Ασκ. 36 (Θέμα 2 | Σεπτ. 16 (γ) + επιπλέον ερωτήματα)

Έστω $X_1, \dots, X_\nu$ τ.δ. από κατανομή με σ.π.π.

$$f(x; a, b) = a b^{a} \frac{1}{x^{a+1}} \mathbb{1}_{[b,+\infty)}(x) \quad , a, b > 0 .$$

α) Να βρεθεί η ε.μ.π. $(\hat{a}, \hat{b})$ του (a, b).

β) Ν.δ.ο. η $\hat{a}_\nu$ και $\hat{b}_\nu$ είναι συνεπείς εκτιμήτριες των a, b.

γ) Να βρεθούν οι ασυμπτωτικές κατανομές των $\hat{a}_\nu$ και $\hat{b}_\nu$.

Λύση

α) Δείξαμε στο Μαθ. 14 ότι $\hat{b} = X_{(1)}$.

$$L(a,b) = a^\nu \left(\prod_{i=1}^{\nu} x_i\right)^{-a-1} b^{\nu a} \mathbb{1}_{(0, x_{(1)}]}(b).$$

Άρα $$\max_a \max_b L(a,b) = \max_a L(a, x_{(1)}) = \max_a a^\nu \left(\prod_{i=1}^{\nu} x_i\right)^{-a-1} x_{(1)}^{\nu a}$$

$$\overset{\log(\cdot)}{=} \max_a \{ \underbrace{\nu \log a - (a+1) \sum_{i=1}^{\nu} \log x_i + \nu a \log x_{(1)}}_{f(a)} \}.$$

$$f'(a) = \frac{\nu}{a} - \sum_{i=1}^{\nu} \log x_i + \nu \log x_{(1)}$$

Άρα $$f'(a) = 0 \Leftrightarrow \frac{\nu}{a} = \sum_{i=1}^{\nu} \log x_i - \nu \log x_{(1)}.$$

$$\Leftrightarrow a^* = \frac{\nu}{\sum_{i=1}^{\nu} \log x_i - \nu \log x_{(1)}} = \frac{1}{\overline{\log x} - \log x_{(1)}}$$

Φανερά το a^* είναι και θέση ολικού μεγίστου, αφού η $f(a)$ είναι γνήσια κοίλη.

β) Δείχνουμε πρώτα ότι η $\hat{b}_\nu = X_{(1)}$ είναι συνεπής, για το b

Έστω $\varepsilon > 0$. $$P\left(|\hat{b}_\nu - b| > \varepsilon\right) \overset{X_{(1)} > b}{=} P\left(X_{(1)} - b > \varepsilon\right) = P\left(X_{(1)} > b + \varepsilon\right)$$

$$= P\left(X_1 > b+\varepsilon, \dots, X_\nu > b+\varepsilon\right) = P\left(X_1 > b+\varepsilon\right)^\nu = \theta^\nu \xrightarrow[\nu \to +\infty]{} 0,$$

αφού $0 < \theta = P(X_1 > b+\varepsilon) < 1$. Αν υποθέταμε ότι $\theta = 1$, τότε το στήριγμα της X_1, θα ήταν το $[b+\varepsilon, +\infty)$, το οποίο είναι άτοπο.

Τώρα θ.δ.ο. η $\hat{a}_\nu = \dfrac{1}{\overline{\log X} - \log X_{(1)}}$ είναι συνεπής εκτιμήτρια του a. Κατ'αρχήν βρίσκουμε την κατανομή της $Y = \log X$.

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right| = f_X(e^y)\, e^y = a b^a e^{-(a+1)y} e^y \mathbb{1}_{[b,+\infty)}(e^y)$$

$$= a b^a e^{-ay} \mathbb{1}_{[\log b, +\infty)}(y)$$

Παρατηρείστε ότι $Y = Z + \log b$, όπου $Z \sim \mathcal{E}xp(a)$.

Από τον Α.Ν.Μ.Α. για την ακολουθία α.ι.τ.μ. $\{\log X_i\}_{i \geq 1}$

με $E \log X_1 = EY_1 = EZ_1 + \log b = \frac{1}{a} + \log b \in \mathbb{R}$,

έχουμε $\overline{\log X} \xrightarrow{P} E\log X_1 = \frac{1}{a} + \log b$, $\forall a, b > 0$.

Επιπλέον δείξαμε ότι $\log X_{(1)} \xrightarrow{P} \log b$, $\forall b > 0$

(συνέπεια της $X_{(1)}$ και Θ.Σ.Α.)

Άρα $\hat{a}_\nu = \dfrac{1}{\overline{\log X} - \log X_{(1)}} \xrightarrow{P} \dfrac{1}{\frac{1}{a} + \log b - \log b} = a$, $\forall a > 0$.

και άρα η $\hat{a}_\nu$ είναι συνεπής.

β) Βρίσκουμε πρώτα την ασυμπτωτική κατανομή της $\hat{b}_\nu = X_{(1)}$.

Έχουμε για $x > 0$, ότι

$$P\left(\nu(X_{(1)} - b) \leq x\right) = P\left(X_{(1)} \leq b + \frac{x}{\nu}\right) = 1 - P\left(X_{(1)} > b + \frac{x}{\nu}\right)$$

$$= 1 - \prod_{i=1}^{\nu} P\left(X_i > b + \frac{x}{\nu}\right) = 1 - P^\nu\left(X_1 > b + \frac{x}{\nu}\right).$$

$$= 1 - \left(\int_{b+\frac{x}{\nu}}^{+\infty} a b^a \frac{1}{t^{a+1}}\, dt\right)^\nu = 1 - \left(\frac{a b^a}{-a}\left[t^{-a}\right]_{b+\frac{x}{\nu}}^{+\infty}\right)^\nu$$

$$= 1 - \left[b^a\left(b + \frac{x}{\nu}\right)^{-a}\right]^\nu = 1 - \left[\left(1 + \frac{\frac{x}{b}}{\nu}\right)^{-a}\right]^\nu \xrightarrow[\nu \to +\infty]{} 1 - e^{-\frac{a}{b}x} \quad (1)$$

Επιπλέον για $x \leq 0$, $P\left(\nu(X_{(1)} - \beta) \leq x\right) = 0 \xrightarrow[\nu\to+\infty]{} 0$. (2)

Απο (1) + (2) έχουμε ότι, $\forall x \in \mathbb{R}$.

$$F_{\nu(X_{(1)}-\beta)}(x) \longrightarrow F_{\mathcal{E}xp\left(\frac{a}{\beta}\right)} \Rightarrow \nu\left(X_{(1)} - \beta\right) \xrightarrow{d} \mathcal{E}xp\left(\frac{a}{\beta}\right) \quad (3)$$

$\forall a, \beta > 0$.

που είναι και η ασυμπτωτική κατανομή της $X_{(1)}$.

Απο τη μέθοδο Δέλτα και επειδή η $\overset{g(x)}{\log x}$ είναι διαφορίσιμη για $x>0$

έχουμε λόγω της (3), ότι $\nu\left(\underset{g(X_{(1)})}{\log X_{(1)}} - \underset{g(\beta)}{\log \beta}\right) \xrightarrow{d} g'(\beta)\, \mathcal{E}xp\left(\frac{a}{\beta}\right)$

Όμως $g'(\beta) = \frac{1}{\beta}$, και άρα $\nu\left(\log X_{(1)} - \log\beta\right) \xrightarrow{d} \frac{1}{\beta}\mathcal{E}xp\left(\frac{a}{\beta}\right) = \mathcal{E}xp(a)$

(4) $(\forall a, \beta > 0)$

Επιπλέον απο Κ.Ο.Θ. και επειδή $Var(\log X_1) = Var Z_1 = Var(\mathcal{E}xp(a)) = \frac{1}{a^2}$

έχουμε

$$\sqrt{\nu}\left(\overline{\log X} - E(\log X_1)\right) = \sqrt{\nu}\left(\overline{\log X} - \frac{1}{a} - \log\beta\right) \xrightarrow[\nu\to\infty]{d} \mathcal{N}\left(0, \frac{1}{a^2}\right) \quad (5)$$

Απο (4) και (5), έχουμε·

$$\sqrt{\nu}\left(\overline{\log X} - \log X_{(1)} - \frac{1}{a}\right) = \sqrt{\nu}\left[\overline{\log X} - \frac{1}{a} - \log\beta - \left(\log X_{(1)} - \log\beta\right)\right]$$

$$= \sqrt{\nu}\left(\overline{\log X} - \frac{1}{a} - \log\beta\right) - \frac{\nu\cdot\left(\log X_{(1)} - \log\beta\right)}{\sqrt{\nu}} \quad (6)$$

Λόγω (4) + Slutsky $\frac{1}{\sqrt{\nu}}\nu\left(\log X_{(1)} - \log\beta\right) \xrightarrow{d} 0 \cdot \mathcal{E}xp(a) = 0$.

Άρα πάλι λόγω Slutsky + (5), (6), έχουμε $\sqrt{\nu}\left(\overline{\log X} - \log X_{(1)} - \frac{1}{a}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{a^2}\right)$

Τώρα τελικά, απο Δέλτα μέθοδο

$$\sqrt{\nu}\left(\frac{1}{\overline{\log X} - \log X_{(1)}} - a\right) = \sqrt{\nu}\left(\phi(W_\nu) - \phi\left(\tfrac{1}{a}\right)\right) \xrightarrow{d} \phi'\left(\tfrac{1}{a}\right)\cdot \mathcal{N}\left(0, \frac{1}{a^2}\right)$$

με $\phi(w) = \frac{1}{w} \Rightarrow \phi'(w) = -\frac{1}{w^2} \Rightarrow \sqrt{\nu}\left(\phi(W_\nu) - \phi\left(\frac{1}{a}\right)\right) \xrightarrow{d} \mathcal{N}(0, a^2)$

$\boxed{\text{Ασκ. 37}}$

Έστω $X_1, \dots, X_\nu$ τ.δ. από $\log N(\mu, \sigma^2)$, όπου $\mu \in \mathbb{R}$, $\sigma^2 > 0$ άγνωστα.
Να βρεθεί η ε.μ.π. $(\hat{\mu}, \hat{\sigma}^2)$ του (μ, σ^2).
(υπόδειξη: κάντε μετασχηματισμό δεδομένων, $\log X_i \sim N(\mu, \sigma^2)$)

Λύση

Έχουμε δείξει ότι αν $Y_1, \dots, Y_\nu$ είναι ένα τ.δ. από $N(\mu, \sigma^2)$,
τότε
$$\boxed{\hat{\mu} = \overline{Y} \quad \text{και} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^{\nu}(Y_i - \overline{Y})^2}{\nu}} \qquad (1)$$

Η συνάρτηση πιθανοφάνειας του αρχικού τ.δ. είναι
$$\boxed{L(\mu, \sigma^2) = \prod_{i=1}^{\nu} f_{X_i}(x_i; \mu, \sigma^2)}, \quad \text{όπου } X_i \sim \log N(\mu, \sigma^2). \qquad (2)$$

Θα μπορούσαμε να υπολογίσουμε τη σ.π.π. των X_i και να λύσουμε το πρόβλημα. Παρατηρούμε όμως ότι
$$X_i \sim \log N(\mu, \sigma^2) \Rightarrow \log X_i \sim N(\mu, \sigma^2).$$
Άρα αν θέσουμε $Y_i = \log X_i \Rightarrow X_i = e^{Y_i}$, $\forall\, 1 \le i \le \nu$,
όπου $\{Y_i\}_{1 \le i \le \nu}$ τ.δ. από $N(\mu, \sigma^2)$, που η λύση ε.μ.π. δίνεται από την (1). Έχουμε $X = g(Y) = e^Y \Leftrightarrow Y = g^{-1}(X) = \log X$.
Άρα
$$f_X(x; \mu, \sigma^2) = f_Y(g^{-1}(x); \mu, \sigma^2) \left| \frac{dg^{-1}(x)}{dx} \right| = f_Y(\log x; \mu, \sigma^2) \cdot \frac{1}{x}$$
Αντικαθιστώντας στη (2), έχουμε
$$L_X(\mu, \sigma^2) = \left(\prod_{i=1}^{\nu} x_i\right)^{-1} \prod_{i=1}^{\nu} f_{Y_i}(\log x_i; \mu, \sigma^2) = \left(\prod_{i=1}^{\nu} x_i\right)^{-1} L_{\log X}(\mu, \sigma^2)$$
Αυτό δείχνει ότι το πρόβλημα μεγιστοποίησης της $L_X(\mu, \sigma^2)$, με δεδομένα τα X_i, είναι ισοδύναμο με το πρόβλημα μεγιστοποίησης της $L_{\log X}(\mu, \sigma^2)$, με δεδομένα τα $\log X_i$. Άρα από (1), και $Y_i = \log X_i$
$$\boxed{(\hat{\mu}, \hat{\sigma}^2) = \left(\overline{\log X}, \frac{\sum_{i=1}^{\nu}(\log X_i - \overline{\log X})^2}{\nu} \right)}$$

Ασκ 38 (i) Να βρεθεί η ε.μ.π. του θ στο πρόβλημα δεδομένων του Hubble, όπου είχαμε υποθέσει ότι $Y_i = \theta X_i + \varepsilon_i$, $1 \le i \le \nu$, $\varepsilon_i \sim N(0, \sigma^2)$ και ανεξάρτητες τ.μ.

($\hat{\theta}$ είναι η ε.μ.π. του ρυθμού διαστολής του σύμπαντος)

(ii) Να λυθεί το πρόβλημα της ε.μ.π. για $Y_i = c + \theta X_i + \varepsilon_i$, όπου $\varepsilon_i \sim N(0, \sigma^2)$, όπου c, θ, σ^2 άγνωστες παράμετροι ($c \in \mathbb{R}$, $\theta \in \mathbb{R}$, $\sigma^2 > 0$).

(iii) Να συγκριθούν τα 2 μοντέλα στην προσαρμογή τους στα δεδομένα του Hubble με το κριτήριο AIC.

Λύση

(i) Είδαμε ότι

$$L(\theta, \sigma^2) = \prod_{i=1}^{\nu} f_{Y_i}(y_i ; \theta, \sigma^2) = (2\pi\sigma^2)^{-\frac{\nu}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{\nu}(y_i - \theta x_i)^2}$$

(τα Y_i, $1 \le i \le \nu$, είναι ανεξ. τ.μ., αλλά όχι ισόνομες).

$$\ell(\theta, \sigma^2) = -\frac{\nu}{2}\log(2\pi) - \frac{\nu}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{\nu}(Y_i - \theta x_i)^2 .$$

Εφαρμόζουμε διαδοχική μεγιστοποίηση.

Σταθεροποιώντας το σ^2 έχουμε $\max\limits_{\theta} \ell(\theta, \sigma^2) = \min\limits_{\theta} \sum_{i=1}^{\nu}(Y_i - \theta x_i)^2$

Η συνάρτηση $f(\theta) = \sum_{i=1}^{\nu}(Y_i - \theta x_i)^2$ είναι γνήσια κυρτή, άρα αν έχει στάσιμο σημείο θα είναι και ολικό ελάχιστο.

$f'(\theta) = -2 \sum_{i=1}^{\nu} x_i (Y_i - \theta x_i)^2$. Άρα $f'(\theta) = 0 \iff$

$$\theta^* = \frac{\sum_{i=1}^{\nu} x_i y_i}{\sum_{i=1}^{\nu} x_i^2} \Rightarrow \boxed{\hat{\theta} = \frac{\sum_{i=1}^{\nu} X_i Y_i}{\sum_{i=1}^{\nu} X_i^2}}$$ είναι η ε.μ.π. του θ.

Επίσης παίρνουμε $\boxed{\hat{\sigma}^2 = \frac{\sum_{i=1}^{\nu}(Y_i - \hat{\theta} X_i)^2}{\nu}}$

(ii) Έχουμε $\ell(c,\theta,\sigma^2) = -\frac{\nu}{2}\log(2\pi) - \frac{\nu}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{\nu}(y_i - c - \theta x_i)^2$.

Εφαρμόζουμε διαδοχική μεγιστοποίηση.

Σταθεροποιώντας το σ^2, έχουμε $\max\limits_{c,\theta} \ell(c,\theta,\sigma^2) = \min\limits_{c,\theta} \sum_{i=1}^{\nu}(y_i - c - \theta x_i)^2$

Θέτουμε $f(c,\theta) = \sum_{i=1}^{\nu}(y_i - c - \theta x_i)^2$.

Παρόμοια $\min\limits_{c\in\mathbb{R},\theta\in\mathbb{R}} \sum_{i=1}^{\nu}(y_i - c - \theta x_i)^2 = \min\limits_{\theta}\min\limits_{c} \sum_{i=1}^{\nu}(y_i - \theta x_i - c)^2$.

Άρα για σταθερό θ, αναζητούμε το ελάχιστο της

$g(c) = \sum_{i=1}^{\nu}(z_i - c)^2$, όπου $z_i = y_i - \theta x_i$, $\forall\, 1 \le i \le \nu$.

Η $g(c)$ είναι <u>γνήσια κυρτή</u>.

Γνωρίζουμε ότι ελαχιστοποιείται για $\boxed{c^* = \bar{z} = \bar{y} - \theta\bar{x}}$.

Παρατηρούμε εδώ ότι $c^* = c^*(\theta)$, δηλ. εξαρτάται από το θ.

Άρα $\min\limits_{c\in\mathbb{R},\theta\in\mathbb{R}} f(c,\theta) = \min\limits_{\theta\in\mathbb{R}} f(c^*(\theta),\theta) = \min\limits_{\theta\in\mathbb{R}} \sum_{i=1}^{\nu}(y_i - \bar{y} + \theta\bar{x} - \theta x_i)^2$

$= \min\limits_{\theta\in\mathbb{R}} \sum_{i=1}^{\nu}(\tilde{y}_i - \theta\tilde{x}_i)^2$, όπου $\tilde{y}_i = y_i - \bar{y}$, $\tilde{x}_i = x_i - \bar{x}$

(τα αντίστοιχα κεντροποιημένα δεδομένα).

Η συναρτησιακή μορφή τώρα είναι ίδια με το πρόβλημα (i)

$$\Rightarrow \boxed{\theta^* = \frac{\sum_{i=1}^{\nu}\tilde{x}_i\tilde{y}_i}{\sum_{i=1}^{\nu}(\tilde{x}_i)^2} = \frac{\sum_{i=1}^{\nu}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{\nu}(x_i - \bar{x})^2}} \Rightarrow$$

$\hat{\theta} = \theta^*$ και $\boxed{\hat{c} = \bar{y} - \hat{\theta}\bar{x}}$ Επίσης όπως πριν

$$\boxed{\hat{\sigma}^2 = \frac{\sum_{i=1}^{\nu}(y_i - \hat{\theta}x_i - \hat{c})^2}{\nu} = \frac{\sum_{i=1}^{\nu}\left[y_i - \bar{y} - \hat{\theta}(x_i - \bar{x})\right]^2}{\nu}}$$

Οι ακριβείς τιμές για τα δεδομένα Hubble και στα 2 φωτήματα δίνονται στο βοηθητικό αρχείο R που αντιστοιχεί στην Άσκηση 38.

(iii) Η επιλογή του καλύτερου μοντέλου με το κριτήριο AIC, δίνεται στο βοηθητικό αρχείο R. Υπάρχει επίσης βοηθητικό γράφημα που δείχνει τις ευθείες ελαχίστων τετραγώνων και στα 2 μοντέλα που αντιστοιχούν σε εκτιμώμενες τιμές των παραμέτρων με τη μέθοδο της μέγιστης πιθανοφάνειας.

Η μελέτη του αρχείου αυτού στην R, συνιστάτε ανεπιφύλακτα και βοηθά στην καλύτερη κατανόηση των αποτελεσμάτων, αλλά είναι και μία πρώτη επαφή/γνωριμία με το λογισμικό R πολύ κατάλληλο για Στατιστικές Αναλύσεις.

Εκτός απο το σχολιασμό των αποτελεσμάτων, σας προτείνω να επαληθεύσετε τα αποτελέσματα που δίνει η R στο παράδειγμα (ii), έχοντας ως πρότυπο, τον έλεγχο που έκανα στο ερώτημα (i). Μπορείτε να κάνετε το ίδιο και για τις τιμές του AIC, αλλά και της λογαριθμοπιθανοφάνειας, που μπορείτε να υπολογίσετε μόνοι σας απο τη θεωρία.

Μπορεί βέβαια το αρχείο αυτό να αποτελέσει και βάση για περισσότερο πειραματισμό απο τη δική σας πλευρά.

Άσκηση 39

Υπόδειξη: Δείτε τη σύνδεση της Άσκησης αυτής με την Άσκηση 36. Τώρα το α είναι σταθερό. Η διαδικασία επίλυσης είναι παρόμοια.