Κεφάλαιο 14

Μέθοδος Newton-Raphson

Θα συζητήσουμε υπολογισμό της εκτιμήτριας μεγίστης πιθανοφάνειας με τη μέθοδο Newton-Raphson. Αν και υπάρχουν περιπτώσεις για τις οποίες η λύση μπορεί να υπολογιστεί ακριβώς, στα περισσότερα παραδείγματα η Ε.Μ.Π. πρέπει να βρεθεί με αναδρομικές αριθμητικές μεθόδους, όπως η Newton-Raphson. Θα συζητήσουμε τις βασικές αρχές με ένα παράδειγμα.

14.1 Παράδειγμα

Τα δεδομένα αναφέρονται σε ώρες επιβίωσης που μετρούν την αντοχή συγκεκριμένων πλοίων σε συνθήκες πίεσης (Τα δεδομένα παρατίθενται στο Παράρτημα). Ένα συγκεκριμένο μοντέλο που χρησιμοποιείται για ανάλυση δεδομένων επιβίωσης είναι η κατανομή Weibull.

$$f(y;\lambda;\theta) = \frac{\lambda y^{\lambda-1}}{\theta^{\lambda}} e^{-(\frac{y}{\theta})^{\lambda}}, y > 0,$$

όπου,

λ: παράμετρος που καθορίζει το σχήμα της κατανομής, και

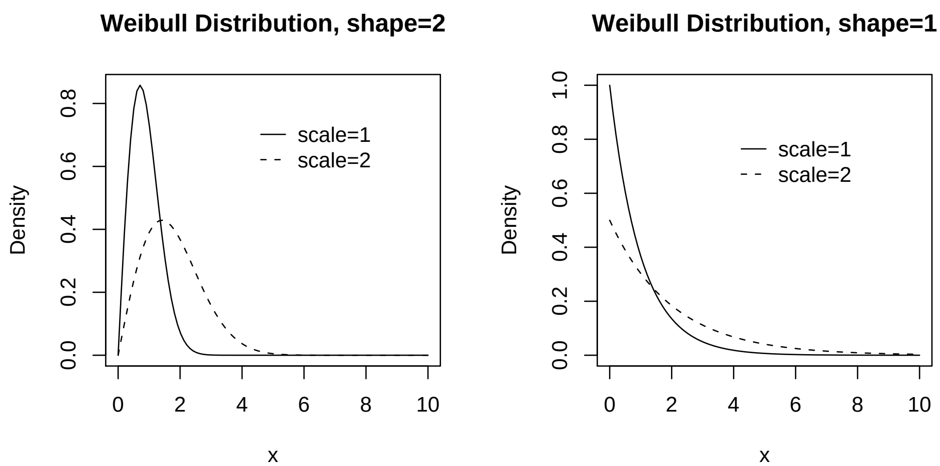
ϑ: παράμετρος που καθορίζει την κλίμακα

Δίνονται τα γραφήματα της κατανομής Weibull για $\lambda = 1, 2$ και $\theta = 1, 2$ (Σχήμα 14.1) για κατανόηση του σχήματος της συνάρτησης πυκνότητας πιθανότητας.

```
> x=seq(0,10, length=100)
> par(mfrow=c(1,2))
> plot(x, dweibull(x, shape=2, scale=1), type="l", ylab="Density")
> lines(x, dweibull(x, shape=2, scale=2), lty=2)
```

```
> title(main="Weibull Distribution, shape=2")
> leg.names<-c("scale=1","scale=2")
> legend(locator(1),leg.names,lty=1:2,bty="n")

> plot(x, dweibull(x, shape=1, scale=1), type="l", ylab="Density")
> lines(x, dweibull(x, shape=1, scale=2), lty=2)
> title(main="Weibull Distribution, shape=1")
> leg.names<-c("scale=1","scale=2")
> legend(locator(1),leg.names,lty=1:2,bty="n")
```

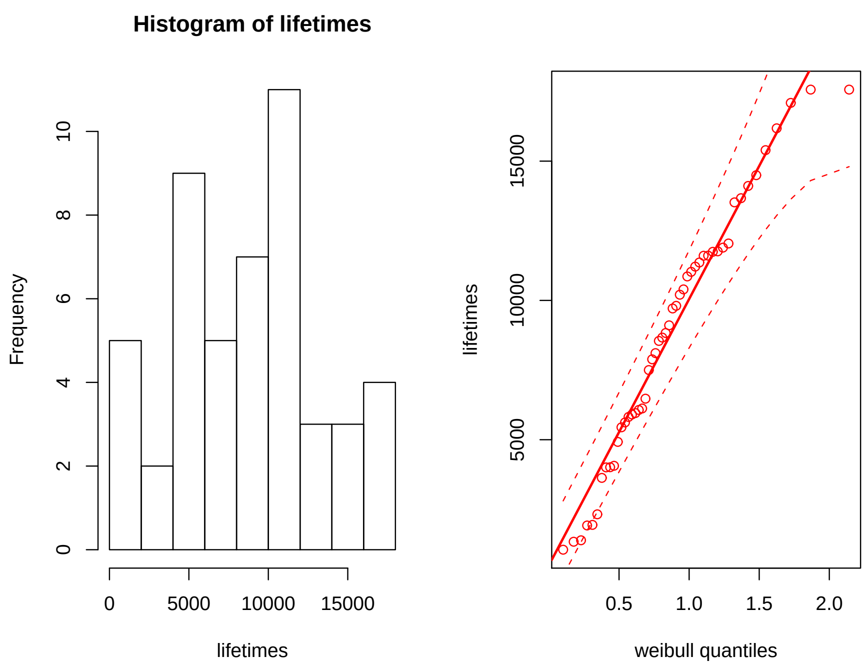


Σχήμα 14.1: Κατανομή Weibull για $\lambda = 1, 2$ και $\theta = 1, 2$.

Επίσης στο Σχήμα 14.2 δίνεται το ιστόγραμμα των σχετικών δεδομένων και το QQ plot με βάση την κατανομή Weibull με $\lambda = 2$.

```
> par(mfrow=c(1,2))
> hist(lifetimes)
> library("car")
> qq.plot(lifetimes, dist="weibull", shape=2)
> qqline(lifetimes, rweibull(49, shape=2))
```

Όπως βλέπουμε από το Σχήμα 14.2, μπορύμε να υποθέσουμε ότι τα δεδομένα ακολουθούν την κατανομή Weibull με $\lambda = 2$ (γαιτί·) Αφού έχουμε αποδεχτεί ότι



Σχήμα 14.2: Ιστόγραμμα και QQ plot των δεδομένων.

η παράμετρος λ είναι γνωστή, έστω τώρα $y_1, \ldots, y_n$ δεδομένα με λ γνωστό (στην περίπτωσή μας $n = 49$). Τότε η από κοινού συνάρτηση πυκνότητας πιθανότητας δίνεται από

$$f(y_1, \ldots, y_n; \theta) = \prod_{i=1}^{n} \frac{\lambda {y_i}^{\lambda-1}}{\theta^{\lambda}} e^{-(\frac{y_i}{\theta})^{\lambda}},$$

οπότε η πιθανοφάνεια δίνεται από

$$L(\theta) = \log f(y_1, \ldots, y_n; \theta) = \sum_{i=1}^{n} \left\{ (\lambda - 1)\log y_i + \log \lambda - \lambda \log \theta - (\frac{y_i}{\theta})^{\lambda} \right\}.$$

Στην R η συνάρτηση λογαριθμικής πιθανοφάνειας $L(\theta)$ μπορεί να οριστεί ως ακολούθως :

```
> loglikelihood <- function(data, theta, lambda=2)
+ { +
```

```
logl<-sum((lambda-1)*log(data)+log(lambda)-lambda*log(theta)-
+    (data/theta)^{lambda})
+ return(logl)
+ }
> theta1 <- seq(7000, 15000, by=100)
> loglik=rep(NA, length(theta1))
> for (i in 1:length(theta1))
+ {
+ loglik[i]=loglikelihood(lifetimes, theta1[i])
+ }
```

Το γράφημα της παρουσιάζεται στο Σχήμα 14.3 και παρατηρούμε ότι υπάρχει τιμή της θ η οποία μεγιστοποιεί την $L(\theta)$. Η συνάρτηση score δίνεται από

$$\frac{dl}{d\theta} = U = \sum_{i=1}^{n} \left\{ -\frac{\lambda}{\theta} + \frac{\lambda {y_i}^{\lambda}}{\theta^{\lambda+1}} \right\} = -\frac{\lambda n}{\theta} + \frac{\lambda \sum_{i=1}^{n} y_i^{\lambda}}{\theta^{\lambda+1}},$$

και στην R ορίζεται ως

```
> get.score <- function(data, theta, lambda=2)
+ { + score<--(lambda*length(data)/theta)+
+ lambda*(sum(data^{lambda}))/(theta^{3})
+ return(score)
+ }
```

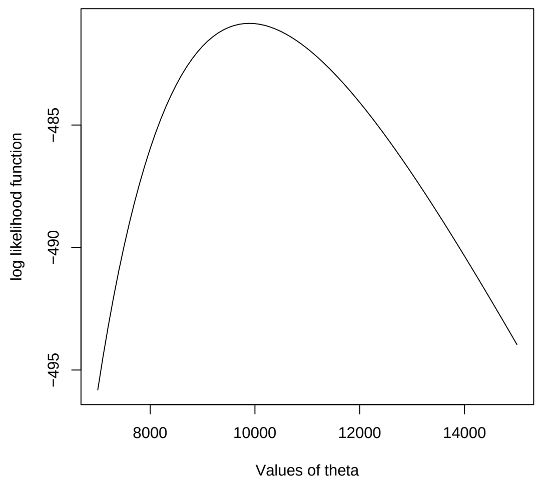
Παρατηρούμε ότι για $\lambda = 2$,

$$U(\theta) = 0 \Rightarrow -\frac{2n}{\theta} + \frac{2\sum_{i=1}^{n} {y_i}^2}{\theta^3} = 0 \Rightarrow \frac{n}{\theta} = \frac{\sum_{i=1}^{n} {y_i}^2}{\theta^3} \Rightarrow \theta^2 = \frac{\sum_{i=1}^{n} {y_i}^2}{n}$$

$$\Rightarrow \theta = \sqrt{\frac{\sum_{i=1}^{n} {y_i}^2}{n}},$$

δηλαδή η Ε.Μ.Π. μπορεί να υπολογιστεί ακριβώς. Θα συγκρίνουμε το ακριβές αποτέλεσμα με εκείνο το οποίο δίνουν οι αναδρομικές μέθοδοι. Πρώτα όμως εξηγούμε τη μέθοδο Newton-Raphson. Γενικά θέλουμε να υπολογίσουμε την τιμή της x, για την οποία $f(x) = 0$. Η εφαπτομένη της $f(x)$ στο σημείο $x^{(m-1)}$ δίνεται από

$$\left[\frac{df}{dx}\right]_{x=x^{(m-1)}} = f'(x^{(m-1)}) = \frac{f(x^{(m)}) - f(x^{(m-1)})}{x^{(m)} - x^{(m-1)}},$$



Σχήμα 14.3: Γράφημα της συνάρτησης λογαριθμικής πιθανοφάνειας.

όπου η απόσταση $x^{(m)} - x^{(m-1)}$ είναι μικρή. Αν το $x^{(m)}$ είναι η λύση της $f(x) = 0$, δηλαδή $f(x^{(m)}) = 0$, έχουμε ότι

$$x^{(m)} = x^{(m-1)} - \frac{f(x^{(m-1)})}{f'(x^{(m-1)})}.$$

Για $m = 1, 2, \ldots,$ και με αρχική τιμή $x^{(1)}$, βρίσκουμε διαδοχικές προσεγγίσεις έτσι ώστε $|x^{(m)} - x^{(m-1)}| < \varepsilon$. Ειδικά, για την εκτιμήτρια μεγίστης πιθανοφάνειας (Ε.Μ.Π),

$$\theta^{(m)} = \theta^{(m-1)} - \frac{U(\theta^{(m-1)})}{U'(\theta^{(m-1)})}.$$

Έχουμε ότι,

$$U(\theta) = -\frac{2n}{\theta} + \frac{2\sum_{i=1}^{n} {y_i}^2}{\theta^3} = 0,$$

και

$$\frac{dU(\theta)}{d\theta} = U'(\theta) = \sum_{i=1}^{n} \left\{ \frac{\lambda}{\theta^2} - \frac{\lambda(\lambda+1){y_i}^{\lambda}}{\theta^{\lambda+2}} \right\} = \frac{2n}{\theta^2} - \frac{2 \cdot 3 \cdot \sum {y_i}^2}{\theta^4},$$

όπου στην τελευταία ισότητα θέτουμε $\lambda = 2$. Αντί να χρησιμοποιήσουμε την $U'(\theta)$, θέτουμε στην παραπάνω αναδρομική σχέση την $E(U'(\theta)) = -J(\theta)$. Μπορεί να αποδειχτεί ότι

$$J(\theta) = -E(U'(\theta)) = \frac{\lambda^2 n}{\theta^2}.$$

Η συνάρτηση $J(\theta)$ ονομάζεται πληροφορία Fischer και γραφεταί στην R ως

```
> get.information <- function(data, theta, lambda=2)
+ {
+ information <- (lambda^{2}*length(data))/(theta^{2})
+ return(information)
+ }
```

Συνεπώς, καταλήγουμε σε μια τροποποίηση του αλγορίθμου Newton-Raphson, ο οποίος ονομάζεται Fisher scoring.Δηλαδή,

$$\theta^{(m)} = \theta(m-1) + \frac{U(\theta^{(m-1)})}{J(\theta^{(m-1)})},$$

Χρησιμοποιώντας την πιο πάνω αναδρομική σχέση μπορούμε να βρούμε την εκτιμήτρια μέγιστης πιθανοφάνειας για το θ. Το τυπικό σφάλμα για το θ δίνεται από τον τύπο

$$s(\hat{\theta}) = \sqrt{\frac{1}{J}},$$

και ένα 95% διάστημα εμπιστοσύνης δίνεται από

$$\hat{\theta} \pm 1.96 \cdot s(\hat{\theta}).$$

Στη συνέχεια παρουσιάζεται πώς εφαρμόζεται η μέθοδος αυτή στην R για να υπολογιστεί η Ε.Μ.Π. για το θ, το τυπικό του σφάλμα και ένα 95% διάστημα εμπιστοσύνης.

```
> ybar <- mean(lifetimes)
> ybar
[1] 8805.694
> theta <- ybar
> it  <- 0   ####iterative count
> del <- 1   ####iterative adjustment
> while(abs(del) > 0.00001 && (it <- it+1) < 10)
+ {
+ del<-get.score(lifetimes,theta)/get.information(lifetimes,theta)
+ theta <- theta+del
+ cat(it,theta,get.score(lifetimes,theta),
```

```
+ get.information(lifetimes,theta),"\n")
+ }
1 9959.204 -0.0001320064 1.976090e-06
2 9892.402 -4.517492e-07 2.002869e-06
3 9892.177 -5.150392e-12 2.002960e-06
4 9892.177 1.734723e-18  2.002960e-06
> sqrt(mean(lifetimes^{2})) ###exact value
[1] 9892.177
> ##########Confidence Interval
> sderror <- sqrt(1/2.002960e-06)
> sderror
[1] 706.5841
> 9892.177-1.96*sderror;    9892.177+1.96*sderror
[1] 8507.272 [1] 11277.08
```

Από τα παραπάνω, παρατηρούμε ότι οι αναδρομικές σχέσεις που ορίζονται από τον αλγόριθμο Fisher-scoring συγκλίνουν στην ακριβή τιμή της $\hat{\theta}$.

Παράρτημα

Τα δεδομένα που χρησιμοποιούνται σε αυτό το κεφάλαιο για την εύρεση της Ε.Μ.Π. με την μέθοδο Newton-Raphson.

```
   lifetimes
1       1051
2       1337
3       1389
4       1921
5       1942
6       2322
7       3629
8       4006
9       4012
10      4063
11      4921
12      5445
13      5620
```

14	5817
15	5905
16	5956
17	6068
18	6121
19	6473
20	7501
21	7886
22	8108
23	8546
24	8666
25	8831
26	9106
27	9711
28	9806
29	10205
30	10396
31	10861
32	11026
33	11214
34	11362
35	11604
36	11608
37	11745
38	11762
39	11895
40	12044
41	13520
42	13670
43	14110
44	14496
45	15395
46	16179
47	17092
48	17568
49	17568