Παρασκευή 16/10/20 Μάθημα 2ο

Τυχαίο δείγμα:

Τυχαίο δείγμα μεγέθους n από την κατανομή F(x,θ) (ή σ.π.π f(x,θ)) καλείται μια συλλογή ανεξάρτητων και ισόνομων τ.μ. $X_1, .. X_n \sim F(x,\theta)$

- είναι στην ουσία μετρήσεις μιας τ.μ. $X \sim f(x;\theta)$.

π.χ) αν μας ενδιαφέρει το εισόδημα X των νοικοκυριών της Ελλάδας. η f(x,θ) είναι η κατανομή των εισοδημάτων όλων των νοικοκυριών της χώρας και οι $X_1, .. X_n$ είναι μετρήσεις που παίρνουμε ρωτώντας στην τύχη n νοικοκυριά

Στατιστική (ή δειγματική) συνάρτηση

Έστω τ.δ. $X_1, X_2, .. X_n \sim F(x,\theta)$.
Στατιστική συνάρτηση (σ.σ) καλείται κάθε μετρήσιμη συνάρτηση $T = T(X) = T(X_1, X_2, .. X_n)$ των $X_1, X_2, .. X_n$ που δεν εξαρτάται από άγνωστες παραμέτρους.

$$\bar{X} = \frac{1}{n}\sum x_i \,, \quad m_r = \frac{1}{n}\sum x_i^r \,, \quad \sigma^2 = \frac{1}{n}\sum(x_i-\bar{x})^2$$

- Μια εκτιμήτρια συνάρτηση T(X) είναι και αυτή μια τυχαία μεταβλητή (κάθε φορά που παίρνουμε ένα άλλο τυχαίο δείγμα X η T δίνει διαφορετική τιμή)

Πρακτική Άσκηση:

Παράγετε 20 δείγματα 20 παρατηρήσεων έκαστο από την $N(2,2^2)$ κατανομή

1) Εκτιμήστε τη μέση τιμή κάθε δείγματος.

2) » την κατανομή της εκτιμήτριας που χρησιμοποιήσατε βάση των υπολογισμένων τιμών της και δώστε την γραφική της παράσταση

3] Επαναλάβετε τα 1 και 2 με εκτιμήτη της μέσης τιμής την συνάρτηση $(X_{min} + X_{max})/2$.

4) Συγκρίνετε τις 2 κατανομές και διατυπώστε τα συμπεράσματα σας.

▶ Πως συνδέονται οι σ.σ ($\bar{X}$, m_r, σ^2, κλπ) που μελετάμε με το γεγονός ότι οι παρατηρήσεις $X_1 \ldots X_n$ είναι ανεξάρτητες και ισόνομες από $f(x,\theta)$. ;

Απάντηση:

Αρχικά παρατηρούμε ότι όλες ορίζονται ως κάποιας μορφής άθροισμα των $X_1, \ldots X_n$.

$X_1 .. X_n$ ανεξάρτητες $\Rightarrow V\left\{\sum_{i=1}^{n} f(X_i)\right. = \sum_{i=1}^{n} V(f(x_i))$

όπου f κάποια συνάρτηση των X_i

- $\bar{X} \Rightarrow f(x) = x$
- $m_r \Rightarrow f(x) = x^r$
- $\sigma^2 \Rightarrow f(x) = (x-\bar{x})^2$

- $E\sum_{i=1}^{n} f(X_i) = \sum_{i=1}^{n} E(f(x_i))$ Ισχύει ασχέτως αν είναι ανεξάρτητες ή όχι

$V(X+Y) = V(X) + V(Y)$

Αν X, Y εξαρτημένες τότε $V(X+Y) = V(X) + V(Y) + Cov(X,Y)$.

$X_1 .. X_n$ ισόνομες $\Rightarrow$ κάθε μια από τις X_i ακολουθεί ίδια κατανομή με τις άλλες

Δηλαδή έχουν όλες ίδια μέση τιμή και διακύμανση

Στην πράξη πρέπει να μπορούμε να επαληθεύσουμε

▷ την τυχαιότητα του δείγματος

▷ την προέλευση του από συγκεκριμένη κατανομή

▷ την ανεξαρτησία τ.μ.

Εκτιμήτρια συνάρτηση:

Εκτιμήτρια συνάρτηση της παραμέτρου θ (ή των παραμέτρων θ) ή μιας παραμετρικής συνάρτησης $g(\theta)$ θα καλείται μια σ.σ $T=T(X)$ η οποία χρησιμοποιείται για την εκτίμηση του θ.

- Για ένα τ.δ X μπορούμε να κατασκευάσουμε πολλές εκτιμήτριες για μια παράμετρο θ.
- Διαισθητικά αναμένουμε ότι μια "καλή εκτιμήτρια" T του θ θα λαμβάνει τιμές
  - γύρω από το θ
  - κοντά στο θ με μεγάλη πιθανότητα

οπότε θέλουμε μια που να έχει τιμή θ ή "σχεδόν" θ και με "μικρή" διασπορά,

Αμερόληπτη εκτιμήτρια:

Μια εκτιμήτρια συνάρτηση T της $g(\theta)$ καλείται αμερόληπτη εάν $E(T)=E(T(X))=g(\theta)$ για κάθε θ

Αποτελεσματικότητα:

Αν T_1, T_2 είναι δυο α.ε η T_1 θα καλείται αποτελεσματικότερη της T_2 εάν ισχύει $V(T_1)<V(T_2)$

Α.Ο.Ε.Δ εκτιμήτρια:

Αν μια α.ε $T=T(X)$ έχει $(\forall\theta)$ τη μικρότερη διασπορά μεταξύ όλων των α.ε του $g(\theta)$ τότε θα καλείται άριστη ή αμερόληπτη ομοιόμορφα ελάχιστης διασποράς (Α.Ο.Ε.Δ).

Μέσο τετραγωνικό σφάλμα:

Αν T είναι μια εκτιμήτρια του $g(\theta)$ η ποσότητα $MSE(T) = E(T-g(\theta))^2 = V(T) + bias^2(T)$ και καλείται Μ.Τ.Σ της T από την $g(\theta)$

Αν T_1, T_2 δυο εκτιμήτριες της $g(\theta)$ η T_1 είναι αποτελεσματικότερη της T_2 εαν $MSE(T_1) < MSE(T_2)$

Παράδειγμα 1

Έστω τ.δ $X_1, \ldots X_n \sim N(\theta, \sigma^2)$ όπου $\theta \in \mathbb{R}$ και σ^2 γνωστή διακύμανση
Να υπολογιστεί το Μ.Τ.Σ του $\bar{X}$

$$MSE(\bar{X}) = E(\bar{X}-\theta)^2 = Var(\bar{X}) + (E(\bar{X}-\theta))^2$$

$$Var(\bar{X}) = \frac{\sigma^2}{n}$$

$$E(\bar{X}) = \theta.$$

οποτε $MSE(\bar{X}) = \frac{\sigma^2}{n}.$

Παράδειγμα 2:

Έστω τ.δ $X_1, \ldots X_n \sim Ber(1, \theta)$ $\theta \in \mathbb{R}$.
Να υπολογιστή το Μ.Τ.Σ του $\bar{X}$ και να βρεθεί ένα άνω φράγμα του και να σχολιαστεί.

$$MSE(\bar{X}) = Var(\bar{X}) + \{E(\bar{X}-\theta)\}^2 = \frac{\theta(1-\theta)}{n} + (\overset{0}{\cancel{\theta-\theta}})^2$$

Η συνάρτηση που θέλω να μεγιστοποιήσω είναι η $g(\theta) = \theta - \theta^2$.

$g'(\theta) = 1 - 2\theta, \Rightarrow 1 - 2\theta = 0 \Rightarrow \theta = \frac{1}{2}.$

$g''(\theta) < 0$ άρα όντως μέγιστο.

$MSE(\bar{X}) \leq g(1/2)/n = \frac{1}{4n}.$

Όσο αυξάνεται το μέγεθος του δείγματος τοσο μειώνεται το μέσο τετραγ. σφάλμα

Μέθοδος μέγιστης πιθανοφάνειας:

Είναι αρκετά ισχυρή διότι οδηγεί σε εκτιμήτριες με πολύ καλές ιδιότητες. Παράγει την πιο πιθανή να συμβεί στην πράξη εκτίμηση.

Συνάρτηση πιθανοφάνειας:

Συνάρτηση πιθανοφάνειας ή πιθανοφάνεια του δείγματος $X_1, X_2, .. X_n$ καλείται η από κοινού σ.π. των $X_1, X_2, .. X_n$ θεωρούμενη ως συνάρτηση του θ, δηλαδή η

$$L_x(\theta) = \prod_{i=1}^{n} f(X_i, \theta)$$

Η παραμετρική μέθοδος εκτίμησης της σ.π.π βασίζεται σε μεγάλο βαθμό στην διαθέσιμη πληροφορία για την οικογένεια που ανήκει η σ.π.π που θέλουμε να εκτιμήσουμε.

π.χ. αν ξέρουμε ότι η υπό εκτίμηση σ.π.π είναι κανονική τότε ξέρουμε ότι $f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

και οι υπό εκτίμηση άγνωστοι είναι οι 2 παράμετροι μ, σ^2

Η μέθοδος της μέγιστης πιθανοφάνειας βασίζεται στην έννοια της σ.π.π., όσο πιο μεγάλη είναι η $f(X_i, \theta)$ τόσο πιο πιθανό είναι να παρατηρηθεί ξανά η συγκεκριμένη παρατήρηση.

$$L = (\theta, X_1, \ldots X_n) = \prod_{i=1}^{n} f(\theta, X_1, \ldots, X_n)$$

Μεγιστοποίηση της L ως προς θ σημαίνει ότι η τιμή του θ που προκύπτει από τη διαδικασία μεγιστοποιεί την πιθανοφάνεια επανεμφάνισης του δείγματος.
Η θ μπορεί να είναι μια τιμή ή διάνυσμα τιμών.

Αναλόγως με την υπο εκτίμηση σ.π.π και τον τύπο δεδομένων οι βασικές μέθοδοι εκτίμησης είναι:

→ μέθοδοι βασισμένες στον αλγόριθμο EM

→ μέθοδοι » » Newton

Εκτιμήτρια μέγιστης πιθανοφάνειας

Μια στατιστική συνάρτηση $\hat{\theta}$ καλείται εκτιμήτρια μέγιστης πιθανοφάνειας (ΕΜΠ) των παραμέτρων θ αν μεγιστοποιεί την πιθανοφάνεια $L_x(\theta)$

δηλ. $L_x(\hat{\theta}) = \sup\limits_{\theta \in \Theta} L_x(\theta)$

• Για ευκολία αναζητούμε το μέγιστο της $\ln(L_x(\theta))$
Λύνουμε την $l'(\theta) = 0$ ελέγχουμε αν $l''(\theta) < 0$

Για $k \geq 2$, $\theta = (\theta_1, \dots \theta_k)$ έχουμε πρόβλημα βελτιστοποίησης στο $\mathbb{R}^k$. Σχηματίζουμε τις $\frac{\partial l(\theta)}{\partial \theta_1} = 0, \dots \frac{\partial l(\theta)}{\partial \theta_k} = 0$

$$\frac{\partial^2 l(\theta)}{\partial \theta_1^2} < 0, \dots \frac{\partial^2 l(\theta)}{\partial \theta_k^2} < 0$$

Είναι πιθανόν να μην υπάρχει πεπερασμένο μέγιστο να υπάρχουν παραπάνω από ένα μέγιστα ή να έχουμε ακριβώς ένα μέγιστο

Αν $\hat{\theta}$ είναι η ΕΜΠ της θ και g αμφιμονοσήμαντη τότε $g(\hat{\theta})$ ΕΜΠ της $g(\theta)$

Ασκηση 1:

Εστω τ.δ $X_1, X_2, \dots X_n$ από την εκθετική κατανομή με σ.π.π
$f(x,\theta) = \theta^{-1} e^{x\theta^{-1}}$, $0 \le x \le +\infty$, $\theta \in (0,+\infty)$
Να υπολογιστή ο ΕΜΠ για την θ.

$$L(\theta) = \prod_{i=1}^{n} f(x_i,\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{\frac{x}{\theta}} = \frac{1}{\theta^n} e^{-\frac{1}{\theta}\Sigma x_i}$$

$$\ln L(\theta) = -\frac{1}{\theta} \Sigma x_i - n\ln\theta$$

$$\frac{\partial}{\partial\theta} \ln(L\theta) = \frac{1}{\theta^2} \Sigma x_i - \frac{n}{\theta}$$

$$\frac{\partial}{\partial\theta} \ln(L\theta) = 0 \Rightarrow \Sigma x_i - \theta n = 0 \Rightarrow \theta = \bar{X}$$

$$\frac{\partial^2 \ln L(\theta)}{\partial\theta^2} = \frac{n}{\theta^2} - \frac{2}{\theta^3} \Sigma x_i \overset{\theta=\bar{X}}{=} \frac{n}{\bar{X}^2} - \frac{2 n \bar{X}}{\bar{X}^2} < 0.$$

Ασκηση 2:

Ενα τ.δ $X_1, \dots X_n$ απο κανονική κατανομή με σ.π.π.

$$f(x,\theta_1,\theta_2) = \frac{1}{\sqrt{\theta_2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta_1)^2}{2\theta_2}} \quad , \theta_1 \in R, \ \theta_2 \in (0,+\infty)$$

Να υπολογιστει ο ΕΜΠ για τις παραμετρων $\theta = (\theta_1, \theta_2)$

$$L(\theta) = \frac{1}{\sqrt{\theta_2^n}} \cdot \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2\theta_2}\Sigma(x_i-\theta_1)^2}$$

$$\text{Log}(L\theta)) = \frac{1}{-2\theta_2} \Sigma(x_i-\theta_1)^2 - \frac{n}{2} \ln\theta_2 - \frac{n}{2} \ln(2\pi)$$

$$\frac{\partial \ln(L(\theta))}{\partial\theta_1} = -\frac{2(-1)\Sigma(x_i-\theta_1)}{\theta_2} = 0 \Rightarrow \sum_{i=1}^{n} (x_i-\theta_1) = 0 \Rightarrow$$

$$\hat{\theta_1} = \bar{x}$$

$$\frac{\partial \ln L(\theta)}{\partial \theta_2} = \frac{-n}{2\theta_2} + \frac{\Sigma(x_i-\theta_i)^2}{2\theta_2^2} = 0 \Rightarrow$$

$$\hat{\theta_2} = \frac{1}{n}\Sigma(x_i-\bar{x})^2$$

Άσκ. 3

Έχουμε παρατηρήσεις $X_1, \dots X_n$ που αφορούν ώρες επιβίωσης συγκεκριμένων πλοίων σε συνθήκες πίεσης.
Ένα συγκεκριμένο μοντέλο που χρησιμοποιείται για ανάλυση δεδομένων επιβίωσης είναι η κατανομή weibull

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

όπου λ καθορίζει το σχήμα της κατανομής και θ την κλίμακα.

Για το συγκεκριμένο παράδειγμα $\lambda=2$.

$$f(x_1, \dots x_n) = \prod_{i=1}^{n} \frac{\lambda x_i^{\lambda-1}}{\theta^\lambda} e^{-\left(\frac{x_i}{\theta}\right)^\lambda}$$

$$L(\theta) = \log f(x_1, \dots x_n) = \sum_{i=1}^{n} \left\{ (\lambda-1)\log x_i + \log\lambda - \lambda\log\theta - \left(\frac{x_i}{\theta}\right)^\lambda \right\}$$

$$\frac{dL\theta}{d\theta} = 0 \Rightarrow \sum_{i=1}^{n} \left(\frac{-\lambda}{\theta} + \frac{\lambda x_i^\lambda}{\theta^{\lambda+1}} \right) = -\frac{\lambda n}{\theta} + \frac{\lambda \Sigma x_i^\lambda}{\theta^{\lambda+1}} = 0$$

$$\xrightarrow{\lambda=2} \frac{-2n}{\theta} + \frac{2}{\theta^3}\Sigma x_i^2 = 0 \Rightarrow \theta^2 = \frac{\Sigma x_i^2}{n}$$

$$\Rightarrow \theta = \sqrt{\frac{1}{n}\Sigma x_i^2}$$

Για να ~~[illegible]~~ υλοποιήσουμε την εκτίμηση στην R:

```
lifetimes <- read.table("lifetimes.txt"),[ ,2]
thetause <- sqrt(mean(lifetimes^2))
loglikelihood <- function(data, theta, lamda =2)
 {
   logl <- sum ((lamda-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) <- logliklihood(lifetimes, theta1[i])
  }
plot(theta1, loglik, type ="1")
```

Όταν δεν μπορούμε να βρούμε μέγιστο με αναλυτικό τρόπο

$$u(\theta) = -\frac{2n}{\theta} + 2\frac{\sum x_i^2}{\theta^3} = 0$$

$$\frac{du(\theta)}{d\theta} = u'(\theta) = \frac{2n}{\theta^2} - \frac{2\cdot 3\sum x_i^2}{\theta^4}$$

Αντί να χρησιμοποιήσουμε την $u'(\theta)$ θέτουμε στην παρακάτω αναδρομική σχέση $E(u'(\theta)) = -J(\theta)$

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

$J(\theta)$ πληροφορία του fischer

$J(\theta)$ στην R

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

Newton-Raphson

Η εφαπτομένη οποιασδήποτε συνάρτησης $f(x)$ στο $x^{(m-1)}$

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

όταν η απόσταση $x^{m}-x^{m-1}$ είναι μικρή

Αν $x^{(m)}$ λύση της $f(x)=0$ τότε

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

Για $m=1,2..$ με αρχική τιμή $x^{(1)}$ βρίσκουμε διαδοχικές προσ. ώστε $|x^{(m)}-x^{(m-1)}|<\varepsilon$.

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

που είναι μια τροποποίηση του αλγόριθμου Newton-Raphson που ονομάζεται Fisher scoring

Άσκηση:

Χρησιμοποιώντας την πιο πάνω αναδρομική σχέση συγκρίνετε την ακριβή τιμή της ΕΜΠ για την ασκ. 3 με την προσεγγιστική που προκύπτει από την εφαρμογή της αριθ. μέθοδο

Έστω ότι τυπικό σφάλμα $s(\hat{\theta})=\sqrt{\frac{1}{J}}$ μπορούμε να υπολογίσουμε ένα διάστημα μέσα στο οποίο να περιμένουμε να είναι η θ_j