

Lab 3: Power and Sample Size Calculations

11/01/2026

Στο εργαστήριο αυτό θα εξασκηθούμε σε υπολογισμούς ισχύος και μεγέθους δείγματος για διαφορετικούς τύπους κλινικών μελετών.

Άσκηση 1 – Μέγεθος δείγματος για προ-καθορισμένη ακρίβεια (precision)

Έστω ότι θέλεις να εκτιμήσεις τη μέση τιμή ενός πληθυσμού, μ . Ο προφανής τρόπος είναι να εκτιμήσεις την παράμετρο μ μέσω του δειγματικού μέσου \bar{X} .

Να βρεθεί το απαιτούμενο μέγεθος δείγματος ώστε το σφάλμα εκτίμησης της μέσης τιμής να βρίσκεται εντός $d = 0.5\sigma$ από την πραγματική τιμή της μέσης τιμής, με επίπεδο εμπιστοσύνης 95%, δηλαδή:

$$\Pr(|\bar{X} - \mu| < d) \geq 0.95.$$

Δηλαδή, υποθέτουμε γνωστή (ή καλά εκτιμημένη) τη διακύμανση σ^2 .

Υπόθεση/προσέγγιση: Υποθέστε ότι, για επαρκώς μεγάλο n ,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

Υπόδειξη

Θεωρητική προσέγγιση

1. Γράψτε την πιθανότητα $\Pr(|\bar{X} - \mu| < d)$ σε όρους της τυποποιημένης κανονικής.

2. Δείξτε ότι:

$$n \geq \left(\frac{\sigma z_{0.975}}{d} \right)^2,$$

και ότι για $d = 0.5\sigma$:

$$n \geq (2z_{0.975})^2.$$

3. Υπολογίστε αριθμητικά το n .

R

```
alpha_ex1 <- 0.05
z_ex1 <- qnorm(1 - alpha_ex1/2)
n_raw_ex1 <- (2 * z_ex1)^2
n_ex1 <- ceiling(n_raw_ex1)
n_ex1
```

Άσκηση 2 – Έλεγχος μίας μέσης τιμής (one-sample t-test) με `pwr.t.test`

Θεωρήστε ότι θέλουμε να ελέγξουμε αν η μέση τιμή μιας συνεχούς μεταβλητής (π.χ. επίπεδο βιοδείκτη) υπερβαίνει ένα κλινικά σημαντικό όριο $\mu_0 = 3$. Θεωρούμε ότι, αν η πραγματική μέση τιμή είναι $\mu_1 = 4$, τότε η θεραπεία/παρέμβαση θεωρείται κλινικά χρήσιμη.

Υποθέτουμε γνωστή (ή καλά εκτιμημένη από προηγούμενη μελέτη/pilot) κοινή τυπική απόκλιση $\sigma = 2$ ως planning value και one-sided έλεγχο με:

- επίπεδο σημαντικότητας $\alpha = 0.05$,
- στοχευόμενη ισχύ $1 - \beta = 0.90$.

Υπόδειξη

Θεωρητική προσέγγιση

1. Διατυπώστε τις υποθέσεις:

$$H_0 : \mu \leq \mu_0 \quad \text{έναντι} \quad H_1 : \mu > \mu_0.$$

2. Ορίστε το effect size (Cohen's d):

$$d = \frac{\mu_1 - \mu_0}{\sigma}.$$

3. Εξηγήστε γιατί στο planning χρησιμοποιούμε ως στόχο τη μ_1 (την πραγματική μέση τιμή υπό την εναλλακτική υπόθεση) και όχι τη μ_0 .

R

```
library(pwr)

mu0_ex2 <- 3
mu1_ex2 <- 4
sigma_ex2 <- 2
```

```

d_ex2 <- (mu1_ex2 - mu0_ex2) / sigma_ex2

res_ex2 <- pwr.t.test(
  d = d_ex2,
  sig.level = 0.05,
  power = 0.90,
  alternative = "greater",
  type = "one.sample"
)

n_ex2 <- ceiling(res_ex2$n)
n_ex2

# achieved power with rounded n
res_power_ex2 <- pwr.t.test(
  d = d_ex2,
  sig.level = 0.05,
  n = n_ex2,
  alternative = "greater",
  type = "one.sample"
)
res_power_ex2$power

```

Άσκηση 3 – Σύγκριση δύο ποσοστών με στόχο την ισχύ (two-arm design)

Θεωρήστε μια τυχαιοποιημένη μελέτη δύο παράλληλων ομάδων (1:1), όπου το πρωτεύον καταληκτικό είναι ένα δυαδικό outcome (π.χ. ανταπόκριση/μη ανταπόκριση).

Στην ομάδα ελέγχου αναμένουμε ποσοστό ανταπόκρισης $p_C = 0.20$ και στην πειραματική ομάδα στόχο $p_T = 0.30$.

Να βρεθεί το απαιτούμενο μέγεθος δείγματος (ανά ομάδα) ώστε να επιτευχθεί power 0.80 σε one-sided έλεγχο με $\alpha = 0.10$.

Υπόδειξη

Θεωρητική προσέγγιση

1. Διατυπώστε τον one-sided έλεγχο:

$$H_0 : p_T \leq p_C \quad \text{έναντι} \quad H_1 : p_T > p_C.$$

2. Υπολογίστε το Cohen's h :

$$h = 2 \arcsin(\sqrt{p_T}) - 2 \arcsin(\sqrt{p_C}).$$

Υπενθύμιση: θετικό h αντιστοιχεί σε μεγαλύτερο ποσοστό στην πειραματική ομάδα.

3. Χρησιμοποιήστε το h στο `pwr.2p.test` για να υπολογίσετε n ανά ομάδα.

R

```
library(pwr)

pC_ex3 <- 0.20
pT_ex3 <- 0.30
alpha_ex3 <- 0.10
power_ex3 <- 0.80
```

```

h_ex3 <- 2 * asin(sqrt(pT_ex3)) - 2 * asin(sqrt(pC_ex3))
# alternatively: h_ex3 <- ES.h(p1 = pT_ex3, p2 = pC_ex3)

res_ex3 <- pwr.2p.test(
  h = h_ex3,
  sig.level = alpha_ex3,
  power = power_ex3,
  alternative = "greater"
)

n_per_group_ex3 <- ceiling(res_ex3$n)
N_total_ex3 <- 2 * n_per_group_ex3

n_per_group_ex3
N_total_ex3

```

Άσκηση 4 – Μελέτη Φάσης II: Μονοσταδιακός (*single-stage*) σχεδιασμός (exact binomial)

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

Δίνονται:

- $p_0 = 0.15$: μη αποδεκτό/ιστορικό επίπεδο ανταπόκρισης,
- $p_1 = 0.40$: επιθυμητό/κλινικά ενδιαφέρον επίπεδο,
- $\alpha = 0.10$: μονόπλευρο σφάλμα τύπου I,
- $\beta = 0.20$: σφάλμα τύπου II (ισχύς $1 - \beta = 0.80$).

Υποθέστε ότι ο αριθμός ανταποκρίσεων X ακολουθεί:

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

Υπόδειξη

Θεωρητική προσέγγιση

1. Διατυπώστε τις υποθέσεις ως design problem:

$$H_0 : p \leq p_0 \quad \text{έναντι} \quad H_1 : p \geq p_1.$$

2. Περιγράψτε μονοσταδιακό κανόνα απόφασης:

- Απορρίπτουμε το H_0 αν $X > r$ (ισοδύναμα, αν $X \geq r + 1$), δηλαδή αν παρατηρήσουμε τουλάχιστον $r + 1$ ανταποκρίσεις σε n ασθενείς.

3. Θέστε τους περιορισμούς σχεδιασμού:

$$\Pr_{p_0}(X > r) \leq \alpha, \quad \Pr_{p_1}(X > r) \geq 1 - \beta.$$

R (Exact binomial design)

```
library(clinfun)

p0_ex4 <- 0.15
p1_ex4 <- 0.40
alpha_ex4 <- 0.10
beta_ex4 <- 0.20

designs_ex4 <- ph2single(
  pu = p0_ex4,
  pa = p1_ex4,
  ep1 = alpha_ex4,
  ep2 = beta_ex4,
  nsoln = 10
)
designs_ex4

# choose smallest n
design_small_ex4 <- designs_ex4[order(designs_ex4$n), ][1, ]
n_ex4 <- design_small_ex4$n
r_ex4 <- design_small_ex4$r
c(n = n_ex4, r = r_ex4)

# achieved alpha and achieved power for chosen (n,r)
alpha_ach_ex4 <- 1 - pbinom(r_ex4, size = n_ex4, prob = p0_ex4) #  $P(X \geq r+1)$ 
power_ach_ex4 <- 1 - pbinom(r_ex4, size = n_ex4, prob = p1_ex4) #  $P(X \geq r+1)$ 

c(alpha_ach = alpha_ach_ex4, power_ach = power_ach_ex4)
```

R (Approximation με arcsine – για σύγκριση)

```
library(pwr)
```



```
h_arc_ex4 <- ES.h(p1_ex4, p0_ex4)

res_arc_ex4 <- pwr.p.test(
  h = h_arc_ex4,
  sig.level = alpha_ex4,
  power = 1 - beta_ex4,
  alternative = "greater"
)
n_arc_ex4 <- ceiling(res_arc_ex4$n)
n_arc_ex4
```

Άσκηση 5 – Δύο ανεξάρτητες μέσες τιμές: αναπαραγωγή πίνακα μεγέθους δείγματος

Στόχος είναι να αναπαράγετε (σε μορφή πίνακα) το απαιτούμενο **συνολικό** μέγεθος δείγματος N για σύγκριση δύο ανεξάρτητων μέσων τιμών με:

- ίση κατανομή (1:1),
- δίπλευρο έλεγχο,
- power = 0.90,
- $\alpha \in \{0.05, 0.10\}$,
- effect sizes $d = 0.25, 0.50, \dots, 1.50$.

Ορισμός effect size (Cohen's d):

$$d = \frac{\mu_1 - \mu_2}{\sigma},$$

όπου σ είναι η κοινή τυπική απόκλιση (planning value).

Υπόδειξη

Θεωρητική προσέγγιση

1. Για κάθε d και κάθε α , χρησιμοποιήστε `pwr.t.test` για να υπολογίσετε το n ανά ομάδα.
2. Θέστε $N = 2\lceil n \rceil$.
3. Παρουσιάστε τα N σε πίνακα με γραμμές τα d και στήλες για $\alpha = 0.05$ και $\alpha = 0.10$.

R

```
library(pwr)
```

```
ES_ex5 <- seq(0.25, 1.50, by = 0.25)
```

```

power_ex5 <- 0.90

tab_ex5 <- data.frame(
  ES = ES_ex5,
  `Alpha = 0.05` = NA_integer_,
  `Alpha = 0.10` = NA_integer_
)

for (i in seq_along(ES_ex5)) {
  d_val <- ES_ex5[i]

  n05 <- pwr.t.test(d = d_val,
    sig.level = 0.05,
    power = power_ex5,
    type = "two.sample",
    alternative = "two.sided")$n

  n10 <- pwr.t.test(d = d_val,
    sig.level = 0.10,
    power = power_ex5,
    type = "two.sample",
    alternative = "two.sided")$n

  tab_ex5[i, "Alpha = 0.05"] <- 2 * ceiling(n05)
  tab_ex5[i, "Alpha = 0.10"] <- 2 * ceiling(n10)
}

tab_ex5

```

Άσκηση 6 – Unequal allocation και επιπτώσεις στο συνολικό δείγμα

Θεωρήστε σύγκριση δύο ανεξάρτητων μέσων τιμών (two-sample design) με:

- effect size $d = 0.5$,
- δίπλευρο έλεγχο $\alpha = 0.05$,
- $\beta = 0.10$ (ισχύς 0.90),
- κατανομή 1:k για $k = 1, \dots, 10$.

Θα μελετήσετε πώς επηρεάζεται το συνολικό sample size N όταν απομακρυνόμαστε από 1:1.

Θα χρησιμοποιηθεί normal approximation μέσω `gsDesign::nNormal` (z-approach).

Υπόδειξη

Θεωρητική προσέγγιση

1. Στη `gsDesign::nNormal`, το ratio είναι n_1/n_2 . Για κατανομή 1:k (δηλαδή $n_1 : n_2 = 1 : k$), περνάμε `ratio = 1/k`.
2. Σχολιάστε ποιο μέλος της διακύμανσης μεταβάλλεται με unequal allocation (ιδέα: $1/n_1 + 1/n_2$).

R

```
library(gsDesign)

d_ex6 <- 0.5
alpha_ex6 <- 0.05
beta_ex6 <- 0.10

k_vals_ex6 <- 1:10
```

```

ar_ex6 <- data.frame(
  k = k_vals_ex6,
  ratio = paste0("1:", k_vals_ex6),
  N = NA_integer_,
  n1 = NA_integer_,
  n2 = NA_integer_
)

for (i in seq_along(k_vals_ex6)) {
  k <- k_vals_ex6[i]

  N_total <- as.numeric(
    nNormal(delta1 = d_ex6, delta0 = 0,
      sd = 1, sd2 = 1,
      alpha = alpha_ex6, beta = beta_ex6,
      sided = 2, ratio = 1/k)
  )

  N_int <- ceiling(N_total)
  n1 <- ceiling(N_int / (1 + k))
  n2 <- N_int - n1

  ar_ex6$N[i] <- N_int
  ar_ex6$n1[i] <- n1
  ar_ex6$n2[i] <- n2
}

ar_ex6

```

Άσκηση 7 – Σχεδιασμός μελέτης επιβίωσης: hazard ratio, accrual, follow-up, dropout

Θεωρήστε τυχαιοποιημένη κλινική μελέτη δύο ομάδων με πρωτεύον καταληκτικό την επιβίωση, όπου ο σχεδιασμός βασίζεται σε log-rank planning.

Υποθέτουμε εκθετική κατανομή επιβίωσης σε κάθε ομάδα.

Δίνονται:

- median survival στην ομάδα ελέγχου: $m_0 = 15$ μήνες,
- median survival στην πειραματική ομάδα: $m_1 = 25$ μήνες,
- ίση κατανομή 1:1,
- $\alpha = 0.05$ (δίπλευρος έλεγχος),
- $\beta = 0.20$ (ισχύς 0.80),
- διάρκεια accrual: $R = 55$ μήνες,
- ελάχιστο follow-up μετά το τέλος accrual: $F = 10$ μήνες,
- συνολική διάρκεια: $T = R + F = 65$ μήνες.

Υπόδειξη

Θεωρητική προσέγγιση

1. Υποθέτοντας εκθετική κατανομή, δείξτε ότι:

$$\lambda = \frac{\log(2)}{m}.$$

2. Υπολογίστε λ_0 , λ_1 και το hazard ratio

$$HR = \frac{\lambda_1}{\lambda_0},$$

και δείξτε ότι υπό εκθετική κατανομή:

$$HR = \frac{m_0}{m_1}.$$

3. Με $R = 55$, $F = 10$, $T = 65$, χρησιμοποιήστε τη `gsDesign::nSurv` για να υπολογίσετε το απαιτούμενο συνολικό sample size N .
4. Υπολογίστε έναν απλό δείκτη ρυθμού στρατολόγησης ως N/R .
5. Επαναλάβετε με dropout hazard $\eta = 0.009$.
6. Επαναλάβετε με ελάχιστο follow-up $F = 100$ (άρα $T = 155$) και σχολιάστε γιατί το N αναμένεται να μειωθεί.

R

```
library(gsDesign)

m0_ex7 <- 15
m1_ex7 <- 25
lambda0_ex7 <- log(2) / m0_ex7
lambda1_ex7 <- log(2) / m1_ex7
HR_ex7 <- lambda1_ex7 / lambda0_ex7

R_ex7 <- 55
F_ex7 <- 10
T_ex7 <- R_ex7 + F_ex7

out1_ex7 <- nSurv(lambdaC = lambda0_ex7,
                  hr = HR_ex7,
                  hr0 = 1,
                  eta = 0,
                  alpha = 0.05,
                  beta = 0.20,
                  sided = 2,
                  ratio = 1,
                  R = R_ex7,
                  minfup = F_ex7,
                  T = T_ex7)

N1_ex7 <- out1_ex7$n
```

```

accrual_rate_ex7 <- N1_ex7 / R_ex7
c(N = N1_ex7, accrual_rate = accrual_rate_ex7)

# dropout
out2_ex7 <- nSurv(lambdaC = lambda0_ex7,
                  hr = HR_ex7,
                  hr0 = 1,
                  eta = 0.009,
                  alpha = 0.05,
                  beta = 0.20,
                  sided = 2,
                  ratio = 1,
                  R = R_ex7,
                  minfup = F_ex7,
                  T = T_ex7)

out2_ex7$n

# extend follow-up
F2_ex7 <- 100
T2_ex7 <- R_ex7 + F2_ex7

out3_ex7 <- nSurv(lambdaC = lambda0_ex7,
                  hr = HR_ex7,
                  hr0 = 1,
                  eta = 0,
                  alpha = 0.05,
                  beta = 0.20,
                  sided = 2,
                  ratio = 1,
                  R = R_ex7,
                  minfup = F2_ex7,
                  T = T2_ex7)

out3_ex7$n

```