

Lab 2: Study designs

11/01/2026

Άσκηση 1 – Μελέτη Φάσης I: Εκτίμηση δόσης-στόχου μέσω λογιστικής παλινδρόμησης

Θεωρήστε μια κλινική δοκιμή Φάσης I για κυτταροτοξικό φάρμακο. Διερευνήθηκαν δέκα προκαθορισμένα επίπεδα δόσης. Κάθε ασθενής έλαβε ακριβώς μία δόση του φαρμάκου (one-dose-per-patient design) και καταγράφηκε αν εμφάνισε dose-limiting toxicity (DLT) εντός προκαθορισμένου χρονικού παραθύρου μετά τη χορήγηση.

Στο αρχείο `trialPh1.csv` υπάρχουν οι μεταβλητές:

- `id`: αναγνωριστικός αριθμός ασθενούς
- `dose`: χορηγηθείσα δόση (αριθμητική μεταβλητή)
- `toxicity`: εμφάνιση σοβαρής τοξικότητας (1 = ναι, 0 = όχι)

Για τους σκοπούς του παρόντος εργαστηρίου, ως δόση-στόχο $D_{0.30}$ ορίζουμε τη δόση για την οποία, βάσει του προσαρμοσμένου μοντέλου δόσης-τοξικότητας, η πιθανότητα DLT ισούται με 30%.

(α) Εισαγωγή δεδομένων και περιγραφικά

1. Εισάγετε τα δεδομένα στο R.
2. Εμφανίστε τις πρώτες 10 γραμμές και τη δομή του dataset.
3. Δώστε περιγραφικά της `dose` (ελάχιστο, μέσο, διάμεσο, μέγιστο) και πίνακα συχνοτήτων για την `toxicity`.
4. Σχολιάστε αν οι παρατηρήσεις της δόσης καλύπτουν “πυκνά” το εύρος ή αν συγκεντρώνονται σε λίγα επίπεδα.

```

# Read the data (ensure the working directory is set appropriately)
trialPh1 <- read.csv("trialPh1.csv")

# Display the first 10 observations
head(trialPh1, 10)

# Display the internal structure of the dataset
str(trialPh1)

# Descriptive statistics for dose
summary(trialPh1$dose)

# Frequency table for the toxicity outcome
table(trialPh1$toxicity)

```

(β) Μοντελοποίηση πιθανότητας τοξικότητας

5. Εκτιμήστε την πιθανότητα τοξικότητας ως συνάρτηση της χορηγούμενης δόσης, χρησιμοποιώντας κατάλληλο γενικευμένο γραμμικό μοντέλο (GLM).
6. Ποιες είναι οι εκτιμήσεις για τα $\hat{\beta}_0$ και $\hat{\beta}_1$;
7. Ερμηνεύστε το $\hat{\beta}_1$ καθώς και το $\exp(\hat{\beta}_1)$.

```

# Fit a logistic regression model for the dose-toxicity relationship
fit_glm <- glm(toxicity ~ log(dose),
               family = binomial,
               data = trialPh1)

# Model summary
summary(fit_glm)

# Extract regression coefficients
beta_hat <- coef(fit_glm)

# Odds ratio associated with a one-unit increase in log(dose)
exp(beta_hat[2])

```

(γ) Εκτίμηση της δόσης-στόχου $D_{0.30}$

8. Βάσει του μοντέλου που χρησιμοποιήσατε, ποια είναι η δόση-στόχος;

```
# Define Logit function
logit <- function(p) log(p / (1 - p))

# Extract regression coefficients
beta0 <- beta_hat[1]
beta1 <- beta_hat[2]

# Compute the target dose corresponding to 30% toxicity (D_0.30)
dose_30 <- exp( (logit(0.30) - beta0) / beta1 )
dose_30
```

(δ) Γραφική απεικόνιση

9. Απεικονίστε την καμπύλη πιθανότητας τοξικότητας από το μοντέλο ως προς το $\log(dose)$.

```
# Inverse logit function for visualization
inv_logit <- function(x) exp(x) / (1 + exp(x))

# Plot toxicity probability vs log dose
curve(inv_logit(beta0 + beta1*x),
      from = min(log(trialPh1$dose)),
      to = max(log(trialPh1$dose)),
      ylim = c(0,0.60),
      las = 1,
      xlab = "Logarithm of dose",
      ylab = "Probability of toxicity",
      main = "Estimated probability of toxicity \n from a logistic regression model")

# Add reference lines for 30% toxicity dose
lines(c(min(log(trialPh1$dose)), log(dose_30)),
      rep(0.3, 2),
      lty = 2, col = "red")
```

```
lines(rep(log(dose_30), 2), c(0, 0.3),  
      lty = 2, col = "red")
```

(ε) Ηθική αξιολόγηση

10. Σχολιάστε γιατί η προσέγγιση που ακολούθησαν οι ερευνητές δεν θεωρείται ηθικά αποδεκτή σε σύγχρονες κλινικές δοκιμές, αναφερόμενοι:

- στην έκθεση πολλών ασθενών σε δυνητικά υπερ-τοξικές δόσεις,
- στην αρχή ελαχιστοποίησης κινδύνου,
- και στον στόχο των σύγχρονων σχεδιασμών Φάσης I.

Άσκηση 2 – Κλασικός σχεδιασμός 3+3 (three-up / three-down)

Οι μελέτες Φάσης I εξετάζουν τη σχέση δόσης–τοξικότητας και στοχεύουν στον προσδιορισμό δόσης με προκαθορισμένη αποδεκτή πιθανότητα dose-limiting toxicity (DLT), συνήθως γύρω στο 20–33% ανά πρωτόκολλο. Οι αποφάσεις για κλιμάκωση ή τερματισμό της κλιμάκωσης λαμβάνονται διαδοχικά, με κύριο περιορισμό την ασφάλεια.

Ο κλασικός σχεδιασμός 3+3 (three-up/three-down) είναι αλγοριθμικός και rule-based. Στην άσκηση θα μελετήσουμε τον σχεδιασμό αυτό ως στοχαστική διαδικασία και θα υπολογίσουμε τα operating characteristics (OCs) του.

Υποθέστε ότι κάθε ασθενής εμφανίζει DLT ανεξάρτητα με πιθανότητα p_i στο επίπεδο δόσης i .

Θυμηθείτε ότι:

- Ξεκινάμε με 3 ασθενείς στη δόση i
- Αν 0/3 έχουν τοξικότητα → κλιμάκωση στη επόμενη δόση (νέα τριάδα ασθενών).
- Αν 1/3 έχουν τοξικότητα → εντάσσονται 3 επιπλέον ασθενείς στην ίδια δόση (σύνολο 6):
 - Αν 1/6 έχουν τοξικότητα → κλιμάκωση στην επόμενη δόση (νέα τριάδα ασθενών).
 - Αν $\geq 2/6$ έχουν τοξικότητα → τερματισμός κλιμάκωσης (η MTD ορίζεται ως η προηγούμενη δόση).
- Αν $\geq 2/3$ στην αρχική τριάδα έχουν τοξικότητα → τερματισμός κλιμάκωσης (η MTD ορίζεται ως η προηγούμενη δόση).

Θεωρήστε ότι η πιθανότητα τοξικότητας στην i -οστή δόση είναι p_i , ανεξάρτητη μεταξύ ασθενών. Συμβολίστε με $b(k; n, p)$ την πιθανότητα να παρατηρηθούν k τοξικότητες σε n ασθενείς (διωνυμική κατανομή).

(α) Θεωρητική απόδειξη

1. Δείξτε ότι η δεσμευμένη (conditional) πιθανότητα κλιμάκωσης πέρα από τη δόση i , δεδομένου ότι φτάσαμε στη δόση i , είναι:

$$P_i = b(0; 3, p_i) + b(1; 3, p_i) b(0; 3, p_i).$$

Αναλύστε διεξοδικά τα δύο σενάρια που οδηγούν σε κλιμάκωση και γράψτε αναλυτικά τις πιθανότητες.

2. Δείξτε ότι η μη δεσμευμένη (unconditional) πιθανότητα να ξεπεραστεί η δόση i είναι:

$$Q_i = \prod_{k=1}^i P_k.$$

(β) Υπολογιστικό μέρος στο R

3. Υπολογίστε αριθμητικά το P_i για $p_i = 0.20$ και $p_i = 0.30$, χρησιμοποιώντας τη συνάρτηση `dbinom`.

```
# Conditional probability of dose escalation (Pi) in the 3+3 design
p <- c(0.20, 0.30)

p0 <- dbinom(0, 3, p)
p1 <- dbinom(1, 3, p)

Pi <- p0 + p1 * p0
Pi
```

4. Υπολογίστε το Q_2 όταν $p_1 = 0.20$ και $p_2 = 0.30$.

```
# Unconditional probability of escalation beyond dose 2 (Q2)
Q2 <- Pi[1] * Pi[2]
Q2
```

5. Ερμηνεύστε το Q_2 : τι σημαίνει για την πιθανότητα να προχωρήσει η μελέτη σε ακόμη υψηλότερες δόσεις;

(γ) Γενική συνάρτηση λειτουργικών χαρακτηριστικών

6. Γράψτε μια συνάρτηση στο R που:

- δέχεται ως είσοδο ένα διάνυσμα πιθανοτήτων τοξικότητας p (ένα στοιχείο για κάθε δόση),
- υπολογίζει τα P_i , Q_i και OC_i για κάθε δόση,
- επιστρέφει `data.frame` με στήλες $dose$, p , P_i , Q_i , OC_i .

7. Χρησιμοποιώντας αυτή τη συνάρτηση, υπολογίστε την πιθανότητα να συνεχίσει η μελέτη πέρα από τη δεύτερη δόση, όταν:

- $p_1 = 0.20$ (πρώτη δόση),
- $p_2 = 0.30$ (δεύτερη δόση).

```
# Operating characteristics of the 3+3 dose-escalation design
OC_3plus3 <- function(p) {
  Pi <- dbinom(0, 3, p) + dbinom(1, 3, p) * dbinom(0, 3, p)
  Qi <- cumprod(Pi)
  OC <- 1 - Qi
  data.frame(dose = seq_along(p), p = p, Pi = Pi, Qi = Qi, OC = OC)
}

OC_3plus3(c(0.20, 0.30))
```

(δ) Operating characteristics του σχεδιασμού 3+3 σε δύο σενάρια

Θεωρήστε δύο σενάρια “αληθινών” πιθανοτήτων *dose-limiting toxicity* (DLT) σε 10 επίπεδα δόσης (Piantadosi, Fig. 12.1, pp. 343, 3rd edition):

- **Σενάριο A:** οι πιθανότητες DLT δίνονται από το διάνυσμα p_A
- **Σενάριο B:** οι πιθανότητες DLT δίνονται από το διάνυσμα p_B

```
p_A <- c(0.001, 0.01, 0.02, 0.05, 0.11, 0.19, 0.28, 0.39, 0.53, 0.68)
p_B <- seq(0.10, 0.55, by = 0.05)
```

8. Εφαρμόστε τη συνάρτηση `OC_3plus3()` που υλοποιήσατε στο προηγούμενο ερώτημα και υπολογίστε, για κάθε επίπεδο δόσης i , τις ποσότητες P_i , Q_i και $OC_i = 1 - Q_i$.

9. Να φτιάξετε γράφημα που να απεικονίζει τη σωρευτική πιθανότητα τερματισμού της κλιμάκωσης OC_i ως συνάρτηση του επιπέδου δόσης i , και για τα δύο σενάρια (A και B) στο ίδιο διάγραμμα.

- Άξονας x: επίπεδο δόσης
- Άξονας y: σωρευτική πιθανότητα τερματισμού (OC_i)

10. Υποθέτοντας ότι η βέλτιστη δόση είναι αυτή με περίπου 30% τοξικότητα, σχολιάστε:

- Σε ποιο περίπου επίπεδο δόσης τείνει να σταματά ο σχεδιασμός στο σενάριο A;
- Σε ποιο στο σενάριο B;
- Πόσο καλά εντοπίζει ο σχεδιασμός 3+3 τη βέλτιστη δόση σε κάθε σενάριο;

Άσκηση 3 – Μελέτη Φάσης II: Σχεδιασμός ενός σταδίου

Θεωρήστε μια μη συγκριτική (*single-arm*) μελέτη Φάσης II με πρωτεύον καταληκτικό το ποσοστό ανταπόκρισης p .

Θέλουμε να αξιολογήσουμε αν η νέα θεραπεία είναι αρκετά αποτελεσματική ώστε να αξίζει περαιτέρω ανάπτυξη, συγκριτικά με ένα ιστορικό επίπεδο αποτελεσματικότητας.

Δίνονται:

- $p_0 = 0.15$: μη αποδεκτό ποσοστό ανταπόκρισης,
- $p_1 = 0.40$: επιθυμητό/κλινικά ενδιαφέρον ποσοστό ανταπόκρισης,
- $n = 16$: μέγεθος δείγματος,
- $\alpha = 0.10$: μονόπλευρο επίπεδο σημαντικότητας.

Υποθέστε ότι ο αριθμός ανταποκρίσεων X ακολουθεί διωνυμική κατανομή,

$$X \sim \text{Bin}(n, p).$$

(α) Διατύπωση υποθέσεων

1. Διατυπώστε τις υποθέσεις H_0 και H_1 που αντιστοιχούν στο παραπάνω ερώτημα.
2. Περιγράψτε (ποιοτικά) τη λογική ενός κατάλληλου κανόνα απόφασης που βασίζεται στον αριθμό ανταποκρίσεων X . Αιτιολογήστε γιατί αναμένεται η απόφαση να εξαρτάται από “μεγάλες” τιμές του X .

(β) Προσδιορισμός κανόνα απόφασης (ακριβές διωνυμικό τεστ)

3. Βρείτε τον μικρότερο ακέραιο k^* τέτοιο ώστε ο κανόνας απόφασης “απόρριψη της H_0 ” αν $X > k^*$ να έχει σφάλμα τύπου I το πολύ $\alpha = 0.10$ όταν $p = p_0$.
4. Υπολογίστε το ακριβές επίπεδο σημαντικότητας α_{exact} που επιτυγχάνεται με τον κανόνα απόφασης που επιλέξατε και σχολιάστε τη σχέση του με το α (ονομαστικό vs ακριβές).

(γ) Ισχύς υπό το p_1

5. Υπολογίστε την ισχύ του τεστ όταν $p = p_1 = 0,40$.
6. Σχολιάστε αν ο σχεδιασμός με $n = 16$ πετυχαίνει ταυτόχρονα:

- έλεγχο του σφάλματος τύπου I κοντά στο 0.10,
- και επιθυμητή ισχύ περίπου 0.80.

(δ) Κριτική του μονοσταδιακού σχεδιασμού

7. Σχολιάστε συνοπτικά πλεονεκτήματα και περιορισμούς του μονοσταδιακού σχεδιασμού, αναφερόμενοι:

- στην απουσία μηχανισμού πρόωρης διακοπής λόγω ανεπαρκούς αποτελεσματικότητας,
- στην αναμενόμενη έκθεση ασθενών όταν η θεραπεία είναι αναποτελεσματική ($p \approx p_0$),
- και στο πώς αυτά συνδέονται με την ανάγκη για δύο-σταδίων σχεδιασμούς Φάσης II.

Άσκηση 4 – Μελέτη Φάσης II: Σχεδιασμός δύο σταδίων (Simon design)

Θεωρήστε ότι θέλουμε να σχεδιάσουμε μια κλινική μελέτη Φάσης II με δύο στάδια (Simon's two-stage design) για τη θεραπεία του προηγούμενου παραδείγματος, με:

- μη αποδεκτό ποσοστό ανταπόκρισης $p_0 = 0.15$,
- επιθυμητό ποσοστό ανταπόκρισης $p_1 = 0.40$,
- σφάλμα τύπου I $\alpha = 0.10$ (μονόπλευρο),
- σφάλμα τύπου II $\beta = 0.20$ (ισχύς 0.80).

1. Χρησιμοποιώντας τη συνάρτηση `ph2simon` από τη βιβλιοθήκη `clinfun`, προσδιορίστε τον βέλτιστο (optimal) και τον minimax σχεδιασμό δύο σταδίων.

```
# Simon's two-stage Phase II design (optimal and minimax)

# Install once if needed:
# install.packages("clinfun")
library(clinfun)

ph2simon(pu = 0.15, pa = 0.40,
         ep1 = 0.10, ep2 = 0.20,
         nmax = 100)
```

2. Ερμηνεύστε τις παραμέτρους του σχεδιασμού που προκύπτουν από το output του `ph2simon`:

- συνολικό μέγεθος δείγματος n ,
- μέγεθος πρώτου σταδίου n_1 ,
- κανόνα διακοπής μετά το πρώτο στάδιο (τιμές r_1),
- τελικό κανόνα απόφασης (τιμή r),
- αναμενόμενο μέγεθος δείγματος $EN(p_0)$,
- πιθανότητα πρόωρης διακοπής υπό H_0 (PET, Probability of Early Termination).

3. Συγκρίνετε τον δύο-σταδίων σχεδιασμό με τον μονοσταδιακό σχεδιασμό της Άσκησης 3 ως προς:

- συνολικό μέγεθος δείγματος,
- αναμενόμενο μέγεθος δείγματος υπό H_0 ,
- δεοντολογικά πλεονεκτήματα (πρόωρη διακοπή όταν η θεραπεία είναι μη αποτελεσματική).

Άσκηση 5 – Cross-over trial

Σε μια cross-over μελέτη, διερευνήθηκαν οι επιδράσεις της pronethalol στη στηθάγχη. Δώδεκα ασθενείς έλαβαν placebo και pronethalol σε δύο διαδοχικές περιόδους ίσης διάρκειας, με τυχαιοποιημένη σειρά. Καταγράφηκε ο αριθμός επεισοδίων στηθάγχης σε κάθε περίοδο.

Τα δεδομένα περιλαμβάνουν:

- patient: αναγνωριστικό ασθενούς
- placebo: αριθμός επεισοδίων στηθάγχης υπό placebo
- drug: αριθμός επεισοδίων στηθάγχης υπό pronethalol
- diff = placebo - drug: διαφορά επεισοδίων (θετική τιμή = βελτίωση με το φάρμακο)

(α) Εισαγωγή δεδομένων και περιγραφικά

1. Εισάγετε τα δεδομένα στην R και δώστε περιγραφικά για placebo, drug και diff.

```
# Creating the dataset directly
angina <- data.frame(
  patient = 1:12,
  placebo = c(71, 323, 8, 14, 23, 34, 79, 60, 2, 3, 17, 7),
  drug = c(29, 348, 1, 7, 16, 25, 65, 41, 0, 0, 15, 2)
)
angina$diff <- angina$placebo - angina$drug
head(angina)
```

```
##   patient placebo drug diff
## 1      1      71   29   42
## 2      2     323  348  -25
## 3      3       8    1    7
## 4      4      14    7    7
## 5      5      23   16    7
## 6      6      34   25    9
```

2. Ελέγξτε την κατανομή της diff γραφικά (ιστόγραμμα, boxplot, Q-Q plot) και σχολιάστε αν η υπόθεση κανονικότητας είναι εύλογη.

```
summary(angina[, c("placebo", "drug", "diff")])
```

##	placebo	drug	diff
##	Min. : 2.00	Min. : 0.00	Min. : -25.000
##	1st Qu.: 7.75	1st Qu.: 1.75	1st Qu.: 2.750
##	Median : 20.00	Median : 15.50	Median : 7.000
##	Mean : 53.42	Mean : 45.75	Mean : 7.667
##	3rd Qu.: 62.75	3rd Qu.: 32.00	3rd Qu.: 10.250
##	Max. : 323.00	Max. : 348.00	Max. : 42.000

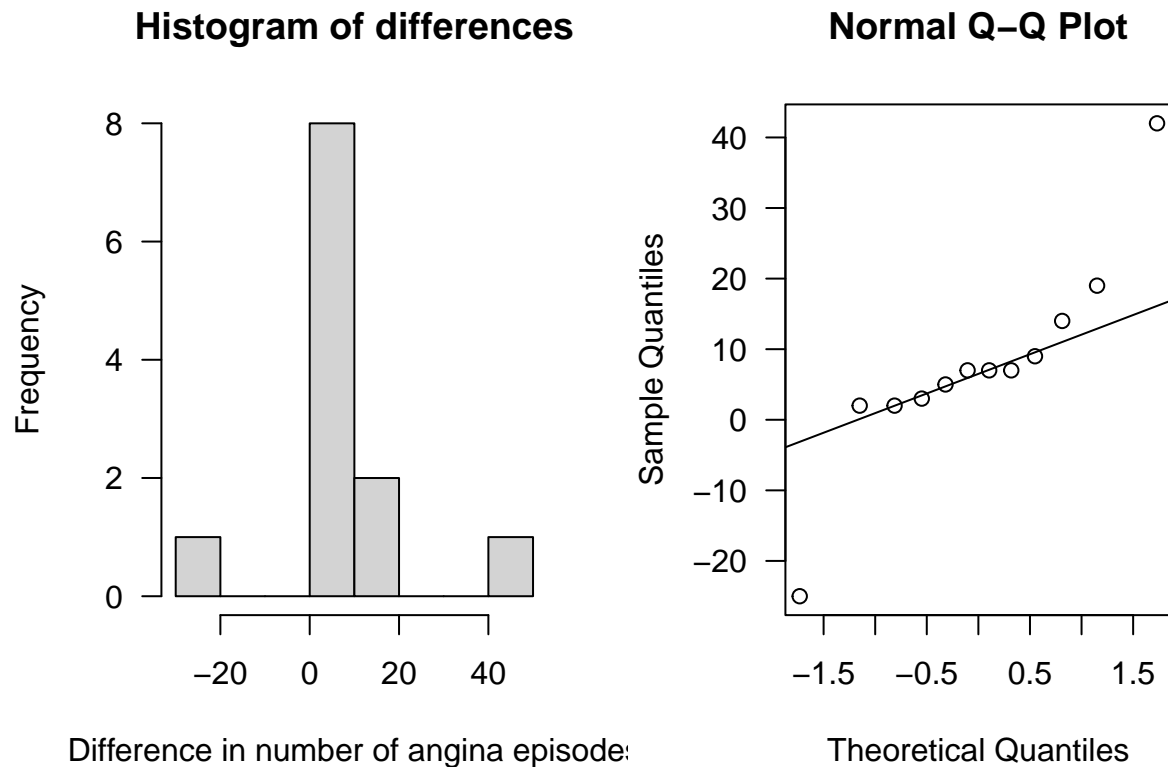
```
par(mfrow = c(1, 2))
```

```
# Histogram with descriptive x-axis and horizontal y-axis labels
```

```
hist(angina$diff,  
      main = "Histogram of differences",  
      xlab = "Difference in number of angina episodes",  
      las = 1)
```

```
# Q-Q plot with horizontal y-axis labels
```

```
qqnorm(angina$diff, las = 1)  
qqline(angina$diff)
```



```
par(mfrow = c(1, 1))
```

(β) «Αφελής» ανάλυση ως *parallel-group* trial

3. Για λόγους σύγκρισης, θεωρήστε (λανθασμένα) ότι τα δεδομένα προέρχονται από έναν *parallel-group* trial με δύο ανεξάρτητες ομάδες (placebo vs drug), δηλαδή αγνοήστε τη δομή cross-over / matched pairs.

- Εφαρμόστε κατάλληλο έλεγχο σύγκρισης μέσων (two-sample t-test),
- Καταγράψτε την εκτίμηση της διαφοράς και το p-value.

Σχολιάστε αν φαίνεται στατιστικά σημαντική διαφορά με αυτή την (λανθασμένη) προσέγγιση.

(γ) Σωστή ανάλυση: *paired (matched) design*

4. Αναλύστε τα δεδομένα λαμβάνοντας υπόψη το matched cross-over design:

- Εφαρμόστε **paired t-test** στη μεταβλητή diff.
- Ελέγξτε την εναλλακτική υπόθεση ότι **το φάρμακο μειώνει τα επεισόδια** (δηλαδή ότι η διαφορά placebo - drug είναι θετική → μονόπλευρος έλεγχος / alternative = "greater").
- Δώστε την εκτίμηση της μέσης διαφοράς, το διάστημα εμπιστοσύνης και το p-value.

5. Συγκρίνετε τα αποτελέσματα της paired και της «αφελούς» ανάλυσης.

(δ) Ερμηνεία και προϋποθέσεις εγκυρότητας

6. Σχολιάστε συνοπτικά:

- Γιατί η paired ανάλυση είναι προτιμητέα σε cross-over μελέτες (σύνδεση με τη μείωση της ενδο-ατομικής μεταβλητότητας σε σχέση με *parallel-group* σχεδιασμό),
- Ποιες πρόσθετες υποθέσεις πρέπει να ισχύουν για να είναι αξιόπιστα τα αποτελέσματα (σταθερότητα νόσου, carryover effects, επαρκές washout period, dropout μεταξύ περιόδων κ.λπ.).