ΦΡΟΝΤΙΣΤΗΡΙΑ 7 ΚΑΙ 8

Άσκηση 1

Έστω X_1, X_2, ..., X_n iid τυχαίες μεταβλητές με $X_1 \sim Bernoulli(p_0),\ \ p_0 \in (0,1)$

Να εξάγετε

α) τη συνάρτηση πιθανοφάνειας στις 3 μορφές τις και

β) τον εκτιμητή μέγιστης πιθανοφάνειας για την άγνωστη παράμετρο p_0.

Δίνεται η συνάρτηση πιθανότητας :

$$f(x;p) = \begin{cases} p^x(1-p)^{1-x} & ,\ x \in \{0,1\} \\ 0 & ,\ \alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases}$$

Λύση

α) Δεδομένου ότι οι τυχαίες μεταβλητές είναι iid η συνάρτηση πιθανότητας του τυχαίου διανύσματος $\begin{pmatrix} X_1 \\ X_2 \\ ... \\ X_n \end{pmatrix}$ μπορεί να γραφτεί ως

$$f(\mathbf{x}) = \prod_{i=1}^{n} f(x_i;p) = f(x_1;p)f(x_2;p)...f(x_n;p) = p^{x_1}(1-p)^{1-x_1}\,p^{x_2}(1-p)^{1-x_2}\ ...\ p^{x_n}(1-p)^{1-x_n}$$

$$= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Η συνάρτηση πιθανοφάνειας θα είναι :

$$L(\mathbf{x},p) = \prod_{i=1}^{n} f(x_i,p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

η λογαριθμική συνάρτηση πιθανοφάνειας :

$$\ln L(\mathbf{x},p)=\ln\left(\prod_{i=1}^{n} f(x_i,p)\right)=\ln\left(p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}\right)=\sum_{i=1}^{n}x_i\ln p+n-\sum_{i=1}^{n}x_i\ln(1-p)$$

και τέλος η μέση λογαριθμική συνάρτηση πιθανοφάνειας :

$$\frac{\ln L(\mathbf{x},p)}{n}=\frac{\ln\left(\prod_{i=1}^{n} f(x_i,p)\right)}{n}=\frac{\ln\left(p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}\right)}{n}=\frac{\sum_{i=1}^{n}x_i\ln p+n-\sum_{i=1}^{n}x_i\ln(1-p)}{n}$$

β) Ο εκτιμητής μέγιστης πιθανοφάνειας για το p, θα πρέπει να ικανοποιεί τις συνθήκες πρώτης και δεύτερης τάξης. Από τις συνθήκες πρώτης τάξης έχουμε :

$$\frac{\partial \ln L(p)}{\partial p}=0\Rightarrow\frac{\sum_{i=1}^{n}x_i}{p}-\frac{n-\sum_{i=1}^{n}x_i}{1-p}=0\Leftrightarrow(1-p)\sum_{i=1}^{n}x_i-p\left(n-\sum_{i=1}^{n}x_i\right)=0\Leftrightarrow\hat{p}_{ML}=\frac{\sum_{i=1}^{n}x_i}{n}$$

Το $\hat{p}_{ML}=\frac{\sum_{i=1}^{n}x_i}{n}\in(0,1)$ ως άθροισμα n μεταβλητών που παίρνουν τιμές στο {0,1,...,M}

Αυτό είναι το κρίσιμο σημείο. Για να δούμε εάν είναι και μεγιστοποιούν θα πρέπει να πάρουμε και συνθήκες δεύτερης τάξης.

Αρκεί να δείξουμε ότι $\frac{\partial^2 \ln L(p)}{\partial p^2}\Big|_{p=\hat{p}_{ML}}<0$ ώστε το $\hat{p}_{ML}=\frac{\sum_{i=1}^{n}x_i}{n}$ να είναι ο εκτιμητής μέγιστης πιθανοφάνειας της άγνωστης παραμέτρου. Έχουμε :

$$\frac{\partial^2 \ln L(p)}{\partial p^2}\Bigg|_{p=\hat{p}_{ML}}=-\frac{\sum_{i=1}^{n}x_i}{p^2}-\frac{n-\sum_{i=1}^{n}x_i}{(1-p)^2}<0\ \ \forall p\in(0,1)\qquad \textit{καθώς}\ 0\le\sum_{i=1}^{n}x_i\le n$$

ως άθροισμα μεταβλητών που παίρνουν τιμή 0 ή 1. Επομένως, το

$$\hat{p}_{ML}=\frac{\sum_{i=1}^{n}x_i}{n}$$

είναι ολικό μέγιστο.

Άσκηση 2

Έστω X_1, X_2, ..., X_n iid τυχαίες μεταβλητές με $X_1 \sim Binomial(M, \theta_0), M > 0\ \theta \in (0,1)$

Να εξάγετε

α) τη συνάρτηση πιθανοφάνειας στις 3 μορφές τις και

β) τον εκτιμητή μέγιστης πιθανοφάνειας για την άγνωστη παράμετρο θ_0. Θεωρείστε το M γνωστό.

γ) Είναι αμερόληπτος ο εκτιμητής ;

Δίνεται η συνάρτηση πιθανότητας :

$$p(x;\theta,M) = \begin{cases} \binom{M}{x}\theta^x (1-\theta)^{M-x},\ x \in \{0,1,...,M\} \\ 0 \qquad\qquad\qquad\qquad ,\alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases}$$

και το διωνυμικό ανάπτυγμα :

$$(a+b)^m = \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} a^y b^{m-y}$$

Λύση

α) Δεδομένου ότι οι τυχαίες μεταβλητές είναι iid η συνάρτηση πιθανότητας του τυχαίου διανύσματος $\begin{pmatrix} X_1 \\ X_2 \\ ... \\ X_n \end{pmatrix}$ μπορεί να γραφτεί ως

$$P(\mathbf{x}) = \prod_{i=1}^{n} P(x_i, \theta) = \prod_{\iota=1}^{n} \binom{M}{x_i} \theta^{x_i} (1-\theta)^{M-x_i}$$

Η συνάρτηση πιθανοφάνειας θα είναι :

$$L_n(\theta)=\prod_{\iota=1}^{n}\binom{M}{x_i}\theta^{x_i}\left(1-\theta\right)^{M-x_i}$$

η λογαριθμική συνάρτηση πιθανοφάνειας :

$$l_n(\theta)=\ln\left(\prod_{\iota=1}^{n}\binom{M}{x_i}\theta^{x_i}\left(1-\theta\right)^{M-x_i}\right)=\sum_{i=1}^{n}\ln\binom{M}{x_i}+\ln\theta\sum_{i=1}^{n}x_i+\ln(1-\theta)(nM-\sum_{i=1}^{n}x_i)$$

και τέλος η μέση λογαριθμική συνάρτηση πιθανοφάνειας :

$$\overline{l}_n(\theta)=\frac{1}{n}\ln\left(\prod_{\iota=1}^{n}\binom{M}{x_i}\theta^{x_i}\left(1-\theta\right)^{M-x_i}\right)=\frac{1}{n}\left(\sum_{i=1}^{n}\ln\binom{M}{x_i}+\ln\theta\sum_{i=1}^{n}x_i+\ln(1-\theta)(nM-\sum_{i=1}^{n}x_i)\right)$$

β) Ο εκτιμητής μέγιστης πιθανοφάνειας για το θ, θα πρέπει να ικανοποιεί τις συνθήκες πρώτης και δεύτερης τάξης. Απο τις συνθήκες πρώτης τάξης έχουμε :

$$\frac{\partial l_n(\theta)}{\partial\theta}=0\Rightarrow\frac{\sum_{i=1}^{n}x_i}{\theta}-\frac{nM-\sum_{i=1}^{n}x_i}{1-\theta}=0\Leftrightarrow(1-\theta)\sum_{i=1}^{n}x_i-\theta nM-\sum_{i=1}^{n}x_i=0\Leftrightarrow\hat{\theta}_n=\frac{\sum_{i=1}^{n}x_i}{nM}$$

Το $\hat{\theta}_n=\frac{\sum_{i=1}^{n}x_i}{nM}\in[0,1]$ ως άθροισμα n μεταβλητών που παίρνουν τιμές στο {0,1,...,M}

Αυτό είναι το κρίσιμο σημείο. Για να δούμε εαν είναι και μεγιστοποιούν θα πρέπει να πάρουμε και συνθήκες δεύτερης τάξης.

Αρκεί να δείξουμε ότι $\left.\frac{\partial^2 l_n(\theta)}{\partial\theta^2}\right|_{\theta=\hat{\theta}_n}<0$ ώστε το $\hat{\theta}_n=\frac{\sum_{i=1}^{n}x_i}{nM}$ να είναι ο εκτιμητής μέγιστης πιθανοφάνειας της άγνωστης παραμέτρου. Έχουμε :

$$\left.\frac{\partial^2 l_n(\theta)}{\partial\theta^2}\right|_{\theta=\hat{\theta}_n}=-\frac{\sum_{i=1}^{n}x_i}{\theta^2}-\frac{nM-\sum_{i=1}^{n}x_i}{(1-\theta)^2}<0\qquad \kappa\alpha\theta\acute{\omega}\varsigma\ 0\le x_i\le M\Leftrightarrow 0\le\sum_{i=1}^{n}x_i\le nM$$

γ) Ο εκτιμητής είναι αμερόληπτος εάν

$$E(\hat{p}_{ML})=p_0$$

Έχουμε :

$$E(\hat{p}_{ML}) = E\left(\frac{\sum_{i=1}^{n} x_i}{nM}\right) = \frac{\sum_{i=1}^{n} E(x_i)}{nM}$$

Παρατηρώ ότι προκύπτει η ανάγκη υπολογισμού της πρώτης ροπής της κατανομής. Έχουμε :

$$E(X) = \sum_{x=0}^{M} x \binom{M}{x} p^x (1-p)^{M-x} = \sum_{x=1}^{M} x \binom{M}{x} p^x (1-p)^{M-x} = \sum_{x=1}^{M} x \frac{M!}{x!(M-x)!} p^x (1-p)^{M-x}$$

$$= \sum_{x=1}^{M} \frac{M!}{(x-1)!(M-x)!} p^x (1-p)^{M-x}$$

Εφόσον ο πρώτος όρος (για x=0) ισούται με το μηδέν.

Θέτω $y = x - 1 \;\kappa\alpha\iota\; n = M - 1$.

Συνεπώς :

$$E(X) = \sum_{y=o}^{n} \frac{(n+1)!}{(y)!(n-y)!} p^{y+1} (1-p)^{n-y} = (n+1) p \sum_{y=o}^{n} \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}$$

$$Mp \sum_{y=o}^{n} \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}$$

Παρατηρώ, πως αν αντικαταστήσω στο διωνυμικό ανάπτυγμα $a = p, b = 1 - p$, έχω :

$$\sum_{y=o}^{n} \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y} = \left[p + (1-p)\right]^n = 1$$

Επομένως

$$E(X) = Mp$$

Έχουμε λοιπόν πως :

$$E(\hat{p}_{ML}) = E\left(\frac{\sum_{i=1}^{n} x_i}{nM}\right) = \frac{\sum_{i=1}^{n} E(x_i)}{nM} = \frac{\sum_{i=1}^{n} Mp_0}{nM} = \frac{nMp_0}{nM} = p_0$$

και επομένως ο εκτιμητής είναι αμερόληπτος.

Άσκηση 3

Έστω $X_1, X_2, ..., X_n$ iid τυχαίες μεταβλητές με $X_1 \sim Poisson(\lambda_0)\ ,\lambda_0 > 0$

Να εξάγετε

α) τη συνάσtηση πιθανοφάνειας στις 3 μορφές τις και

β) τον εκτιμητή μέγιστης πιθανοφάνειας για την άγνωστη παράμετρο λ_0.

Δίνεται η συνάρτηση πιθανότητας :

$$p(x;\theta,\mathrm{M}) = \begin{cases} \dfrac{\lambda^x}{x!} e^{-\lambda} & , x \in \{0,1,2,...\} \\ 0 & ,\alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases}$$

Λύση

α) Δεδομένου ότι οι τυχαίες μεταβλητές είναι iid η συνάρτηση πιθανότητας του τυχαίου διανύσματος $\begin{pmatrix} X_1 \\ X_2 \\ ... \\ X_n \end{pmatrix}$ μπορεί να γραφτεί ως

$$P(\mathbf{x}) = \prod_{i=1}^{n} p(x_i;\lambda) = \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} e^{-n\lambda}$$

Η συνάρτηση πιθανοφάνειας θα είναι :

$$L(\mathbf{x},\lambda) = \prod_{i=1}^{n} p(x_i,\lambda) = \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} e^{-n\lambda}$$

η λογαριθμική συνάρτηση πιθανοφάνειας :

$$\ln L(\mathbf{x},\lambda)=\ln\left(\prod_{i=1}^{n}p(x_i,\lambda)\right)=\ln\left(\frac{\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}e^{-n\lambda}\right)=\sum_{i=1}^{n}x_i\ln\lambda-n\lambda-\ln\prod_{i=1}^{n}x_i!$$

και τέλος η μέση λογαριθμική συνάρτηση πιθανοφάνειας :

$$\frac{\ln L(\mathbf{x},\lambda)}{n}=\frac{\ln\left(\prod_{i=1}^{n}p(x_i,\lambda)\right)}{n}=\frac{\ln\left(\frac{\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}e^{-n\lambda}\right)}{n}=\frac{\sum_{i=1}^{n}x_i\ln\lambda-n\lambda-\ln\prod_{i=1}^{n}x_i!}{n}$$

β) Ο εκτιμητής μέγιστης πιθανοφάνειας για το λ, θα πρέπει να ικανοποιεί τις συνθήκες πρώτης και δεύτερης τάξης. Απο τις συνθήκες πρώτης τάξης έχουμε :

$$\frac{\partial\ln L(\lambda)}{\partial\lambda}=0\Rightarrow\frac{\sum_{i=1}^{n}x_i}{\lambda}-n=0\Leftrightarrow\hat{\lambda}_{ML}=\frac{\sum_{i=1}^{n}x_i}{n}$$

Το $\hat{\lambda}_{ML}=\frac{\sum_{i=1}^{n}x_i}{n}>0$ ως άθροισμα n μεταβλητών που παίρνουν τιμές στο {0,1,...,}

Αυτό είναι το κρίσιμο σημείο. Για να δούμε εαν είναι και μεγιστοποιούν θα πρέπει να πάρουμε και συνθήκες δεύτερης τάξης.

Αρκεί να δείξουμε ότι $\left.\frac{\partial^2\ln L(\lambda)}{\partial\lambda^2}\right|_{\lambda=\hat{\lambda}_{ML}}<0$ ώστε το $\hat{\lambda}_{ML}=\frac{\sum_{i=1}^{n}x_i}{n}$ να είναι ο εκτιμητής μέγιστης πιθανοφάνειας της άγνωστης παραμέτρου. Έχουμε :

$$\left.\frac{\partial^2\ln L(\lambda)}{\partial\lambda^2}\right|_{\lambda=\hat{\lambda}_{ML}}=-\frac{\sum_{i=1}^{n}x_i}{\hat{\lambda}_{ML}{}^2}<0$$ για n αρκούντως μεγάλο.

Άσκηση 4

Έστω X_1, X_2, ..., X_n iid τυχαίες μεταβλητές με $X_1 \sim Gamma(a_0, b_0), a_0, b_0 > 0$ για $\forall i = 1,...,n$.
Θεωρήστε γνωστό το $a_0 = 2$ και άγνωστη την παράμετρο $b_0 > 0$.

α) Να εξαχθούν οι συναρτήσεις πιθανοφάνειας (και στις 3 μορφές)

β) Να εξαχθεί ο εκτιμητής μέγιστης πιθανοφάνειας (MLE) για το b_0

γ) Να δείξετε ότι $E(X^k) = \dfrac{\Gamma(\alpha + k)}{\Gamma(\alpha)} b^{-k}$ και να εξετάσετε αν ο εκτιμητής είναι αμερόληπτος.

Δίνεται η pdf της Γάμμα κατανομής :

$$f(x;a,b) = \begin{cases} \dfrac{b^a}{\Gamma(\alpha)} x^{a-1} e^{-bx}, x \geq 0 \\ 0 \qquad\qquad\qquad , \alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases}$$

και η συνάρτηση Γάμμα :

$$\int_0^\infty t^{\alpha-1} e^{-t} dt = \Gamma(a) = (a-1)!$$

Λύση

Αφού $a_0 = 2$ η συνάρτηση πυκνότητας γίνεται :

$$f(x;b) = \begin{cases} b^2 x e^{-bx} & , x \geq 0 \\ 0 & , \alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases}$$

α) Παρατηρώ ότι το στήριγμα είναι το suppP=$[0,\infty)$ (αφού αυτό είναι το μικρότερο διάστημα στο οποίο όταν ολοκληρώνω τη συνάρτηση πυκνότητας παίρνω μονάδα).

Δεδομένου ότι οι τυχαίες μεταβλητές είναι iid η συνάρτηση πυκνότητας του τυχαίου διανύσματος $\begin{pmatrix} X_1 \\ X_2 \\ ... \\ X_n \end{pmatrix}$ μπορεί να γραφτεί ως

$$f(\mathbf{x}) = \prod_{i=1}^{n} f(x_i, b) = \prod_{i=1}^{n} b^2 x_i e^{-bx_i}$$

Η συνάρτηση πιθανοφάνειας θα είναι :

$$L_n(b) = \prod_{i=1}^{n} f(x_i, b) = \prod_{i=1}^{n} b^2 x_i e^{-bx_i}$$

η λογαριθμική συνάρτηση πιθανοφάνειας :

$$l_n(b) = \ln\left(\prod_{i=1}^{n} f(x_i, b)\right) = \sum_{i=1}^{n} \ln\left(b^2 x_i e^{-bx_i}\right) = 2\sum_{i=1}^{n} \ln b + \sum_{i=1}^{n} \ln x_i - b\sum_{i=1}^{n} x_i = 2n\ln b + \sum_{i=1}^{n} \ln x_i - b\sum_{i=1}^{n} x_i$$

και τέλος η μέση λογαριθμική συνάρτηση πιθανοφάνειας :

$$\overline{l}_n(b) = \frac{1}{n}\left(\ln\left(\prod_{i=1}^{n} f(x_i, b)\right)\right) = \frac{1}{n}\left(2n\ln b + \sum_{i=1}^{n} \ln x_i - b\sum_{i=1}^{n} x_i\right) = 2\ln b + \frac{\sum_{i=1}^{n} \ln x_i - b\sum_{i=1}^{n} x_i}{n}$$

Ο εκτιμητής μέγιστης πιθανοφάνειας για το θ, θα πρέπει να ικανοποιεί τις συνθήκες πρώτης και δεύτερης τάξης. Από τις συνθήκες πρώτης τάξης έχουμε :

$$\frac{\partial \overline{l}_n(b)}{\partial b} = 0 \Rightarrow \frac{2}{b} - \frac{\sum_{i=1}^{n} x_i}{n} = 0 \Leftrightarrow \hat{b}_n = \frac{2n}{\sum_{i=1}^{n} x_i}$$

Το $\hat{b}_n$ είναι το κρίσιμο σημείο. Για να δούμε εάν είναι και μεγιστοποιούν θα πρέπει να πάρουμε και συνθήκες δεύτερης τάξης.

Αρκεί να δείξουμε ότι $\left.\frac{\partial^2 \overline{l}_n(b)}{\partial b^2}\right|_{b=\hat{b}_n} < 0$ ώστε το $\hat{b}_n = \frac{2n}{\sum_{i=1}^{n} x_i}$ να είναι ο εκτιμητής μέγιστης πιθανοφάνειας της άγνωστης παραμέτρου. Έχουμε :

$$\left.\frac{\partial^2 \overline{l}_n(b)}{\partial b^2}\right|_{b=\hat{b}_n} = -\frac{2}{\left(\hat{b}_n\right)^2} < 0$$

Τέλος, παρατηρώ ότι $\hat{b}_n = \dfrac{2n}{\sum_{i=1}^{n} x_i} > 0$ αφού $n \in N$ και $x_i \in R^+$ και συνεπώς οι εκτιμήσεις που δίνει βρίσκονται εντός του παραμετρικού χώρου.

Συμπερασματικά, ο $\hat{b}_n$ είναι ο εκτιμητής μέγιστης πιθανοφάνειας για την άγνωστη παράμετρο b_0.

γ) (Για την απόδειξη $E(X^k) = \dfrac{\Gamma(\alpha+k)}{\Gamma(\alpha)} b^{-k}$, δείτε φροντιστήριο 7, Άσκηση 5)

Γνωρίζουμε ότι ο εκτιμητής είναι αμερόληπτος εάν ισχύει ότι $E(\hat{b}_n) = b_0$. Έχουμε :

$$E(\hat{b}_n) = E\left(\frac{2n}{\sum_{i=1}^{n} x_i}\right) \neq \frac{2n}{\sum_{i=1}^{n} Ex_i} = \frac{2n}{\sum_{i=1}^{n} \frac{2}{b_0}} = \frac{2n}{\frac{2n}{b_0}} = b_0$$

αφού η πρώτη ροπή , δίνεται από το $E(X^k) = \dfrac{\Gamma(\alpha+k)}{\Gamma(\alpha)} b^{-k}$ για k=1 :

$$E(X) = \frac{(b_0)^{-1}}{\Gamma(2)!}\Gamma(2+1)! = \frac{2!}{1!b_0} = \frac{2}{b_0}$$

Συνεπώς, $E(\hat{b}_n) \neq b_0$, άρα ο εκτιμητής είναι μεροληπτικός