

Lab 2: Solutions

11/01/2026

Άσκηση 1 – Μελέτη Φάσης I (Εύρεση βέλτιστης δόσης)

1. Εισάγετε τα δεδομένα στο R και εξετάστε βασικά descriptives για τη δόση και την τοξικότητα.

```
# Read the data
trialPh1 <- read.csv("trialPh1.csv")

# Display the first 10 rows of the data
head(trialPh1, 10)
```

```
##      id toxicity      dose
## 1    1         0 0.2500000
## 2    2         0 0.2750000
## 3    3         0 0.3025000
## 4    4         0 0.3327500
## 5    5         0 0.3660250
## 6    6         1 0.4026275
## 7    7         1 0.4428902
## 8    8         1 0.4871793
## 9    9         1 0.5358972
## 10 10         1 0.5894869
```

```
# Show the internal structure of the dataset:
str(trialPh1)
```

```
## 'data.frame':    2000 obs. of  3 variables:
## $ id          : int  1 2 3 4 5 6 7 8 9 10 ...
## $ toxicity: int  0 0 0 0 0 1 1 1 1 1 ...
## $ dose       : num  0.25 0.275 0.302 0.333 0.366 ...
```

```
# Produce a frequency table of the toxicity variable  
table(trialPh1$toxicity)
```

```
##  
##      0      1  
## 1582  418
```

2. Εκτιμήστε την πιθανότητα τοξικότητας ως συνάρτηση της χορηγούμενης δόσης, χρησιμοποιώντας κατάλληλο γενικευμένο γραμμικό μοντέλο (GLM) με δυαδικό outcome.

Εφαρμόζουμε logistic regression με λογαριθμική μετατροπή της δόσης:

```
# Perform logistic regression model  
fit_glm <- glm(toxicity ~ log(dose),  
               family = binomial,  
               data = trialPh1)  
  
# Show model summary  
summary(fit_glm)
```

```
##
## Call:
## glm(formula = toxicity ~ log(dose), family = binomial, data = trialPh1)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.9125     0.2341   12.44  <2e-16 ***
## log(dose)     4.9316     0.2890   17.07  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 2050.5  on 1999  degrees of freedom
## Residual deviance: 1633.7  on 1998  degrees of freedom
## AIC: 1637.7
##
## Number of Fisher Scoring iterations: 5
```

```
# Extract coefficients
beta_hat = coef(fit_glm)

# Exponentiate coefficient for log(dose)
exp(beta_hat)[2]
```

```
## log(dose)
##  138.6054
```

Οι εκτιμημένοι συντελεστές β_0, β_1 είναι θετικοί, υποδηλώνοντας αύξηση της πιθανότητας τοξικότητας με την αύξηση της δόσης.

3. Βάσει του μοντέλου που χρησιμοποιήσατε, ποια είναι η βέλτιστη δόση;

Ορίστε ως βέλτιστη δόση (Maximum Tolerated Dose, MTD) τη δόση που αντιστοιχεί σε πιθανότητα τοξικότητας $p = 0.30$.

Av

$$\text{logit}(p) = \beta_0 + \beta_1 \log(D),$$

τότε, για δεδομένο p ,

$$D = \exp \left\{ \frac{\text{logit}(p) - \beta_0}{\beta_1} \right\}.$$

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

# Calculate dose associated with 30% toxicity
dose_30 <- exp( (logit(0.30) - beta_hat[1]) / beta_hat[2] )
dose_30
```

```
## (Intercept)
##      0.466553
```

Απεικόνιση της εκτιμώμενης πιθανότητας τοξικότητας:

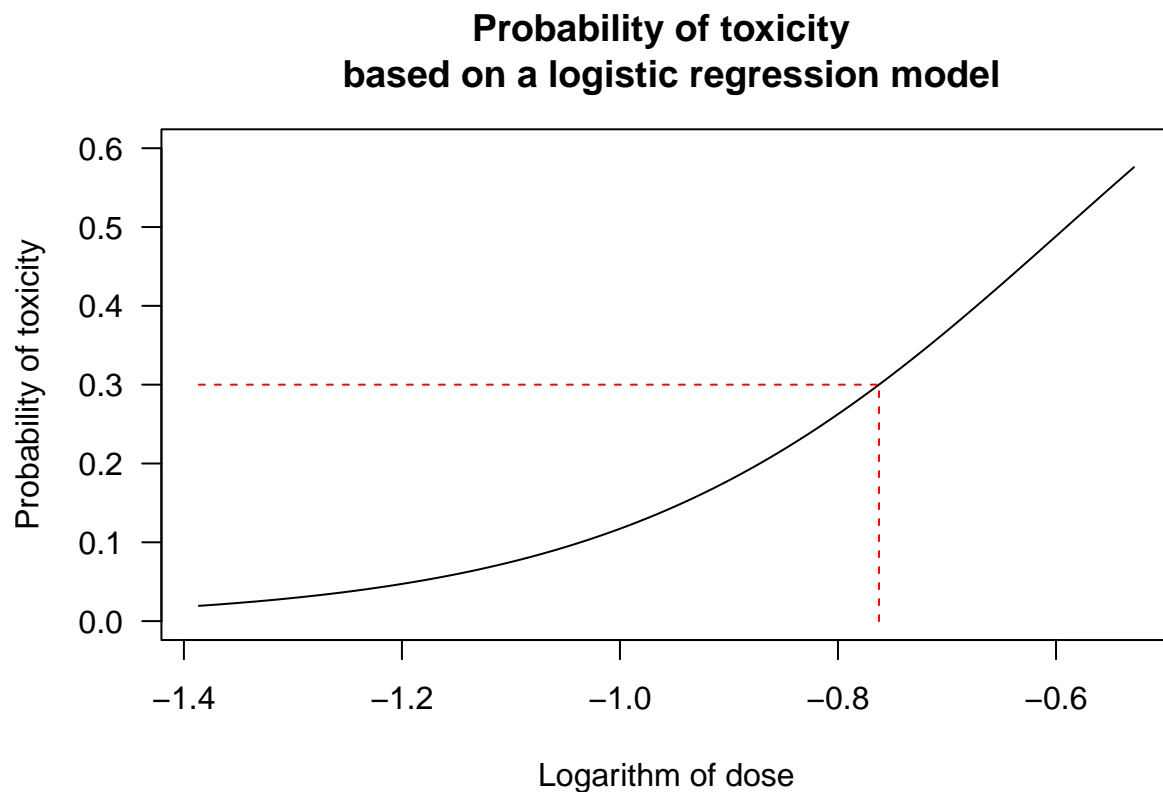
```
# Inverse logit for plotting
inv_logit <- function(x) exp(x) / (1 + exp(x))

# Define parameters for clearer code
beta0 <- beta_hat[1]
beta1 <- beta_hat[2]

# 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 = "Probability of toxicity \n based on 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")
```



5. Η προσέγγιση αυτή:

- εκθέτει μεγάλο αριθμό ασθενών σε υψηλές δόσεις με σημαντικό κίνδυνο σοβαρής τοξικότητας,
- δεν επιβάλλει μηχανισμούς έγκαιρης διακοπής,
- δεν είναι συμβατή με τις σύγχρονες αρχές δεοντολογίας που απαιτούν ελαχιστοποίηση του κινδύνου και προσεκτική κλιμάκωση δόσης.

Σύγχρονα designs (π.χ. 3+3, CRM) στοχεύουν στη σταδιακή αύξηση της δόσης και στην έκθεση όσο το δυνατόν λιγότερων ασθενών σε μη ασφαλείς ή μη αποτελεσματικές δόσεις.

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

(α) Θεωρητική απόδειξη

1.

- Σενάριο 1: 0 τοξικότητες σε 3 ασθενείς: $b(0; 3, p_i)$.
- Σενάριο 2: 1 τοξικότητα στους πρώτους 3 **και** 0 τοξικότητες στους επιπλέον 3:

$$b(1; 3, p_i) \times b(0; 3, p_i).$$

Άρα:

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

2. Για να συνεχίσει η δοκιμή πέρα από την i -οστή δόση, πρέπει να έχει περάσει επιτυχώς όλες τις προηγούμενες, οπότε:

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

Η λειτουργική χαρακτηριστική (OC) στην ή πριν από την i -οστή δόση είναι:

$$OC_i = 1 - Q_i.$$

(β) Υπολογιστικό μέρος στο R

3.

```
OC_3plus3 <- function(p, doses = NULL) {  
  if (is.null(doses)) {  
    doses <- paste0("Dose ", seq_along(p))  
  }  
  Pi <- dbinom(0, size = 3, prob = p) +  
    dbinom(1, size = 3, prob = p) * dbinom(0, size = 3, prob = p)
```

```

Qi <- cumprod(Pi)
OC <- 1 - Qi

data.frame(
  dose = doses,
  p     = p,
  Pi    = Pi,
  Qi    = Qi,
  OC    = OC
)
}

```

Παράδειγμα με $p_1 = 0.20$, $p_2 = 0.30$:

```

res_2doses <- OC_3plus3(c(0.20, 0.30))
res_2doses

```

##	dose	p	Pi	Qi	OC
## 1	Dose 1	0.2	0.708608	0.7086080	0.2913920
## 2	Dose 2	0.3	0.494263	0.3502387	0.6497613

Η πιθανότητα συνέχισης πέρα από τη δεύτερη δόση είναι Q_2 , ενώ η OC (πιθανότητα διακοπής μέχρι και τη δεύτερη δόση) είναι $1 - Q_2$.

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

```

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)

res_A <- OC_3plus3(p_A, doses = paste0("Dose ", 1:10))
res_B <- OC_3plus3(p_B, doses = paste0("Dose ", 1:10))

res_A

```

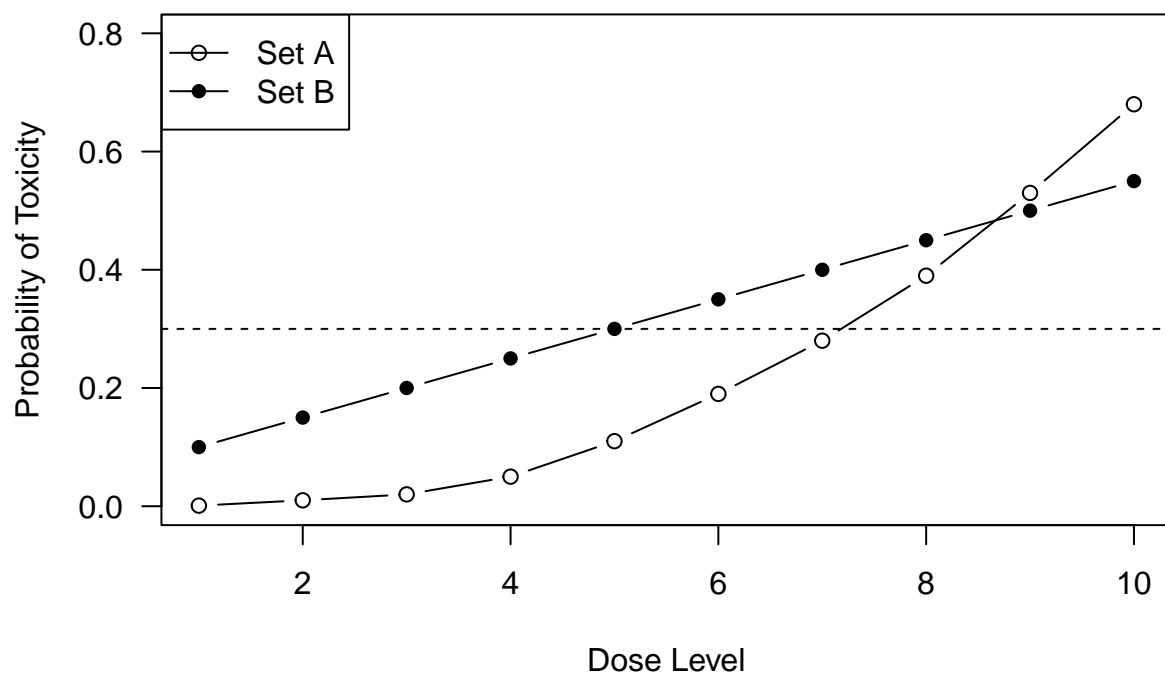
##	dose	p	Pi	Qi	OC
## 1	Dose 1	0.001	0.9999880	0.9999880290	1.197103e-05
## 2	Dose 2	0.010	0.9988287	0.9988167445	1.183256e-03
## 3	Dose 3	0.020	0.9954272	0.9942494030	5.750597e-03
## 4	Dose 4	0.050	0.9734421	0.9678442672	3.215573e-02
## 5	Dose 5	0.110	0.8892430	0.8606487027	1.393513e-01
## 6	Dose 6	0.190	0.7301877	0.6284351061	3.715649e-01
## 7	Dose 7	0.280	0.5357811	0.3367036405	6.632964e-01
## 8	Dose 8	0.390	0.3257988	0.1096976310	8.903024e-01
## 9	Dose 9	0.530	0.1402889	0.0153893552	9.846106e-01
## 10	Dose 10	0.680	0.0396131	0.0006096201	9.993904e-01

res_B

##	dose	p	Pi	Qi	OC
## 1	Dose 1	0.10	0.9061470	9.061470e-01	0.0938530
## 2	Dose 2	0.15	0.8137924	7.374155e-01	0.2625845
## 3	Dose 3	0.20	0.7086080	5.225385e-01	0.4774615
## 4	Dose 4	0.25	0.5998535	3.134466e-01	0.6865534
## 5	Dose 5	0.30	0.4942630	1.549250e-01	0.8450750
## 6	Dose 6	0.35	0.3964555	6.142089e-02	0.9385791
## 7	Dose 7	0.40	0.3093120	1.899822e-02	0.9810018
## 8	Dose 8	0.45	0.2343184	4.451632e-03	0.9955484
## 9	Dose 9	0.50	0.1718750	7.651242e-04	0.9992349
## 10	Dose 10	0.55	0.1215721	9.301779e-05	0.9999070

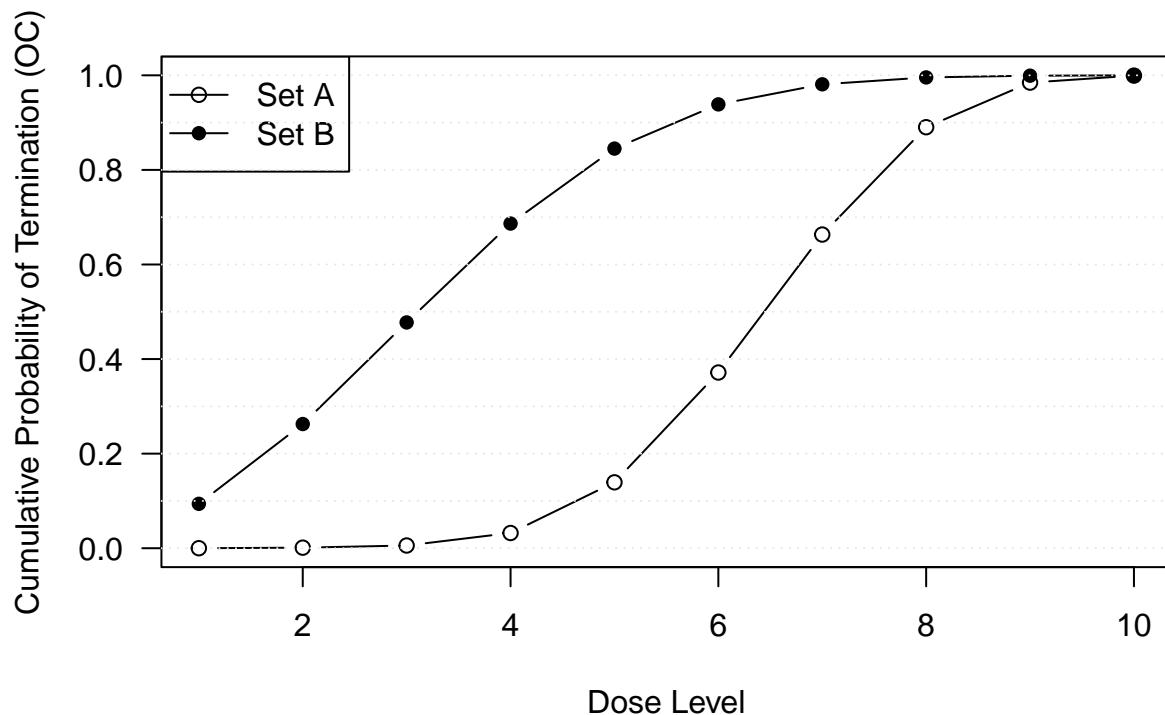
Πιθανότητες τοξικότητας:

```
plot(1:10, p_A, type = "b", pch = 21,
     ylim = c(0, 0.8),
     xlab = "Dose Level",
     ylab = "Probability of Toxicity",
     las = 1)
lines(1:10, p_B, type = "b", pch = 16)
legend("topleft", legend = c("Set A", "Set B"),
     pch = c(21, 16), lty = 1)
abline(h = 0.30, lty = 2)
```

Λειτουργικές χαρακτηριστικές:

```
plot(1:10, res_A$OC, type = "b", pch = 21,
     ylim = c(0, 1),
     xlab = "Dose Level",
     ylab = "Cumulative Probability of Termination (OC)",
     las = 1)
lines(1:10, res_B$OC, type = "b", pch = 16)
legend("topleft", legend = c("Set A", "Set B"),
      pch = c(21, 16), lty = 1)
abline(h = seq(0, 1, by = 0.1), col = "gray90", lty = "dotted")
```



Σχόλιο:

- **Σενάριο Α:** Η καμπύλη τοξικότητας είναι απότομη στα υψηλά επίπεδα. Η μελέτη τείνει να σταματά μεταξύ της 6ης και 7ης δόσης. Ο σχεδιασμός έχει σχεδόν 90% πιθανότητα να σταματήσει πριν ή στην 8η δόση.
- **Σενάριο Β:** Οι πιθανότητες τοξικότητας αυξάνονται γραμμικά. Εδώ φαίνεται η αδυναμία του σχεδιασμού:
 - Η βέλτιστη δόση (MTD ~30%) είναι η **5η δόση**.
 - Όμως, το αναμενόμενο σημείο διακοπής είναι η **3η δόση** (όπου $p \approx 0.20$).
 - Ο σχεδιασμός έχει **85% πιθανότητα να σταματήσει ΠΡΙΝ φτάσει στη βέλτιστη δόση** (πριν την 5η).

Συμπέρασμα: Το παράδειγμα Β δείχνει ότι ο σχεδιασμός 3+3 είναι συχνά **υπερβολικά συντηρητικός**: έχει την τάση να σταματάει νωρίς, “υπο-θεραπεύοντας” τους ασθενείς και αποτυγχάνοντας να εντοπίσει την πραγματική MTD όταν η καμπύλη δόσης-τοξικότητας είναι ρηχή.

Άσκηση 3 – Φάση II, μονοσταδιακός σχεδιασμός

3.1 Περιοχή απόρριψης

```
n <- 16
p0 <- 0.15
p1 <- 0.40

k      <- 0:n
p_tail0 <- 1 - pbinom(k, size = n, prob = p0) #  $P(X > k \mid p_0)$ 
p_tail1 <- 1 - pbinom(k, size = n, prob = p1) #  $P(X > k \mid p_1)$ 

bintab <- cbind(
  k = k,
  "P(X>k|p0)" = round(p_tail0, 5),
  "P(X>k|p1)" = round(p_tail1, 5)
)
bintab
```

##		k	P(X>k p0)	P(X>k p1)
##	[1,]	0	0.92575	0.99972
##	[2,]	1	0.71610	0.99671
##	[3,]	2	0.43862	0.98166
##	[4,]	3	0.21011	0.93485
##	[5,]	4	0.07905	0.83343
##	[6,]	5	0.02354	0.67116
##	[7,]	6	0.00559	0.47283
##	[8,]	7	0.00106	0.28394
##	[9,]	8	0.00016	0.14227
##	[10,]	9	0.00002	0.05832
##	[11,]	10	0.00000	0.01914
##	[12,]	11	0.00000	0.00490
##	[13,]	12	0.00000	0.00094
##	[14,]	13	0.00000	0.00013
##	[15,]	14	0.00000	0.00001
##	[16,]	15	0.00000	0.00000
##	[17,]	16	0.00000	0.00000

Βρίσκουμε τον μικρότερο k^* με $P(X > k^* | p_0) \leq 0.10$:

```
idx <- which(p_tail0 <= 0.10)[1]
k_star <- k[idx]
k_star
```

```
## [1] 4
```

```
p_tail0[idx]
```

```
## [1] 0.0790513
```

Άρα η περιοχή απόρριψης είναι: απόρριψη H_0 αν $X \geq 5$.

3.2 Ακριβές α και ισχύς

```
alpha_exact <- p_tail0[idx]
power_exact <- p_tail1[idx]
```

```
alpha_exact
```

```
## [1] 0.0790513
```

```
power_exact
```

```
## [1] 0.8334326
```

Τυπικά:

- $\alpha_{\text{exact}} \approx 0.079$
- Power ≈ 0.83

Έλεγχος με `binom.test`:

```
binom.test(5, 16, p = p0, alternative = "greater")
```

```
##
## Exact binomial test
##
## data: 5 and 16
## number of successes = 5, number of trials = 16, p-value = 0.07905
## alternative hypothesis: true probability of success is greater than 0.15
## 95 percent confidence interval:
## 0.132111 1.000000
## sample estimates:
## probability of success
## 0.3125
```

Άσκηση 4 – Simon’s two-stage design

4.1 Υπολογισμός σχεδιασμού

```
# install.packages("clinfun")
library(clinfun)

p0 <- 0.15
p1 <- 0.40
alpha <- 0.10
beta <- 0.20

sim_design <- ph2simon(pu = p0, pa = p1,
                      ep1 = alpha, ep2 = beta,
                      nmax = 100)

sim_design
```

```
##
## Simon 2-stage Phase II design
##
## Unacceptable response rate: 0.15
## Desirable response rate: 0.4
## Error rates: alpha = 0.1 ; beta = 0.2
##
##      r1 n1 r  n EN(p0) PET(p0)  qLo  qHi
## Minimax 1  9 4 16 11.80 0.5995 0.457 1.000
## Optimal 1  7 4 18 10.12 0.7166 0.000 0.457
```

Τυπικά αποτελέσματα:

- **Optimal:** $n_1 = 7, n = 18, r_1 = 1, r = 4, EN(p_0) \approx 10.1$.
- **Minimax:** $n_1 = 9, n = 16, r_1 = 1, r = 4, EN(p_0) \approx 11.8$.

4.2 Ερμηνεία και Σύγκριση Σχεδιασμών

Ελέγχουμε τις υποθέσεις:

$$H_0 : p \leq 0.15 \quad (\text{αναποτελεσματικό}) \quad \text{vs} \quad H_1 : p \geq 0.40 \quad (\text{αποτελεσματικό})$$

Ανάλυση του Minimax σχεδιασμού:

Το Minimax design ελαχιστοποιεί το **μέγιστο** συνολικό δείγμα (n).

- **Στάδιο 1:** Εγγράφονται $n_1 = 9$ ασθενείς.
 - $An \leq 1$ ανταποκριθεί \rightarrow **Stop for futility** (το φάρμακο δεν δουλεύει).
 - $An \geq 2$ ανταποκριθούν \rightarrow Συνέχεια στο Στάδιο 2.
- **Στάδιο 2:** Εγγράφονται επιπλέον 7 ασθενείς (σύνολο $n = 16$).
 - An στο σύνολο ≤ 4 ανταποκριθούν \rightarrow Αποτυχία μελέτης (μη απόρριψη H_0).
 - $An \geq 5$ ανταποκριθούν \rightarrow Επιτυχία (απόρριψη H_0).

Ερμηνεία δεικτών ασφαλείας:

- $PET(p_0) \approx 0.60$: Υπάρχει 60% πιθανότητα η μελέτη να σταματήσει νωρίς (στους 9 ασθενείς) αν το φάρμακο είναι πράγματι αναποτελεσματικό.
- $EN(p_0) \approx 11.8$: Αν το φάρμακο δεν δουλεύει, αναμένεται να εκτεθούν κατά μέσο όρο 11.8 ασθενείς.

Σύγκριση Optimal vs Minimax (To Trade-off):

Χαρακτηριστικό	Optimal Design	Minimax Design
Στόχος	Ελαχιστοποιεί τον αναμενόμενο αριθμό ασθενών (EN) αν ισχύει η H_0 .	Ελαχιστοποιεί τον μέγιστο αριθμό ασθενών (n) στο χειρότερο σενάριο.
Πλεονέκτημα	Καλύτερη προστασία ασθενών από αναποτελεσματική θεραπεία (χαμηλότερο EN).	Μικρότερο δείγμα αν η μελέτη φτάσει μέχρι το τέλος (“σφιχτό” design).

Χαρακτηριστικό	Optimal Design	Minimax Design
Στο παράδειγμα	Σταματάει πιο εύκολα νωρίς ($PET \approx 0.72$), άρα εκθέτει λιγότερους ασθενείς κατά μέσο όρο ($EN \approx 10$).	Χρειάζεται το μικρότερο δυνατό συνολικό δείγμα ($n = 16$ έναντι 18), αλλά αργεί λίγο περισσότερο να σταματήσει νωρίς.

Συμπέρασμα: Στην πράξη, συχνά προτιμάται το **Optimal** για ηθικούς λόγους (ελαχιστοποίηση έκθεσης σε μη δραστικό φάρμακο), εκτός αν υπάρχουν αυστηροί περιορισμοί στη διαθεσιμότητα ασθενών ή στο budget, οπότε επιλέγεται το **Minimax**.

Άσκηση 5 – Cross-over trial για στηθάγχη

Ορίζουμε τα δεδομένα όπως στον πίνακα του παραδείγματος:

```
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
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
## 7	7	79	65	14
## 8	8	60	41	19
## 9	9	2	0	2
## 10	10	3	0	3
## 11	11	17	15	2
## 12	12	7	2	5

5.1 Περιγραφικά & κανονικότητα

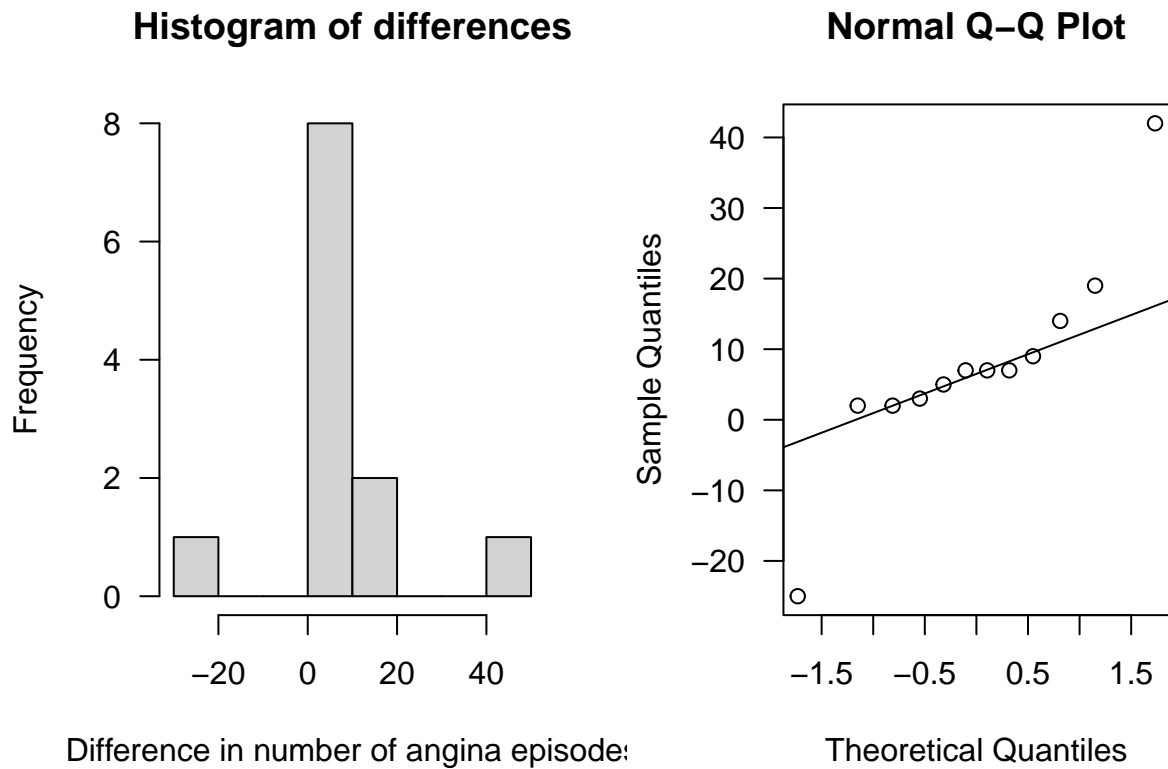
```
summary(angina[, c("placebo", "drug", "diff")])
```

##	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))
```

Η `diff` είναι κυρίως θετική, υποδηλώνοντας μείωση επεισοδίων με το φάρμακο. Η κατανομή είναι σχετικά συμμετρική, αλλά με μικρό n .

5.2 Αφελής two-sample ανάλυση

```
t.test(angina$placebo, angina$drug,
       alternative = "two.sided",
       paired = FALSE)
```

```
##
## Welch Two Sample t-test
##
## data:  angina$placebo and angina$drug
## t = 0.20146, df = 21.835, p-value = 0.8422
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -71.28941  86.62275
## sample estimates:
## mean of x mean of y
##  53.41667  45.75000
```

Η ανάλυση θεωρεί τα δύο δείγματα ανεξάρτητα, αγνοώντας τη συσχέτιση εντός ασθενούς. Το p-value είναι συνήθως μεγαλύτερο και τα διαστήματα εμπιστοσύνης πιο φαρδιά.

5.3 Σωστή paired ανάλυση

```
t.test(angina$placebo, angina$drug,
       alternative = "greater",
       paired = TRUE)
```

```
##
## Paired t-test
##
## data:  angina$placebo and angina$drug
## t = 1.7579, df = 11, p-value = 0.05326
## alternative hypothesis: true mean difference is greater than 0
## 95 percent confidence interval:
##  -0.1655743      Inf
## sample estimates:
## mean difference
##      7.666667
```

ή:

```
t.test(angina$diff,  
       alternative = "greater",  
       mu = 0)
```

```
##  
## One Sample t-test  
##  
## data:  angina$diff  
## t = 1.7579, df = 11, p-value = 0.05326  
## alternative hypothesis: true mean is greater than 0  
## 95 percent confidence interval:  
## -0.1655743      Inf  
## sample estimates:  
## mean of x  
##  7.666667
```

Εκτίμηση μέσης διαφοράς:

```
mean(angina$diff)
```

```
## [1] 7.666667
```

```
sd(angina$diff)
```

```
## [1] 15.10769
```

Η μέση διαφορά είναι θετική (περίπου 7–8 επεισόδια) και το p-value είναι οριακό (0.053). Παρότι τυπικά > 0.05 , υπάρχει ισχυρή ένδειξη κλινικής βελτίωσης που χανόταν τελείως στην unpaired ανάλυση ($p=0.84$).

Σύγκριση:

- Το paired t-test εκμεταλλεύεται τη θετική συσχέτιση μεταξύ μετρήσεων στον ίδιο ασθενή και έχει μεγαλύτερη ισχύ.
- Η αφελής two-sample ανάλυση υποτιμά τη δύναμη του σχεδιασμού και δίνει λιγότερο ακριβή διαστήματα εμπιστοσύνης.

Υποθέσεις cross-over:

- σταθερή ένταση νόσου κατά τη διάρκεια της μελέτης,
- απουσία ή ελάχιστα carryover effects (επαρκής washout),
- θετική συσχέτιση εντός ασθενούς ($\rho > 0$),
- απουσία σημαντικής διαφοροποίησης λόγω περιόδου χορήγησης.

Η παραβίαση αυτών μπορεί να οδηγήσει σε μεροληπτικές ή αναξιόπιστες εκτιμήσεις.