", "B"),
                 block.sizes = 1:4)[1:N, ]
  results$BR[i] <- mean(br$treatment == "A")

  # urn
  trt_urn <- urnModel(N)$treatment
  results$urn[i] <- mean(trt_urn == "A")
}

colMeans(results)

```

SR	BR	urn
0.5067857	0.5050000	0.5010714

```
apply(results, 2, sd)
```

SR	BR	urn
0.13273499	0.03889959	0.07794322

### Συμπέρασμα:

Παρατηρούμε τα εξής:

1. Οι **μέσες τιμές** (colMeans) είναι όλες κοντά στο 0.5, άρα και οι τρεις μέθοδοι είναι αμερόληπτες.
2. Η **τυπική απόκλιση** (SD) όμως διαφέρει σημαντικά:

- Το **BR (Block Randomization)** έχει τη μικρότερη SD (κοντά στο 0), διότι εγγυάται την ισορροπία.
- Η **SR (Simple Randomization)** έχει τη μεγαλύτερη SD, δείχνοντας ότι σε μικρά δείγματα ( $N = 14$ ) υπάρχει μεγάλο ρίσκο ανισορροπίας.
- Το **Urn** βρίσκεται κάπου ενδιάμεσα.