Ονοματεπώνυμο εξεταζομένου:
Αριθμός μητρώου:

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

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

$$f_{X,Y}(x,y) = \begin{cases} cx^2y, & \text{όταν } 0 \leq x \leq y \leq 1, \\ 0, & \text{διαφορετικά,} \end{cases}$$

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 διασπορά $Var[Y]$,
(4) η μέση τιμή $E[3X-2]$,
(5) η πιθανότητα $P(X \leq \frac{1}{2})$,
(6) η δεσμευμένη μέση τιμή $E[X|Y=\frac{1}{2}]$.

Λύση: (1) Έχουμε:

$$\begin{aligned} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1 \\ \Rightarrow \quad & \int_0^1\int_0^y cx^2ydxdy = 1 \\ \Rightarrow \quad & c\int_0^1 y\int_0^y x^2dxdy = 1 \\ \Rightarrow \quad & c\int_0^1 y\left[\frac{x^3}{3}\right]_{x=0}^{x=y} dy = 1 \\ \Rightarrow \quad & \frac{c}{3}\int_0^1 y^4dy = 1 \\ \Rightarrow \quad & \frac{c}{3}\left[\frac{y^5}{5}\right]_{y=0}^{y=1} = 1 \\ \Rightarrow \quad & \frac{c}{15} = 1 \\ \Rightarrow \quad & c = 15. \end{aligned}$$

(2) Επίσης

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

Οπότε για $x \notin [0,1]$ έχουμε $f_X(x) = 0$, ενώ για $x \in [0,1]$ έχουμε

$$\begin{aligned} f_X(x) &= \int_x^1 f_{X,Y}(x,y)dy \\ &= \int_x^1 15x^2ydy \\ &= 15x^2\int_x^1 ydy \\ &= 15x^2\left[\frac{y^2}{2}\right]_{y=x}^{y=1} \\ &= \frac{15x^2(1-x^2)}{2}. \end{aligned}$$

Τελικά:

$$f_X(x) = \begin{cases} \frac{15x^2(1-x^2)}{2}, & \text{αν } x \in [0,1], \\ 0, & \text{διαφορετικά.} \end{cases}$$

Ομοίως:

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

Οπότε για $y \notin [0,1]$ έχουμε $f_Y(y) = 0$, ενώ για $y \in [0,1]$ έχουμε

$$\begin{aligned} f_Y(y) &= \int_0^y f_{X,Y}(x,y)dx \\ &= \int_0^y 15x^2 y dx \\ &= 15y \int_0^y x^2 dx \\ &= 15y \left[\frac{x^3}{3}\right]_{x=0}^{x=y} \\ &= 5y^4. \end{aligned}$$

Τελικά:

$$f_Y(y) = \begin{cases} 5y^4, & \text{αν } y \in [0,1], \\ 0, & \text{διαφορετικά.} \end{cases}$$

(6) Για τη διασπορά της Y έχουμε

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

Είναι

$$\begin{aligned} E[Y^2] &= \int_{-\infty}^{\infty} y^2 f_Y(y)dy \\ &= \int_0^1 5y^6 dy \\ &= 5\int_0^1 y^6 dy \\ &= 5\left[\frac{y^7}{7}\right]_{y=0}^{y=1} \\ &= \frac{5}{7} \end{aligned}$$

και

$$\begin{aligned} E[Y] &= \int_{-\infty}^{\infty} y f_Y(y) dy \\ &= \int_0^1 5y^5 dy \\ &= 5\int_0^1 y^5 dy \\ &= 5\left[\frac{y^6}{6}\right]_{y=0}^{y=1} \\ &= \frac{5}{6}, \end{aligned}$$

οπότε

$$Var[Y] = \frac{5}{7} - \frac{25}{36} = \frac{5}{252}.$$

(4) Είναι $E[3X-2] = 3E[X]-2$, οπότε βρίσκουμε πρώτα την $E[X]$. Είναι

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_0^1 \frac{15x^3(1-x^2)}{2} dx \\ &= \frac{15}{2}\int_0^1 x^3 dx - \frac{15}{2}\int_0^1 x^5 dx \\ &= \frac{15}{2}\left[\frac{x^4}{4}\right]_{x=0}^{x=1} - \frac{15}{2}\left[\frac{x^6}{6}\right]_{x=0}^{x=1} \\ &= \frac{15}{24}. \end{aligned}$$

Άρα:

$$E[3X-2] = 3E[X] - 2 = \frac{21}{24} = \frac{7}{8}.$$

(5) Έχουμε

$$\begin{aligned} P(X \leq \frac{1}{2}) &= \int_{-\infty}^{\frac{1}{2}} f_X(x) dx \\ &= \int_0^{\frac{1}{2}} \frac{15x^2(1-x^2)}{2} dx \\ &= \frac{15}{2}\int_0^{\frac{1}{2}} x^2 dx - \frac{15}{2}\int_0^{\frac{1}{2}} x^4 dx \\ &= \frac{15}{2}\left[\frac{x^3}{3}\right]_{x=0}^{x=\frac{1}{2}} - \frac{15}{2}\left[\frac{x^5}{5}\right]_{x=0}^{x=\frac{1}{2}} \\ &= \frac{17}{32}. \end{aligned}$$

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

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \begin{cases} \frac{15x^2y}{5y^4}, & \text{αν } 0 \le x \le y \le 1, \\ 0, & \text{διαφορετικά,} \end{cases} \\ &= \begin{cases} \frac{3x^2}{y^3}, & \text{αν } 0 \le x \le y \le 1, \\ 0, & \text{διαφορετικά.} \end{cases} \end{aligned}$$

Γενικά είναι

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

οπότε

$$E\left[X|Y=\frac{1}{2}\right] = \int_0^{\frac{1}{2}} 24x^2dx = \left[\frac{24x^3}{3}\right]_{x=0}^{x=\frac{1}{2}} = 1.$$

Θέμα 2ο: (3 βαθμοί) (1) Δίνεται τυχαίο δείγμα $Y_1, Y_2, \ldots, Y_{50}$ από την κατανομή με συνάρτηση πυκνότητας πιθανότητας $f_Y(y) = \frac{1}{3\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{18}}$. Να βρεθεί εκτιμήτρια μέγιστης πιθανοφάνειας $\hat{\mu}$ για το μ.

(2) Έστω $X_1, X_2, \ldots, X_{64}$ ανεξ. ισόνομες συνεχείς τ.μ. με την ομοιόμορφη κατανομή στο $[0,2]$ με μέση τιμή $E[X_i] = 1$ και $E[X_i^2] = \frac{4}{3}$. Να βρεθεί προσεγγιστικά η πιθανότητα το άθροισμα τους να είναι στο διάστημα $[64 - \frac{8}{\sqrt{3}}, 64 + \frac{16}{\sqrt{3}}]$ (Δίνονται $\Phi(1) = 0.8413$, $\Phi(1.5) = 0.9332$, $\Phi(2) = 0.9773$).

Λύση: (1) Η συνάρτηση πιθανοφάνειας είναι

$$\begin{aligned} f(y_1, y_2, \ldots, y_{50}; \mu) &= \prod_{i=1}^{50} f_{Y_i}(y_i) \\ &= \left(\frac{1}{3\sqrt{2\pi}}\right)^{50} e^{-\sum_{i=1}^{50}\frac{(y_i-\mu)^2}{18}}. \end{aligned}$$

Για να τη βελτιστοποιήσουμε θεωρούμε το λογάριθμο της πιθανοφάνειας

$$\log f(y_1, y_2, \ldots, y_{50}; \mu) = -50\log(3\sqrt{2\pi}) - \sum_{i=1}^{50}\frac{(y_i-\mu)^2}{18}.$$

Έχουμε

$$\frac{d}{d\mu}\log f(y_1, y_2, \ldots, y_{50}; \mu) = \frac{1}{9}\sum_{i=1}^{50}(y_i - \mu) = \frac{1}{9}\left(\sum_{i=1}^{50} y_i - 50\mu\right).$$

Λύνοντας την $\frac{d}{d\mu}\log f(y_1, y_2, \ldots, y_{50}; \mu) = 0$ παίρνουμε την υποψήφια εκτίμηση μέγιστης πιθανοφάνειας

$$\hat{\mu} = \frac{\sum_{i=1}^{50} y_i}{50},$$

η οποία αντιστοιχεί πράγματι σε μέγιστο του λογάριθμου πιθανοφάνειας, αφού

$$\frac{d^2}{d\mu^2}\log f(y_1, y_2, \ldots, y_{50}; \mu) = -\frac{50}{9} < 0.$$

Επομένως η εκτιμήτρια μέγιστης πιθανοφάνειας είναι η

$$\hat{\mu} = \frac{\sum_{i=1}^{50} Y_i}{50}.$$

(2) Έχουμε ότι $Var[X_i] = E[X_i^2] - (E[X_i])^2 = \frac{1}{3}$. Έστω $S_{64} = \sum_{i=1}^{64} X_i$. Τότε $E[S_{64}] = 64E[X_i] = 64$ και $Var[S_{64}] = 64Var[X_i] = \frac{64}{3}$. Από το κεντρικό οριακό θεώρημα έχουμε ότι η $\frac{S_{64}-E[S_{64}]}{\sqrt{Var[S_{64}]}}$ ακολουθεί προσεγγιστικά την τυποποιημένη κανονική κατανομή. Επομένως έχουμε:

$$\begin{aligned}
P\left(64 - \frac{8}{\sqrt{3}} \le S_{64} \le 64 + \frac{16}{\sqrt{3}}\right) &= P\left(-1 \le \frac{S_{64} - E[S_{64}]}{\sqrt{Var[S_{64}]}} \le 2\right) \\
&\simeq P(-1 \le Z \le 2),\ Z \sim \mathcal{N}(0,1) \\
&= \Phi(2) - \Phi(-1) \\
&\overset{\Phi(-z)+\Phi(-z)=1}{=} \Phi(1) + \Phi(2) - 1 \\
&= 0.8186.
\end{aligned}$$

Θέμα 3ο: (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 μέση τιμή $E[X]$,
(2) η πιθανότητα $P(X = 100, Y = 0)$,
(3) η πιθανότητα $P(X = 100, Y = 1)$,
(4) η πιθανότητα $P(Y = 0)$,
(5) η δεσμευμένη μέση τιμή $E[Y|X = 100]$,
(6) η δεσμευμένη πιθανότητα $P(X = 100|Y = 0)$.

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

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

(2) Έχουμε, χρησιμοποιώντας τον πολλαπλασιαστικό νόμο

$$\begin{aligned}
P(X = 100, Y = 0) &= P(\text{Επιλέγεται το κατ. Γ, Επιλέγεται άνδρας, Επιλέγεται άνδρας}) \\
&= \frac{1}{3} \cdot \frac{30}{100} \cdot \frac{29}{99} = \frac{29}{990}.
\end{aligned}$$

(3) Έχουμε:

$$\begin{aligned}
P(X = 100, Y = 1) &= P(\text{Επιλέγεται το κατ. Γ, Επιλέγεται άνδρας, Επιλέγεται γυναίκα}) \\
&\quad + P(\text{Επιλέγεται το κατ. Γ, Επιλέγεται γυναίκα, Επιλέγεται άνδρας}) \\
&= \frac{1}{3} \cdot \frac{30}{100} \cdot \frac{70}{99} + \frac{1}{3} \cdot \frac{70}{100} \cdot \frac{30}{99} = \frac{14}{99},
\end{aligned}$$

με χρήση του θεωρήματος ολικής πιθανότητας και του πολλαπλασιαστικού νόμου.

(4) Με βάση το θεώρημα ολικής πιθανότητας έχουμε:

$$\begin{aligned}
P(Y = 0) &= P(\text{Επιλέγεται το κατ. Α})P(\text{Επιλέγ. Ανδ., Ανδ.}|\text{Επιλέχθ. το Α}) \\
&\quad + P(\text{Επιλέγεται το κατ. Β})P(\text{Επιλέγ. Ανδ., Ανδ.}|\text{Επιλέχθ. το Β}) \\
&\quad + P(\text{Επιλέγεται το κατ. Γ})P(\text{Επιλέγ. Ανδ., Ανδ.}|\text{Επιλέχθ. το Γ}) \\
&= \frac{1}{3} \cdot \frac{25}{50} \cdot \frac{24}{49} + \frac{1}{3} \cdot \frac{30}{75} \cdot \frac{29}{74} + \frac{1}{3} \cdot \frac{30}{100} \cdot \frac{29}{99}.
\end{aligned}$$

(5) Έχουμε:

$$\begin{aligned} P(Y=0|X=100) &= P(\text{Επιλέγεται άνδρας, Επιλέγεται άνδρας}|\text{Επιλέχθηκε το κατ. Γ}) \\ &= \frac{30}{100}\cdot\frac{29}{99} = \frac{29}{330}. \end{aligned}$$

Ομοίως

$$\begin{aligned} P(Y=2|X=100) &= P(\text{Επιλέγεται γυναίκα, Επιλέγεται γυναίκα}|\text{Επιλέχθηκε το κατ. Γ}) \\ &= \frac{70}{100}\cdot\frac{69}{99} = \frac{161}{330}. \end{aligned}$$

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

$$P(Y=1|X=100) = 1 - P(Y=0|X=100) - P(Y=2|X=100) = \frac{14}{33}.$$

Άρα

$$E[Y|X=100] = \sum_y yP(Y=y|X=100) = 0\cdot\frac{29}{330} + 1\cdot\frac{14}{33} + 2\cdot\frac{161}{330} = \frac{7}{5}.$$

(6) Με βάση τον κανόνα του Bayes έχουμε:

$$\begin{aligned} P(X=100|Y=0) &= \frac{P(X=100, Y=0)}{P(Y=0)} \\ &= \frac{\frac{29}{990}}{\frac{1}{3}\cdot\frac{25}{50}\cdot\frac{24}{49} + \frac{1}{3}\cdot\frac{30}{75}\cdot\frac{29}{74} + \frac{1}{3}\cdot\frac{30}{100}\cdot\frac{29}{99}}. \end{aligned}$$

Θέμα 4ο: (2 βαθμοί) Μια εταιρεία λανσάρει τα προϊόντα της με ένα κουπόνι το καθένα. Ένας αγοραστής των προϊόντων συλλέγει τα κουπόνια και κάθε φορά που αγοράζει ένα προϊόν είναι εξίσου πιθανό να έχει ένα από τα 10 είδη κουπονιών K1, K2, ..., K10.

(1) Να βρεθεί η πιθανότητα να έχει βρει τουλάχιστον μια φορά το είδος κουπονιού K1, αν έχει αγοράσει 6 προϊόντα.
(2) Να βρεθεί ο αναμενόμενος αριθμός διαφορετικών ειδών κουπονιών που θα έχει μαζέψει, αν έχει αγοράσει 6 προϊόντα.

Λύση: (1) Έχουμε:

$$\begin{aligned} & P(\text{τουλάχιστον ένα K1 μεταξύ των 6 προϊόντων}) \\ =\ & 1 - P(\text{κανένα K1 μεταξύ των 6 προϊόντων}) \\ =\ & 1 - P(\text{το 1ο όχι K1, το 2ο όχι K1, ..., το 6ο όχι K1}) \\ =\ & 1 - \left(\frac{9}{10}\right)^6. \end{aligned}$$

(2) Έστω τυχαίες μεταβλητές X_i, $i = 1, 2, \ldots, 10$ με

$$X_i = \begin{cases} 1, & \text{αν ένα τουλάχιστον Ki μεταξύ των 6 προϊόντων,} \\ 0, & \text{διαφορετικά.} \end{cases}$$

Από το (1) έχουμε ότι $E[X_1] = P(X_1 = 1) = 1 - \left(\frac{9}{10}\right)^6$ και το ίδιο ισχύει για όλα τα $E[X_i]$, $i = 1, 2, \ldots, 10$. Ο αναμενόμενος αριθμός διαφορετικών ειδών κουπονιών που θα έχει μαζέψει αν έχει αγοράσει 6 προϊόντα είναι $E[X_1 + X_2 + \ldots + X_{10}] = 10\left(1 - \left(\frac{9}{10}\right)^6\right)$.