Πιθανότητες - Στατιστική Τμ. Πληροφορικής & Τηλεπικοινωνιών, 4/9/2015
Θέματα - Λύσεις

Θέμα 1ο: (3 βαθμοί) Έστω (X, Y) διδιάστατη συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας

$$f_{X,Y}(x,y) = \begin{cases} cxy^2, & \text{όταν } 0 \leq x \leq y \leq 2, \\ 0, & \text{διαφορετικά,} \end{cases}$$

όπου c σταθερά. Να υπολογιστούν:

(1) η σταθερά c,
(2) οι περιθώριες συναρτήσεις πυκνότητας πιθανότητας $f_X(x)$ και $f_Y(y)$ των X και Y αντίστοιχα,
(3) η πιθανότητα $P(X \leq 1)$,
(4) η μέση τιμή $E[2X-1]$,
(5) η δεσμευμένη μέση τιμή $E[X|Y=1]$,
(6) η διασπορά $Var[Y]$.

Λύση: (1) Έχουμε:

$$\begin{aligned}
& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1 \\
\Rightarrow\ & \int_0^2\int_0^y cxy^2 dxdy = 1 \\
\Rightarrow\ & c\int_0^2 y^2 \int_0^y x dxdy = 1 \\
\Rightarrow\ & c\int_0^2 y^2 \left[\frac{x^2}{2}\right]_{x=0}^{x=y} dy = 1 \\
\Rightarrow\ & \frac{c}{2}\int_0^2 y^4 dy = 1 \\
\Rightarrow\ & \frac{c}{2}\left[\frac{y^5}{5}\right]_{y=0}^{y=2} = 1 \\
\Rightarrow\ & \frac{32c}{10} = 1 \\
\Rightarrow\ & c = \frac{10}{32} = \frac{5}{16}.
\end{aligned}$$

(2) Επίσης

$$f_X(x) \ = \ \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy.$$

Οπότε για $x \notin [0,2]$ έχουμε $f_X(x) = 0$, ενώ για $x \in [0,2]$ έχουμε

$$\begin{aligned} f_X(x) &= \int_x^2 f_{X,Y}(x,y)dy \\ &= \int_x^2 \frac{5xy^2}{16} dy \\ &= \frac{5x}{16} \int_x^2 y^2 dy \\ &= \frac{5x}{16} \left[\frac{y^3}{3} \right]_{y=x}^{y=2} \\ &= \frac{5x(8-x^3)}{48}. \end{aligned}$$

Τελικά:

$$f_X(x) = \begin{cases} \frac{5x(8-x^3)}{48}, & \text{αν } x \in [0,2], \\ 0, & \text{διαφορετικά.} \end{cases}$$

Ομοίως:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx.$$

Οπότε για $y \notin [0,2]$ έχουμε $f_Y(y) = 0$, ενώ για $y \in [0,2]$ έχουμε

$$\begin{aligned} f_Y(y) &= \int_0^y f_{X,Y}(x,y)dx \\ &= \int_0^y \frac{5xy^2}{16} dx \\ &= \frac{5y^2}{16} \int_0^y x dx \\ &= \frac{5y^2}{16} \left[\frac{x^2}{2} \right]_{x=0}^{x=y} \\ &= \frac{5y^4}{32}. \end{aligned}$$

Τελικά:

$$f_Y(y) = \begin{cases} \frac{5y^4}{32}, & \text{αν } y \in [0,2], \\ 0, & \text{διαφορετικά.} \end{cases}$$

(3) Έχουμε

$$\begin{aligned} P(X \leq 1) &= \int_{-\infty}^1 f_X(x)dx \\ &= \int_0^1 \frac{5x(8-x^3)}{48} dx \\ &= \frac{40}{48} \int_0^1 x dx - \frac{5}{48} \int_0^1 x^4 dx \\ &= \frac{40}{48} \left[\frac{x^2}{2} \right]_{x=0}^{x=1} - \frac{5}{48} \left[\frac{x^5}{5} \right]_{x=0}^{x=1} \\ &= \frac{19}{48}. \end{aligned}$$

(4) Είναι $E[2X-1] = 2E[X]-1$, οπότε βρίσκουμε πρώτα την $E[X]$. Είναι

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x)dx \\ &= \int_0^2 \frac{5x^2(8-x^3)}{48}dx \\ &= \frac{40}{48}\int_0^2 x^2 dx - \frac{5}{48}\int_0^2 x^5 dx \\ &= \frac{40}{48}\left[\frac{x^3}{3}\right]_{x=0}^{x=2} - \frac{5}{48}\left[\frac{x^6}{6}\right]_{x=0}^{x=2} \\ &= \frac{10}{9}. \end{aligned}$$

Άρα:

$$E[2X-1] = 2E[X]-1 = \frac{11}{9}.$$

(5) Για τη δεσμευμένη συνάρτηση πυκνότητας $f_{X|Y}(x|y)$ έχουμε

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \begin{cases} \frac{5xy^2/16}{5y^4/32}, & \text{αν } 0 \le x \le y \le 2, \\ 0, & \text{διαφορετικά,} \end{cases} \\ &= \begin{cases} \frac{2x}{y^2}, & \text{αν } 0 \le x \le y \le 2, \\ 0, & \text{διαφορετικά.} \end{cases} \end{aligned}$$

Γενικά είναι

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx,$$

οπότε

$$E[X|Y=1] = \int_0^1 2x^2 dx = \left[\frac{2x^3}{3}\right]_{x=0}^{x=1} = \frac{2}{3}.$$

(6) Για τη διασπορά της Y έχουμε

$$Var[Y] = E[Y^2] - (E[Y])^2.$$

Είναι

$$\begin{aligned} E[Y^2] &= \int_{-\infty}^{\infty} y^2 f_Y(y)dy \\ &= \int_0^2 \frac{5y^6}{32}dy \\ &= \frac{5}{32}\int_0^2 y^6 dy \\ &= \frac{5}{32}\left[\frac{y^7}{7}\right]_{y=0}^{y=2} \\ &= \frac{20}{7} \end{aligned}$$

και

$$\begin{aligned} E[Y] &= \int_{-\infty}^{\infty} y f_Y(y) dy \\ &= \int_0^2 \frac{5y^5}{32} dy \\ &= \frac{5}{32} \int_0^2 y^5 dy \\ &= \frac{5}{32} \left[\frac{y^6}{6} \right]_{y=0}^{y=2} \\ &= \frac{5}{3}, \end{aligned}$$

οπότε

$$Var[Y] = \frac{20}{7} - \frac{25}{9} = \frac{5}{63}.$$

Θέμα 2ο: (3 βαθμοί) Μια εταιρεία έχει 3 καταστήματα Α,Β και Γ, με 50, 75 και 100 υπαλλήλους αντίστοιχα. Οι γυναίκες υπάλληλοι στα Α,Β και Γ είναι 25, 45 και 70 αντίστοιχα. Για ένα εκπαιδευτικό σεμινάριο επιλέγεται τυχαία ένα κατάστημα (το καθένα με πιθανότητα 1/3) και εν συνεχεία δύο διακεκριμένοι υπάλληλοι από αυτό. Έστω X το πλήθος των υπαλληλών του καταστήματος που επιλέχθηκε και Y το πλήθος των γυναικών που επιλέχθηκαν. Να υπολογιστούν:

(1) η πιθανότητα $P(X = 50, Y = 2)$,
(2) η πιθανότητα $P(X = 50, Y = 1)$,
(3) η δεσμευμένη μέση τιμή $E[Y|X = 50]$,
(4) η μέση τιμή $E[X]$,
(5) η πιθανότητα $P(Y = 2)$,
(6) η δεσμευμένη πιθανότητα $P(X = 50|Y = 2)$.

Λύση: (1) Έχουμε, χρησιμοποιώντας τον πολλαπλασιαστικό νόμο

$$\begin{aligned} P(X = 50, Y = 2) &= P(\text{Επιλέγεται το κατ. Α, Επιλέγεται γυναίκα, Επιλέγεται γυναίκα}) \\ &= \frac{1}{3} \cdot \frac{25}{50} \cdot \frac{24}{49} = \frac{4}{49}. \end{aligned}$$

(2) Έχουμε:

$$\begin{aligned} P(X = 50, Y = 1) &= P(\text{Επιλέγεται το κατ. Α, Επιλέγεται άνδρας, Επιλέγεται γυναίκα}) \\ &\quad + P(\text{Επιλέγεται το κατ. Α, Επιλέγεται γυναίκα, Επιλέγεται άνδρας}) \\ &= \frac{1}{3} \cdot \frac{25}{50} \cdot \frac{25}{49} + \frac{1}{3} \cdot \frac{25}{50} \cdot \frac{25}{49} = \frac{25}{147}, \end{aligned}$$

με χρήση του θεωρήματος ολικής πιθανότητας και του πολλαπλασιαστικού νόμου.

(3) Έχουμε:

$$\begin{aligned} P(Y = 0|X = 50) &= P(\text{Επιλέγεται άνδρας, Επιλέγεται άνδρας}|\text{Επιλέχθηκε το κατ. Α}) \\ &= \frac{25}{50} \cdot \frac{24}{49} = \frac{12}{49}. \end{aligned}$$

Ομοίως

$$\begin{aligned} P(Y = 2|X = 50) &= P(\text{Επιλέγεται γυναίκα, Επιλέγεται γυναίκα}|\text{Επιλέχθηκε το κατ. Α}) \\ &= \frac{25}{50} \cdot \frac{24}{49} = \frac{12}{49}. \end{aligned}$$

Τέλος, αφού οι δυνατές τιμές της τ.μ. Y είναι μόνο οι 0,1,2, έχουμε

$$P(Y=1|X=50) \;=\; 1-P(Y=0|X=50)-P(Y=2|X=50)=\frac{25}{49}.$$

Άρα

$$E[Y|X=50]=\sum_y yP(Y=y|X=50)=0\cdot\frac{12}{49}+1\cdot\frac{25}{49}+2\cdot\frac{12}{49}=1.$$

(4) Έχουμε $P(X=50)=P(X=75)=P(X=100)=\frac{1}{3}$, οπότε

$$E[X]=\sum_x xP(X=x)=50\cdot\frac{1}{3}+75\cdot\frac{1}{3}+100\cdot\frac{1}{3}=75.$$

(5) Με βάση το θεώρημα ολικής πιθανότητας έχουμε:

$$\begin{aligned}
P(Y=2) \;&=\; P(\text{Επιλέγεται το κατ. Α})P(\text{Επιλέγ. Γυν., Γυν.}|\text{Επιλέχθ. το Α})\\
&\quad +P(\text{Επιλέγεται το κατ. Β})P(\text{Επιλέγ. Γυν., Γυν.}|\text{Επιλέχθ. το Β})\\
&\quad +P(\text{Επιλέγεται το κατ. Γ})P(\text{Επιλέγ. Γυν., Γυν.}|\text{Επιλέχθ. το Γ})\\
&=\; \frac{1}{3}\cdot\frac{25}{50}\cdot\frac{24}{49}+\frac{1}{3}\cdot\frac{45}{75}\cdot\frac{44}{74}+\frac{1}{3}\cdot\frac{70}{100}\cdot\frac{69}{99}.
\end{aligned}$$

(6) Με βάση τον κανόνα του Bayes έχουμε:

$$\begin{aligned}
P(X=50|Y=2) \;&=\; \frac{P(X=50,Y=2)}{P(Y=2)}\\
&=\; \frac{\frac{4}{49}}{\frac{1}{3}\cdot\frac{25}{50}\cdot\frac{24}{49}+\frac{1}{3}\cdot\frac{45}{75}\cdot\frac{44}{74}+\frac{1}{3}\cdot\frac{70}{100}\cdot\frac{69}{99}}.
\end{aligned}$$

Θέμα 3ο: (3 βαθμοί) (1) Έστω $X_1,X_2,\ldots,X_{100}$ ανεξ. ισόνομες συνεχείς τ.μ. με την ομοιόμορφη κατανομή στο $[0,1]$ με μέση τιμή $E[X_i]=\frac{1}{2}$ και $E[X_i^2]=\frac{1}{3}$. Να βρεθεί προσεγγιστικά η πιθανότητα το άθροισμα τους να είναι στο διάστημα $[50-\frac{10}{\sqrt{3}},50+\frac{5}{\sqrt{3}}]$ (Δίνονται $\Phi(1)=0.8413$, $\Phi(1.5)=0.9332$, $\Phi(2)=0.9773$).

(2) Δίνεται τυχαίο δείγμα $Y_1,Y_2,\ldots,Y_{100}$ από την κατανομή με συνάρτηση πυκνότητας πιθανότητας $f_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2}}$. Να βρεθεί εκτιμήτρια μέγιστης πιθανοφάνειας $\hat{\mu}$ για το μ.

Λύση: (1) Έχουμε ότι $Var[X_i]=E[X_i^2]-(E[X_i])^2=\frac{1}{12}$. Έστω $S_{100}=\sum_{i=1}^{100}X_i$. Τότε $E[S_{100}]=100E[X_i]=50$ και $Var[S_{100}]=100Var[X_i]=\frac{25}{3}$. Από το κεντρικό οριακό θεώρημα έχουμε ότι η $\frac{S_{100}-E[S_{100}]}{\sqrt{Var[S_{100}]}}$ ακολουθεί προσεγγιστικά την τυποποιημένη κανονική κατανομή. Επομένως έχουμε:

$$\begin{aligned}
P\left(50-\frac{10}{\sqrt{3}}\le S_{100}\le 50+\frac{5}{\sqrt{3}}\right) \;&=\; P\left(-2\le\frac{S_{100}-E[S_{100}]}{\sqrt{Var[S_{100}]}}\le 1\right)\\
&\simeq\; P(-2\le Z\le 1),\; Z\sim\mathcal{N}(0,1)\\
&=\; \Phi(1)-\Phi(-2)\\
&\overset{\Phi(-z)+\Phi(-z)=1}{=}\; \Phi(1)+\Phi(2)-1\\
&=\; 0.8186.
\end{aligned}$$

(2) Η συνάρτηση πιθανοφάνειας είναι

$$\begin{aligned} f(y_1, y_2, \ldots, y_{100}; \mu) &= \prod_{i=1}^{100} f_{Y_i}(y_i) \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^{100} e^{-\sum_{i=1}^{100} \frac{(y_i-\mu)^2}{2}}. \end{aligned}$$

Για να τη βελτιστοποιήσουμε θεωρούμε το λογάριθμο της πιθανοφάνειας

$$\log f(y_1, y_2, \ldots, y_{100}; \mu) = -100 \log(\sqrt{2\pi}) - \sum_{i=1}^{100} \frac{(y_i-\mu)^2}{2}.$$

Έχουμε

$$\frac{d}{d\mu} \log f(y_1, y_2, \ldots, y_{100}; \mu) = \sum_{i=1}^{100} (y_i - \mu) = \sum_{i=1}^{100} y_i - 100\mu.$$

Λύνοντας την $\frac{d}{d\mu} \log f(y_1, y_2, \ldots, y_{100}; \mu) = 0$ παίρνουμε την υποψήφια εκτίμηση μέγιστης πιθανοφάνειας

$$\hat{\mu} = \frac{\sum_{i=1}^{100} y_i}{100},$$

η οποία αντιστοιχεί πράγματι σε μέγιστο του λογάριθμου πιθανοφάνειας, αφού

$$\frac{d^2}{d\mu^2} \log f(y_1, y_2, \ldots, y_{100}; \mu) = -100 < 0.$$

Επομένως η εκτιμήτρια μέγιστης πιθανοφάνειας είναι η

$$\hat{\mu} = \frac{\sum_{i=1}^{100} Y_i}{100}.$$

Θέμα 4ο: (2 βαθμοί) Ένας ερασιτέχνης ψαράς ψαρεύει με καλάμι σε μια λίμνη και κάθε φορά που πιάνει ένα ψάρι είναι εξίσου πιθανό να είναι ένα από τα 5 είδη ψαριών Ψ1, Ψ2, Ψ3, Ψ4, Ψ5 που υπάρχουν σε αυτή τη λίμνη.

(1) Να βρεθεί η πιθανότητα να έχει ψαρέψει τουλάχιστον μια φορά από το είδος Ψ1, αν έχει πιάσει 8 ψάρια.

(2) Να βρεθεί ο αναμενόμενος αριθμός διαφορετικών ειδών ψαριών που θα έχει μαζέψει, αν έχει πιάσει 8 ψάρια.

Λύση: (1) Έχουμε:

$$\begin{aligned} & P(\text{τουλάχιστον ένα } \Psi 1 \text{ μεταξύ των 8 ψαριών}) \\ = \ & 1 - P(\text{κανένα } \Psi 1 \text{ μεταξύ των 8 ψαριών}) \\ = \ & 1 - P(\text{το 1ο όχι } \Psi 1, \text{ το 2ο όχι } \Psi 1, \ldots, \text{ το 8ο όχι } \Psi 1) \\ = \ & 1 - \left(\frac{4}{5}\right)^8. \end{aligned}$$

(2) Έστω τυχαίες μεταβλητές X_i, $i = 1, 2, 3, 4, 5$ με

$$X_i = \begin{cases} 1, & \text{αν ένα τουλάχιστον } \Psi\text{i μεταξύ των 8 ψαριών}, \\ 0, & \text{διαφορετικά.} \end{cases}$$

Από το (1) έχουμε ότι $E[X_1] = P(X_1 = 1) = 1 - \left(\frac{4}{5}\right)^8$ και το ίδιο ισχύει για όλα τα $E[X_i]$, $i = 1, 2, 3, 4, 5$. Ο αναμενόμενος αριθμός διαφορετικών ειδών ψαριών που θα έχει μαζέψει αν έχει πιάσει 8 ψάρια είναι $E[X_1 + X_2 + X_3 + X_4 + X_5] = 5\left(1 - \left(\frac{4}{5}\right)^8\right)$.