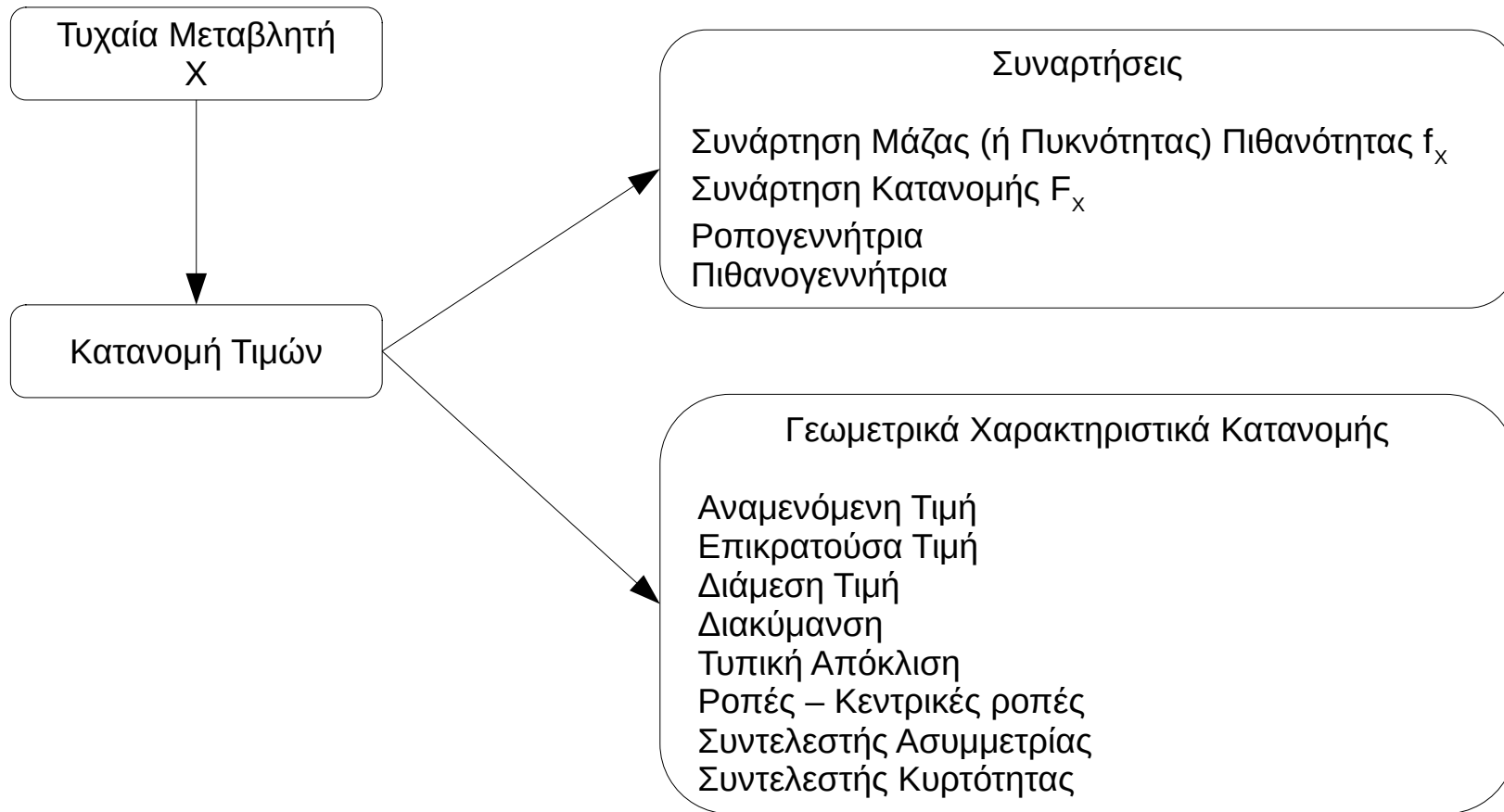


Θεωρία Πιθανοτήτων και Στατιστική

Διδάσκων: Επαμεινώνδας Διαμαντόπουλος
Επικοινωνία: epdiaman@ee.duth.gr

Γραφική Σύνοψη Προηγούμενου Μαθήματος



Μάθημα 4^ο

Περιεχόμενα 4^{ου} μαθήματος

- Ισόνομες τυχαίες μεταβλητές.
- Κοινή συνάρτηση πιθανότητας.
- Ανεξαρτησία τυχαίων μεταβλητών.
- Περιθώρια συνάρτηση πιθανότητας.
- Δεσμευμένη συνάρτηση πιθανότητας.
- Συνδιακύμανση – Συντελεστής συσχέτισης
- Βασικές κατανομές διακριτών τυχαίων μεταβλητών.
 - Κατανομή Bernoulli
 - Διωνυμική κατανομή
 - Υπεργεωμετρική κατανομή
 - Κατανομή Poisson
 - Γεωμετρική κατανομή (2 εκδοχές)
 - Αρνητική διωνυμική.

Γνωστικοί στόχοι 4^{ου} μαθήματος

Στο τέλος αυτού του μαθήματος, ο φοιτητής πρέπει να είναι σε θέση :

- Για δύο τυχαίες μεταβλητές X, Y , να υπολογίζει
 - την από κοινού συνάρτηση πιθανότητας.
 - Την περιθώρια συνάρτηση πιθανότητας.
 - Τη δεσμευμένη συνάρτηση πιθανότητας.
 - Τη συνδιακύμανση και το συντελεστή συσχέτισης.
- Να αναγνωρίζει την κατανομή που ακολουθεί μία διακριτή τυχαία μεταβλητή, αξιολογώντας τη διαδικασία με την οποία αυτή λαμβάνει τιμές.
- Να μπορεί να επιλύει προβλήματα υπολογισμού πιθανοτήτων που αναφέρονται σε μία από τις βασικές διακριτές κατανομές που θα αναφερθούν.

Σχέσεις και συσχετίσεις μεταξύ δύο τυχαίων μεταβλητών

Ισόνομες Τυχαίες Μεταβλητές

Δύο τυχαίες μεταβλητές X , Y , λέγονται **ισόνομες** (identically distributed ή i.d.) όταν

$F_X = F_Y$, δηλαδή

$$F_X(k) = F_Y(k) \text{ ή } P(X \leq k) = P(Y \leq k), k \in \mathbb{R}.$$

Σημείωση

Η αθροιστική συνάρτηση κατανομής είναι απλά μία συνάρτηση από το \mathbb{R} στο $[0, 1]$ και δεν συνδέεται με το πείραμα ή το δειγματοχώρο πάνω στον οποίο ορίζεται μία τυχαία μεταβλητή.

Δύο τυχαίες μεταβλητές μπορεί να είναι ισόνομες ακόμα και αν ορίζονται σε δύο τελείως διαφορετικούς δειγματοχώρους ή στον ίδιο δειγματοχώρο αλλά με διαφορετικό τρόπο.

Για παράδειγμα αν το πείραμα είναι η ρίψη 3 κερμάτων και ορίσουμε:

$X = \{\text{το πλήθος των } K\}$ και $Y = \{\text{το πλήθος των } \Gamma\}$, τότε οι τ. μ. X , Y είναι ισόνομες (γιατί;) ενώ ως τυχαίες μεταβλητές αφορούν διαφορετικά αντικείμενα.

Κοινή, περιθώρια και δεσμευμένη
συνάρτηση πιθανότητας

Κοινή Συνάρτηση Μάζας Πιθανότητας

Ορισμός (κοινή σ.μ.π.)

Έστω δύο διακριτές τυχαίες μεταβλητές X, Y , με συναρτήσεις μάζας πιθανότητας f_X, f_Y . Η **κοινή συνάρτηση (μάζας) πιθανότητας** (joint probability mass function ή joint pmf) των X, Y είναι η συνάρτηση $f_{X,Y}: \mathbb{R} \times \mathbb{R} \rightarrow [0, 1]$ που ορίζεται ως εξής:

$$f_{X,Y}(x, y) = P(X = x, Y = y).$$

Ιδιότητες κοινής συνάρτησης πιθανότητας

- 1) $0 \leq f_{X,Y}(x, y) \leq 1$
- 2) $\sum_x \sum_y f_{X,Y}(x, y) = 1.$

Κοινή Συνάρτηση Πυκνότητας Πιθανότητας

Ορισμός (κοινή σ.π.π.)

Έστω δύο συνεχείς τυχαίες μεταβλητές X, Y . Οι X, Y ονομάζονται αμοιβαία συνεχείς, αν υπάρχει $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, τέτοια ώστε, για κάθε $A \subseteq \mathbb{R}^2$:

$$P((X, Y) \in A) = \iint_A f_{XY}(x, y) dx dy$$

Η συνάρτηση f_{XY} ονομάζεται **κοινή συνάρτηση (πυκνότητας) πιθανότητας** (joint probability density function ή joint pdf) των X, Y .

Ιδιότητες κοινής συνάρτησης πιθανότητας

1) $0 \leq f_{X, Y}(x, y) \leq 1$

2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1$

Κοινή Συνάρτηση Κατανομής

Ορισμός (κοινή σ.κ.)

Έστω δύο τυχαίες μεταβλητές X, Y , με κατανομές F_X, F_Y . Η **κοινή συνάρτηση κατανομής** (joint cumulative distribution function ή joint cdf) των X, Y είναι η συνάρτηση $F_{X,Y}: \mathbb{R} \times \mathbb{R} \rightarrow [0, 1]$ που ορίζεται ως εξής:

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y).$$

Σημείωση: Ο ορισμός αυτός αφορά τόσο τις διακριτές όσο και τις συνεχείς τ.μ.

$$\sum_{k \leq x} \sum_{l \leq y} P(X=k, Y=l)$$
$$\int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(a, b) db da.$$

Περιθώρια Συνάρτηση Πιθανότητας

Από την κοινή συνάρτηση μάζας πιθανότητας είναι δυνατόν να ανακτηθεί η **περιθώρια (marginal) συνάρτηση πιθανότητας** των δύο επιμέρους τυχαίων μεταβλητών.

X, Y διακριτές: $f_X(x) = \sum_y f_{X,Y}(x, y),$ $f_Y(y) = \sum_x f_{X,Y}(x, y)$

X, Y συνεχείς: $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy,$ $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)dx$

$$f_{X,Y} = P(X=x, Y=y)$$

Άσκηση

Δίνονται οι τ.μ. $X = 0, 1, 2$ και $Y = 1, 2$.

(α) Να αποδείξετε ότι η συνάρτηση $f_{X,Y}(x, y) = (x + y) / 15$, αποτελεί αποδεκτή κοινή σ.π. των X, Y .

(β) Να βρείτε την $F_{X,Y}(3/2, 1)$.

(γ) Να βρείτε τις περιθώριες σ.π. f_X, f_Y .

Λύση

(α) Αν $x \in \{0, 1, 2\}, y \in \{1, 2\}$, $f_{X,Y}(x, y) = \frac{x+y}{15} \leq \frac{2+2}{15} < 1$ και $f_{X,Y}(x, y) = \frac{x+y}{15} \geq \frac{0+1}{15} > 0$.

$$\begin{aligned} \sum_x \sum_y f_{X,Y}(x, y) &= f(0, 1) + f(0, 2) + f(1, 1) + f(1, 2) + f(2, 1) + f(2, 2) = \\ &= \frac{1}{15} + \frac{2}{15} + \frac{2}{15} + \frac{3}{15} + \frac{3}{15} + \frac{4}{15} = \frac{15}{15} = 1 \end{aligned}$$

$$\begin{aligned} f_X(2) &= f(2, 1) + f(2, 2) \\ &= \frac{3}{15} + \frac{4}{15} \\ &= \frac{7}{15} \end{aligned}$$

(β) $F_{X,Y}(\frac{3}{2}, 1) = P(X \leq \frac{3}{2}, Y \leq 1) = f(0, 1) + f(1, 1) = \frac{1}{15} + \frac{2}{15} = \frac{3}{15} = \frac{1}{5}$.

(γ) $f_X(x) = \sum_y f_{X,Y}(x, y)$: $f_X(0) = f(0, 1) + f(0, 2) = \frac{1}{15} + \frac{2}{15} = \frac{3}{15}$, $f_X(1) = f(1, 1) + f(1, 2) = \frac{3}{15}$ *

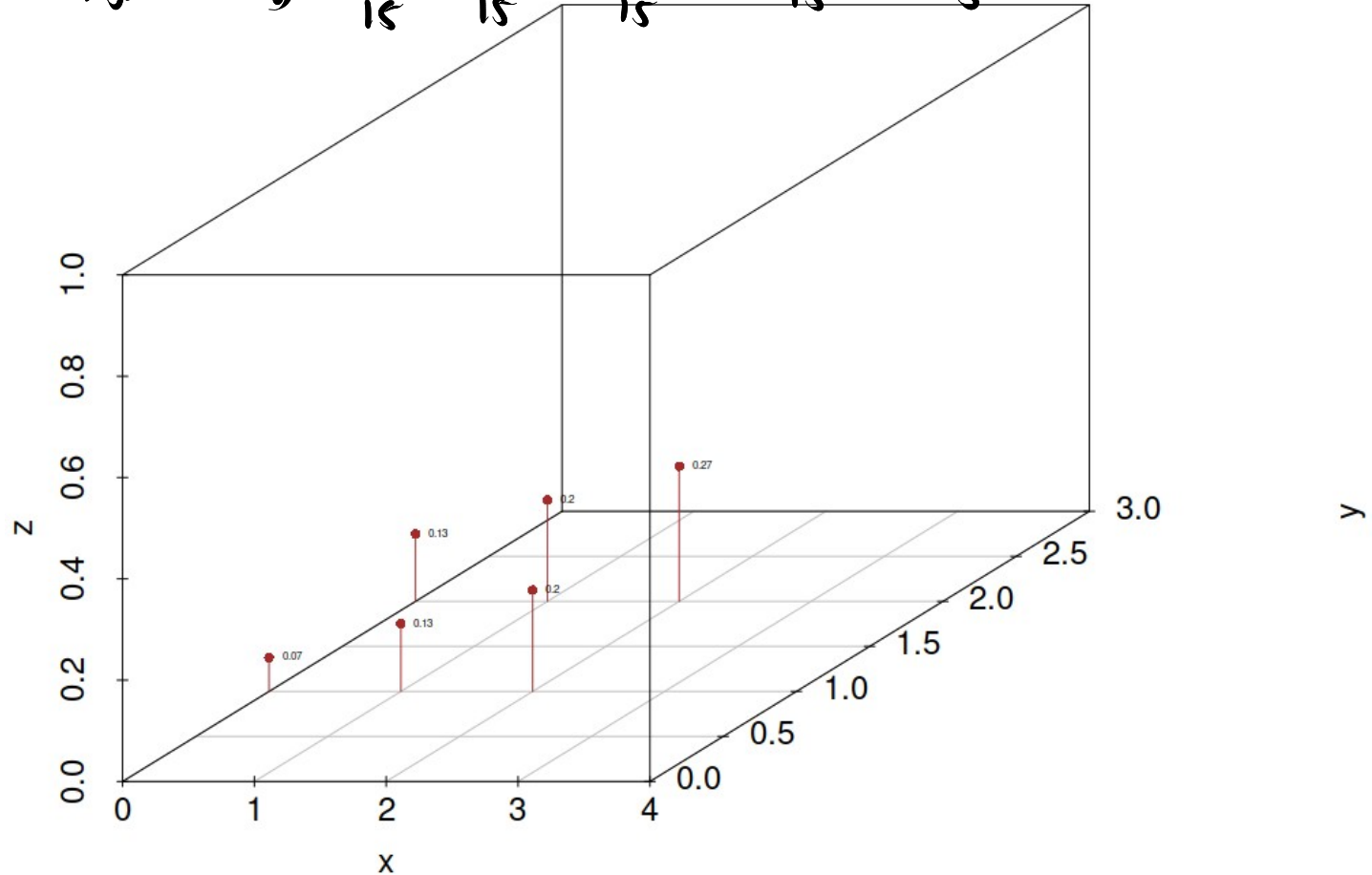
άλυτη

⊛ Evaluar: $f_X(x) = \sum_y f_{X,Y}(x,y) = f(x,1) + f(x,2) = \frac{x+1}{15} + \frac{x+2}{15} = \frac{2x+3}{15}, x=0,1,2$

Άσκηση

$f_Y(y) = \sum_x f_{X,Y}(x,y) = f(0,y) + f(1,y) + f(2,y) = \frac{y}{15} + \frac{1+y}{15} + \frac{2+y}{15} = \frac{3+3y}{15} = \frac{1+y}{5}, y=1,2.$

$f_{X,Y}(x,y) = (x+y)/15,$



$$P(X \leq a, Y \leq b) = \int_{-\infty}^a \int_{-\infty}^b f_{X,Y}(x,y) dy dx$$

Άσκηση

Έστω δύο αμοιβαία συνεχείς τυχαίες μεταβλητές X, Y με κοινή συνάρτηση πυκνότητας πιθανότητας

$f_{X,Y}(x, y) = x + cy^2, 0 \leq x, y \leq 1$ και $f_{X,Y}(x, y) = 0$, αλλού.

(α) Να δείξετε ότι $c = 3/2$.

(β) Να βρείτε την $P(0 \leq X \leq 1/2, 0 \leq Y \leq 1/2)$. $= \int_0^{1/2} \int_0^{1/2} (x + \frac{3}{2}y^2) dx dy = \int_0^{1/2} \left[\frac{x^2}{2} + \frac{3}{2}y^2 x \right]_{x=0}^{x=1/2} dy = \dots$

(γ) Να δείξετε ότι οι περιθώριες σ.π. είναι $f_X(x) = x + 1/2, f_Y(y) = 3/2 y^2 + 1/2$.

Λύση

(α) $\iint_{-\infty}^{\infty} f(x,y) dx dy = 1 \Leftrightarrow \int_0^1 \int_0^1 (x + cy^2) dx dy = 1 \Leftrightarrow \frac{1}{2} + \frac{c}{3} = 1 \Leftrightarrow \boxed{c = \frac{3}{2}}$

$\int_0^1 \int_0^1 (x + cy^2) dx dy = \int_0^1 \left[\frac{x^2}{2} + cy^2 x \right]_{x=0}^{x=1} dy = \int_0^1 \left(\frac{1}{2} + cy^2 \right) dy =$
 $= \left[\frac{1}{2}y + c \frac{y^3}{3} \right]_{y=0}^{y=1} = \frac{1}{2} + \frac{c}{3}$

$$f(x, y) = x + \frac{3}{2}y^2, \quad 0 \leq x, y \leq 1$$

$$(a) f_x(x) = \int_{-\infty}^{+\infty} f_{x,y}(x, y) dy = \int_0^1 (x + \frac{3}{2}y^2) dy = \left[xy + \frac{3}{2} \frac{y^3}{3} \right]_{y=0}^{y=1} = x + \frac{1}{2}.$$

$$f_y(y) = \int_{-\infty}^{+\infty} f_{x,y}(x, y) dx = \int_0^1 (x + \frac{3}{2}y^2) dx = \left[\frac{x^2}{2} + \frac{3}{2}y^2 x \right]_{x=0}^{x=1} = \frac{1}{2} + \frac{3}{2}y^2.$$

Άσκηση

Έστω δύο αμοιβαία συνεχείς τυχαίες μεταβλητές X, Y με κοινή συνάρτηση πυκνότητας πιθανότητας $f_{X,Y}(x, y) = cx^2y, 0 \leq y \leq x \leq 1$ και $f_{X,Y}(x, y) = 0$, αλλού.

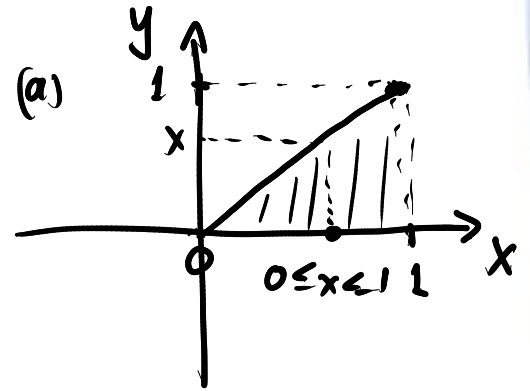
(α) Αναπαραστήστε γραφικά το σύνολο $\{(x, y) \in \mathbb{R}^2, f_{X,Y}(x, y) \neq 0\}$.

(β) Να βρείτε τη σταθερά c .

(γ) Να βρείτε τις περιθώριες συναρτήσεις πυκνότητας πιθανότητας f_X, f_Y .

(δ) Να βρείτε την $P(Y \leq X/2)$.

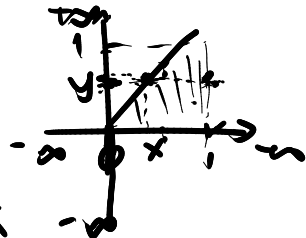
(ε) $P(Y \leq X/4 \mid Y \leq X/2)$.



Λύση

$$\begin{aligned} (β) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy &= \int_0^1 \int_0^x c \cdot x^2 \cdot y dy dx = c \int_0^1 x^2 \left[\frac{y^2}{2} \right]_{y=0}^{y=x} dx = c \int_0^1 \frac{x^4}{2} dx \\ &= \frac{c}{2} \frac{x^5}{5} \Big|_{x=0}^{x=1} = \frac{c}{10}. \text{ Άρα } \frac{c}{10} = 1 \Leftrightarrow \boxed{c=10}. \end{aligned}$$

$$f_{x,y}(x,y) = 10x^2y$$



$$0 \leq y \leq x \leq 1.$$

$$\begin{aligned} \textcircled{a} f_X(x) &= \int_{-\infty}^{\infty} f_{x,y}(x,y) dy = \int_0^x 10x^2y dy = 10x^2 \int_0^x y dy = 10x^2 \left[\frac{y^2}{2} \right]_{y=0}^{y=x} \\ &= 10x^2 \cdot \frac{x^2}{2} = 5x^4. \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_y^1 10x^2y dx = 10y \int_y^1 x^2 dx = 10y \left[\frac{x^3}{3} \right]_{x=y}^{x=1} = 10y \left(\frac{1}{3} - \frac{y^3}{3} \right) \\ &= \frac{10}{3} (y - y^4). \end{aligned}$$

Ανεξαρτησία Τυχαίων Μεταβλητών

Ανεξαρτησία Τυχαίων Μεταβλητών

Ορισμός (ανεξαρτησία τυχαίων μεταβλητών)

Δύο τ. μ. λέγονται **(στοχαστικά) ανεξάρτητες** αν τα ενδεχόμενα $\{X = x\}$ και $\{Y = y\}$ για διακριτές ΤΜ ($x, y \in \mathbb{R}$) ή $\{X \in (\alpha, \beta)\}$ και $\{Y \in (\gamma, \delta)\}$ για συνεχείς ΤΜ ($\alpha, \beta, \gamma, \delta \in \mathbb{R}$) είναι ανεξάρτητα μεταξύ τους.

Αν οι δύο τυχαίες μεταβλητές X, Y είναι στοχαστικά ανεξάρτητες τότε τα $\{X = x\}$ και $\{Y = y\}$ είναι ανεξάρτητα μεταξύ τους και ισχύει

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$$

για κάθε $x, y \in \mathbb{R}$.

Ανεξαρτησία Τυχαίων Μεταβλητών

Δύο τυχαίες μεταβλητές X, Y , είναι **(στοχαστικά) ανεξάρτητες** (independent) αν για κάθε $x, y \in \mathbb{R}$, ισχύει

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y),$$

ή ισοδύναμα, όταν

$$P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y).$$

Σημείωση

Η συνθήκη $P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y)$ είναι ισοδύναμη με την $P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$, δηλαδή με την

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y), \quad x, y \in \mathbb{R},$$

Η παραπάνω ισότητα ισοδυναμεί με την στοχαστική ανεξαρτησία και στην περίπτωση των συνεχών τ.μ. χωρίς ωστόσο να έχει το νόημα σημειακών πιθανοτήτων.

Ισόνομες τ.μ. \nRightarrow Ανεξάρτητες τ.μ.

Παρατήρηση

Δύο τυχαίες μεταβλητές X, Y , μπορεί να είναι *ισόνομες* αλλά όχι *ανεξάρτητες*.

Παράδειγμα

Σε ένα σάκο υπάρχουν 2 μπάλες, μία με τον αριθμό 0 και μία με τον αριθμό 1. Εκτελούμε πείραμα με δύο επιλογές μπάλας και ορίζουμε να είναι

X : Ο αριθμός της μπάλας στην 1^η επιλογή και Y : Ο αριθμός της μπάλας στη 2^η επιλογή.

Αν η δειγματοληψία γίνει **με επανάθεση** τότε $\Omega = \{00, 01, 10, 11\}$ και υπολογίζουμε

$$P(X = 0) = P(X = 1) = P(Y = 0) = P(Y = 1) = 2/4 = 0,5.$$

Στην περίπτωση αυτή οι X, Y είναι **ισόνομες και ανεξάρτητες**.

Αν ωστόσο η δειγματοληψία γίνει **χωρίς επανάθεση**, τότε $\Omega = \{01, 10\}$ και υπολογίζουμε

$$P(X = 0) = P(X = 1) = P(Y = 0) = P(Y = 1) = 1/2 = 0,5.$$

Στην περίπτωση αυτή οι X, Y είναι **ισόνομες και εξαρτημένες**.

Ανεξαρτησία Τυχαίων Μεταβλητών

Παράδειγμα 1

Έστω ότι ρίχνουμε ένα ζάρι και ορίζουμε $X = \{1 \text{ αν η ρίψη είναι άρτιος και } 0 \text{ αν είναι περιττός}\}$ και $Y = \{1 \text{ αν η ρίψη είναι πρώτος και } 0 \text{ αν είναι σύνθετος}\}$.

(α) Να βρεθεί η κοινή συνάρτηση μάζας πιθανότητας.

(β) Είναι οι X, Y ανεξάρτητες τυχαίες μεταβλητές;

Ανεξαρτησία Τυχαίων Μεταβλητών

Λύση

Καθώς, τόσο η X όσο και η Y είναι διακριτές τ.μ. μπορούμε να καταγράψουμε το σύνολο των συνδυασμών των τιμών τους

(α) Για την κοινή συνάρτηση μάζας πιθανότητας υπολογίζουμε

$$P(X = 0, Y = 0) = 1/6, P(X = 1, Y = 0) = 2/6,$$

$$P(X = 0, Y = 1) = 2/6, P(X = 1, Y = 1) = 1/6$$

Ρίψη	1	2	3	4	5	6
X	0	1	0	1	0	1
Y	0	1	1	0	1	0

(β) $P(X = 1, Y = 0) = 2/6 = 1/3$

$$P(X = 1) \cdot P(Y = 0) = 3/6 \cdot 3/6 = 9/36 = 1/4.$$

Οι τυχαίες μεταβλητές είναι (στοχαστικά) εξαρτημένες.

Ανεξαρτησία Τυχαίων Μεταβλητών

Παράδειγμα 2

Έστω ένα κέρμα 2 φορές και ορίζουμε $X = \{\text{πλήθος } K \text{ στις δύο ρίψεις}\}$. Ρίχνουμε το κέρμα άλλες 2 φορές και ορίζουμε $Y = \{\text{πλήθος } \Gamma \text{ στις δύο ρίψεις}\}$ (α) Είναι οι X, Y ανεξάρτητες τυχαίες μεταβλητές; (β) Να βρεθεί η πιθανότητα $P(X < 2, Y > 1)$

Λύση

Ανεξαρτησία Τυχαίων Μεταβλητών

Παράδειγμα 2

Έστω ένα κέρμα 2 φορές και ορίζουμε $X = \{\text{πλήθος } K \text{ στις δύο ρίψεις}\}$. Ρίχνουμε το κέρμα άλλες 2 φορές και ορίζουμε $Y = \{\text{πλήθος } \Gamma \text{ στις δύο ρίψεις}\}$ (α) Είναι οι X, Y ανεξάρτητες τυχαίες μεταβλητές; (β) Να βρεθεί η πιθανότητα $P(X < 2, Y > 1)$

Λύση

Είναι $X = 0, 1, 2$ και $Y = 0, 1, 2$ και

$$P(X = 0) = P(Y = 0) = \frac{1}{4}, P(X = 1) = P(Y = 1) = \frac{1}{2}, P(X = 2) = P(Y = 2) = \frac{1}{4}.$$

(α) Οι X, Y λαμβάνουν τιμές από διαφορετικές ρίψεις κερμάτων άρα είναι ανεξάρτητες τυχαίες μεταβλητές.

$$(β) P(X < 2, Y > 1) = P(X < 2) \cdot P(Y > 1) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}.$$

Άσκηση

Δίνονται οι τ.μ. $X = 0, 1, 2$ και $Y = 1, 2$, με κοινή συνάρτηση πιθανότητας $f_{X,Y}(x, y) = (x + y) / 15$. Να βρείτε τις περιθώριες σ.π. f_X, f_Y . Είναι οι X, Y ανεξάρτητες;

Λύση

$$f_X(x) = \frac{2x+3}{15}$$

$$f_Y(y) = \frac{1+y}{5}$$

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) : \text{Επιθυμητή ανεξαρτησία}$$

$$f_X(0) = \frac{3}{15}, f_Y(1) = \frac{2}{5}, f_{X,Y}(0,1) = \frac{1}{15}$$

$$f_{X,Y}(0,1) = \frac{1}{15} \neq \frac{6}{75} = f_X(0) \cdot f_Y(1) \Rightarrow X, Y : \text{εξαρτημένοι.}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Δεσμευμένη συνάρτηση κατανομής

Έστω X, Y τυχαίες μεταβλητές με κοινή συνάρτηση μάζας πιθανότητας $f_{X,Y}$. Από την κοινή συνάρτηση μάζας πιθανότητας, είναι δυνατόν να οριστεί και η **δεσμευμένη συνάρτηση μάζας πιθανότητας** της X δοθέντος του $Y = y$, ως εξής:

$$f_{X|Y}(x|y) = f_{X,Y}(x,y) / f_Y(y) = \frac{P(X=x, Y=y)}{P(Y=y)}$$

Επιπλέον, η **δεσμευμένη αθροιστική συνάρτηση κατανομής** της X δεδομένου ότι $Y = y$, ορίζεται να είναι

$$F_{X|Y}(x|y) = P(X \leq x | Y = y) = \sum_{k \leq x} P(X = k | Y = y)$$

Ανεξαρτησία Τυχαίων Μεταβλητών

Δύο τ. μ. είναι (στοχαστικά) ανεξάρτητες αν για κάθε $y \in \mathbb{R}$,

$$f_{x|y}(x | y) = f_x(x)$$

Πράγματι: $f_{x|y}(x | y) = f_{x,y}(x, y) / f_y(y) = f_x(x) \cdot f_y(y) / f_y(y) = f_x(x)$

Ασκήσεις

Άσκηση 1

Έστω ότι η κοινή συνάρτηση μάζας πιθανότητας των X, Y δίνεται από τις σχέσεις:

$$f_{X,Y}(0, 0) = 0,4, \quad f_{X,Y}(0, 1) = 0,2, \quad f_{X,Y}(1, 0) = 0,1, \quad f_{X,Y}(1, 1) = 0,3.$$

(α) Να βρεθούν οι περιθώριες συναρτήσεις πιθανότητας των X, Y .

(β) Να βρεθεί η $f_{X|Y}(x | y = 1)$

Λύση

$$X=0,1, Y=0,1 \quad (α) f_X(x) = \sum_y f_{X,Y}(x,y) = f(x,0) + f(x,1)$$

$$f_X(0) = f(0,0) + f(0,1) = 0,4 + 0,2 = 0,6$$

$$f_X(1) = f(1,0) + f(1,1) = 0,1 + 0,3 = 0,4$$

		Y		
		0	1	
X	0	0,4	0,2	0,6
	1	0,1	0,3	0,4
		0,5	0,5	

$\left. \begin{array}{l} \text{Row 1} \\ \text{Row 2} \end{array} \right\} f_X(x)$
 $\left. \begin{array}{l} \text{Col 2} \\ \text{Col 3} \end{array} \right\} f_Y(y)$

$$(β) f_{X|Y}(x | y=1) = \frac{f_{X,Y}(x, y=1)}{f_Y(y=1)}$$

$$f_{X|Y}(0 | y=1) = \frac{0,2}{0,5} = \frac{2}{5}, \quad f_{X|Y}(1 | y=1) = \frac{0,3}{0,5} = \frac{3}{5} \text{ Άλυτη}$$

Ασκήσεις

Άσκηση 2

Ο πίνακας δίνει την κοινή συνάρτηση πιθανότητας για το ζεύγος των τ.μ. (X, Y) .

(α) Να βρεθούν οι περιθώριες και οι δεσμευμένες συναρτήσεις πιθανότητας.

(β) Να βρεθούν οι $P(X = Y)$, $P(X > 2Y)$, $P(X > Y)$, $P(X \geq Y \mid Y \geq 2)$.

Λύση

$f_{X,Y}$	$Y = 1$	$Y = 2$	$Y = 3$
$X = 1$	1/12	1/6	0
$X = 2$	0	1/9	1/5
$X = 3$	1/18	1/4	2/15

Δεσμευμένη Μέση Τιμή

Δεσμευμένη Μέση Τιμή

Ορισμός

Η δεσμευμένη σ.π. της τ.μ. X ως προς την τιμή y_0 της τ.μ. Y ορίζεται ως

$$f_x(x_i | y_0) = P(X = x_i | Y = y_0) = \frac{P(X = x_i, Y = y_0)}{P(Y = y_0)} = \frac{f_{X,Y}(x_i, y_0)}{f_Y(y_0)}$$

όπου $f_{X,Y}(x, y) = P(X = x, Y = y)$ η κοινή συνάρτηση μάζας πιθανότητας των τ.μ. X, Y .

Ορισμός

Η δεσμευμένη μέση τιμή (conditional expected value) μιας διακριτής τυχαίας μεταβλητής X , δοθέντος της $Y = y_0$, ορίζεται ως:

$$E(X | Y = y_0) = \sum_{x_i} x_i \cdot f_x(x_i | y_0)$$

$$P(X=0 | Y=0) = \frac{P(X=0, Y=0)}{P(Y=0)} = \frac{1/6}{3/6} = \frac{1}{3}$$

Δεσμευμένη Μέση Τιμή

Άσκηση 1

Έστω ότι ρίχνουμε ένα ζάρι και ορίζουμε $X = \{1 \text{ αν η ρίψη είναι άρτιος και } 0 \text{ αν είναι περιττός}\}$ και $Y = \{1 \text{ αν η ρίψη είναι πρώτος και } 0 \text{ αν είναι σύνθετος}\}$.

(α) Να βρεθούν τα $E(X)$, $E(Y)$.

(β) Να βρεθούν τα $E(X | Y = 0)$, $E(X | Y = 1)$, $E(Y | X = 0)$, $E(Y | X = 1)$.

Λύση

Πείραμα	1	2	3	4	5	6
X	0	1	0	1	0	1
Y	0	1	1	0	1	0

Handwritten annotations: A red bracket above the table spans from the first column (X=0) to the fifth column (X=1), labeled 'Y=0'. Another red bracket above the table spans from the second column (X=1) to the sixth column (X=0), labeled 'X=1'.

$$\begin{aligned} \text{(α)} \quad E(X) &= 0 \cdot P(X=0) + 1 \cdot P(X=1) = \\ &= 0 \cdot 0,5 + 1 \cdot 0,5 = 0,5 \end{aligned}$$

$$\begin{aligned} E(Y) &= 0 \cdot P(Y=0) + 1 \cdot P(Y=1) = \\ &= 0 \cdot 0,5 + 1 \cdot 0,5 = 0,5. \end{aligned}$$

$$\text{(β)} \quad E(X | Y=0) = 0 \cdot P(X=0 | Y=0) + 1 \cdot P(X=1 | Y=0) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$

$$E(Y | X=1) = 0 \cdot P(Y=0 | X=1) + 1 \cdot P(Y=1 | X=1) = 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{1}{3}$$

Δεσμευμένη Μέση Τιμή

Άσκηση 2

Η κοινή συνάρτηση πιθανότητας των X, Y δίνεται από τον τύπο

$$f_{X,Y}(x, y) = (x + y) / 15, \quad x = 0, 1, 2 \text{ και } y = 1, 2.$$

Να υπολογιστούν οι δεσμευμένες μέσες τιμές $E(X | Y = 1)$ και $E(Y | X = 1)$.

Λύση

Ανεξαρτησία και αναμενόμενη τιμή 1/2

Θεώρημα 1

Αν X, Y ανεξάρτητες τ.μ. τότε $E(X | Y = y) = E(X)$.

Απόδειξη (για διακριτές τ.μ.)

Πράγματι, $E(X | Y = y) = \sum_{x_i} x_i \cdot P(X = x_i | Y = y)$

(ορισμός $E(X)$)

$$= \sum_{x_i} x_i \cdot P(X = x_i) \cdot P(Y = y) / P(Y = y)$$

(X, Y ανεξάρτητες)

$$= \sum_{x_i} x_i \cdot P(X = x_i)$$

$$= E(X).$$

$$E(X+Y) = E(X) + E(Y).$$

Ανεξαρτησία και αναμενόμενη τιμή 2/2

Θεώρημα 2

Αν X, Y ανεξάρτητες τ.μ. τότε $E(X \cdot Y) = E(X) \cdot E(Y)$.

Απόδειξη

Αν X, Y ανεξάρτητες, τότε $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ και

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y)xy dx dy \\ &= \int_{-\infty}^{\infty} f_X(x)f_Y(y)xy dx dy \\ &= \int_{-\infty}^{\infty} f_X(x)xdx \int_{-\infty}^{\infty} f_Y(y)ydy \\ &= E(X)E(Y) \end{aligned}$$

$$\begin{aligned} E[XY] &= \sum_{x,y} xy \Pr(X=x, Y=y) \\ &= \sum_x \sum_y xy \Pr(X=x, Y=y) \\ &= \sum_x \sum_y xy \Pr(X=x) \Pr(Y=y) \\ &= \sum_x x \Pr(X=x) \sum_y y \Pr(Y=y) \\ &= \sum_x x \Pr(X=x) E[Y] \\ &= E[Y] \sum_x x \Pr(X=x) \\ &= E[Y] E[X] \end{aligned}$$

Σημείωση: Όπως έχουμε δει, η αντίστοιχη ιδιότητα της αναμενόμενης τιμής για το άθροισμα ισχύει χωρίς κάποια προϋπόθεση για τις τ.μ. X, Y . Δηλαδή για κάθε τ.μ. X, Y ισχύει $E(X + Y) = E(X) + E(Y)$.

Ασκήσεις στις ανεξάρτητες τ.μ.

Άσκηση 1

Έστω ένα κέρμα 2 φορές και ορίζουμε $X = \{\text{πλήθος } K \text{ στις δύο ρίψεις}\}$. Ρίχνουμε το κέρμα άλλες 2 φορές και ορίζουμε $Y = \{\text{πλήθος } \Gamma \text{ στις δύο ρίψεις}\}$ (α) Είναι οι X, Y ανεξάρτητες τυχαίες μεταβλητές; (β) Να βρεθεί η πιθανότητα $P(X < 2, Y > 1)$

Ασκήσεις στις ανεξάρτητες τ.μ.

Άσκηση 2

Ο πίνακας δίνει την κοινή συνάρτηση πιθανότητας για το ζεύγος των τ.μ. (X, Y) .

(α) Να βρεθούν οι περιθώριες συναρτήσεις πιθανότητας.

(β) Είναι οι τ.μ. X, Y ανεξάρτητες;

(γ) Να βρεθεί η δεσμευμένη κατανομή του X , δοθέντος $Y = 2$ και η δεσμευμένη κατανομή του Y , δοθέντος $X = 3$.

(a)

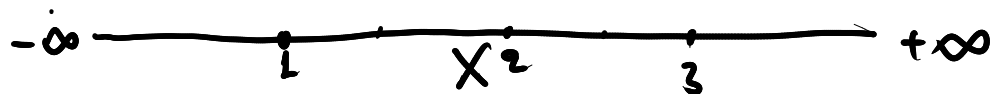
$f_{X,Y}(x,y)$		$\downarrow Y$				f_X
		1	2	3	4	
X	1	0,06	0,12	0,24	0,18	0,60
	2	0,02	0	0,08	0,06	0,16
	3	0	0,08	0,16	0	0,24
		0,08	0,20	0,48	0,24	

$$(β) f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

$$f_{X,Y}(3,1) = 0$$

$$f_X(3) \cdot f_Y(1) = 0,24 \cdot 0,06 \neq 0 = f_{X,Y}(3,1) \Rightarrow X, Y \text{ ελκρυφίνω.}$$

$$(δ) F_{X|Y=2}(x) = P(X \leq x | Y=2)$$



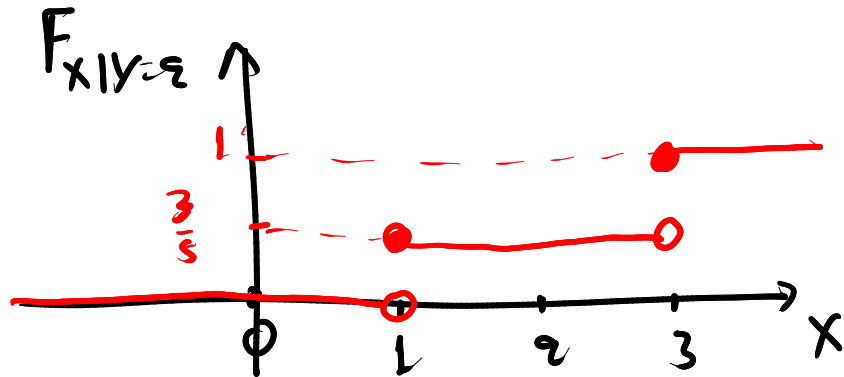
$$P(X \leq x | Y=2)$$

$$A_v \quad x < 1 \Rightarrow F_{X|Y=2}(x) = 0$$

$$A_v \quad 1 \leq x < 2 \Rightarrow F_{X|Y=2}(x) = \frac{P(X=1|Y=2)}{P(Y=2)} = \frac{0,12}{0,20} = \frac{3}{5}$$

$$A_v \quad 2 \leq x < 3 \Rightarrow F_{X|Y=2}(x) = \frac{3}{5}$$

$$A_v \quad x \geq 3 \Rightarrow F_{X|Y=2}(x) = \frac{P(X=1|Y=2) + P(X=2|Y=2) + P(X=3|Y=2)}{P(Y=2)} = \frac{5}{5} = 1$$



- Συνδιακύμανση
- Συντελεστής συσχέτισης

Συνδιακύμανση Τυχαίων Μεταβλητών

Η συνδιακύμανση (covariance) είναι μία στατιστική ποσότητα με την οποία ποσοτικοποιείται το είδος της συμμεταβολής των δύο μεταβλητών, δηλαδή το είδος της μεταβολής στις τιμές μίας συνεχούς τυχαίας μεταβλητής καθώς μία άλλη μεταβάλλεται.

Ο υπολογισμός της έχει νόημα για ζεύγη μεταβλητών που ορίζονται στο ίδιο πείραμα, άρα οι τιμές τους προσδιορίζονται μαζί με την εξέλιξη του πειράματος.

Ορισμός

Αν X και Y είναι δύο τυχαίες μεταβλητές τότε η **συνδιακύμανση** $\text{Cov}(X, Y)$ ορίζεται ως

$$\sigma^2 \quad \text{Cov}(X, Y) = \mathbf{E}[(X - \bar{\mu}_X)(Y - \bar{\mu}_Y)], \quad \mu_X = \mathbf{E}(X), \quad \mu_Y = \mathbf{E}(Y).$$

$$\text{Var}(X) = E(X - \mu)^2 = E(X^2) - (EX)^2$$

Συνδιακύμανση Τυχαίων Μεταβλητών

$$\text{Cov}(X, Y) = E(X - \mu_x)(Y - \mu_y)$$

$$E(X \cdot Y) = \sum_x \sum_y x \cdot y \cdot P(X=x, Y=y)$$

Ιδιότητες

1. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.

2. $\text{Cov}(X, Y) = E(XY) - \mu_x \cdot \mu_y$ ←

Απόδειξη: $\text{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = E(XY - \mu_x \cdot Y - \mu_y \cdot X + \mu_x \cdot \mu_y)$
 $= E(XY) - \mu_x \cdot E(Y) - \mu_y \cdot E(X) + \mu_x \mu_y = E(XY) - \mu_x \cdot \mu_y$

3. Αν X, Y ανεξάρτητες τυχαίες μεταβλητές τότε $\text{Cov}(X, Y) = 0$.

Απόδειξη: $\text{Cov}(X, Y) = E(XY) - \mu_x \cdot \mu_y = E(X)E(Y) - \mu_x \cdot \mu_y = \mu_x \cdot \mu_y - \mu_x \cdot \mu_y = 0$.

4. $\text{Cov}(X, X) = \text{Var}(X)$

Απόδειξη: $\text{Cov}(X, X) = E(X^2) - \mu_x^2 = \text{Var}(X)$.

5. $\text{Cov}(\alpha X + \beta Y, Z) = \alpha \cdot \text{Cov}(X, Z) + \beta \cdot \text{Cov}(Y, Z)$

6. $\text{Var}(\alpha X + \beta Y) = \alpha^2 \text{Var}(X) + \beta^2 \text{Var}(Y) + 2\alpha\beta \text{Cov}(X, Y)$.

7. $\text{Var}(\alpha X - \beta Y) = \alpha^2 \text{Var}(X) + \beta^2 \text{Var}(Y) - 2\alpha\beta \text{Cov}(X, Y)$.

8. Αν X, Y ανεξάρτητες τ.μ. τότε $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

Συνδιακύμανση Τυχαίων Μεταβλητών

Παράδειγμα

Δύο τυχαίες μεταβλητές, X , Y παίρνουν τιμές $X = 1, 2, 3$ και $Y = 1, 2$. Η κοινή συνάρτηση μάζας πιθανότητας παρουσιάζεται στον πίνακα. Να βρεθεί η συνδιακύμανση των X , Y .

Λύση

$$E(X) = 1 \cdot 0,3 + 2 \cdot 0,4 + 3 \cdot 0,3 = 2$$

$$E(Y) = 1 \cdot 0,4 + 2 \cdot 0,6 = 1,6$$

$f_{X,Y}(x, y)$		$f_X(x)$			Σύνολο
		$X = 1$	$X = 2$	$X = 3$	
$f_Y(y)$	$Y = 1$	0,2	0,2	0	0,4
	$Y = 2$	0,1	0,2	0,3	0,6
Σύνολο		0,3	0,4	0,3	1

$$E(X \cdot Y) = 1 \cdot 1 \cdot 0,2 + 1 \cdot 2 \cdot 0,1 + 2 \cdot 1 \cdot 0,2 + 2 \cdot 2 \cdot 0,2 + 3 \cdot 1 \cdot 0 + 3 \cdot 2 \cdot 0,3 = 3,4$$

$$Cov(X, Y) = E(XY) - E(X) \cdot E(Y) = 3,4 - 2 \cdot 1,6 = 3,4 - 3,2 = 0,2$$

$$\text{ή } Cov(X, Y) = E(X - \mu_X)(Y - \mu_Y) = (1 - 2) \cdot (1 - 1,6) \cdot 0,2 + (1 - 2) \cdot (2 - 1,6) \cdot 0,1 + \dots + (3 - 2) \cdot (2 - 1,6) \cdot 0,3 = \dots$$

Συνδιακύμανση Τυχαίων Μεταβλητών

Παράδειγμα

Δύο τυχαίες μεταβλητές, X , Y παίρνουν τιμές $X = 1, 2, 3$ και $Y = 1, 2$. Η κοινή συνάρτηση μάζας πιθανότητας παρουσιάζεται στον πίνακα. Να βρεθεί η συνδιακύμανση των X , Y .

Λύση

Υπολογίζουμε,

$$E(X) = \sum_x x \cdot f_x(x) = 1 \cdot 0,3 + 2 \cdot 0,4 + 3 \cdot 0,3 = 2,$$

$$E(Y) = \sum_y y \cdot f_y(y) = 1 \cdot 0,4 + 2 \cdot 0,6 = 1,6,$$

$f_{X,Y}(x, y)$	$X = 1$	$X = 2$	$X = 3$	Σύνολο	
$Y = 1$	0,25	0,2	0	0,4	$f_Y(y)$
$Y = 2$	0,05	0,2	0,3	0,6	
Σύνολο	0,3	0,4	0,3	1	
	$f_X(x)$				

$$\begin{aligned} E(XY) &= \sum_x \sum_y x \cdot y \cdot f_{X,Y}(x, y) = 1 \cdot 1 \cdot 0,25 + 1 \cdot 2 \cdot 0,05 + 2 \cdot 1 \cdot 0,2 + 2 \cdot 2 \cdot 0,2 + 3 \cdot 1 \cdot 0 + 3 \cdot 2 \cdot 0,3 = \\ &= 0,25 + 0,1 + 0,4 + 0,8 + 0 + 1,8 = 3,35 \end{aligned}$$

$$\text{Υπολογίζουμε } \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 3,35 - 3,2 = 0,15.$$

Συντελεστής Συσχέτισης Pearson

Ο συντελεστής συσχέτισης του Pearson είναι το κατάλληλο στατιστικό για την ανίχνευση της γραμμικής σχέσης δύο ποσοτικών μεταβλητών, για συνεχείς ή αριθμητικές διακριτές μεταβλητές

$$\rho_{X,Y} = \text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mathbf{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

Ισοδύναμα:

$$\rho_{X,Y} = \frac{\mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)}{\sqrt{\mathbf{E}(X^2) - \mathbf{E}(X)^2} \cdot \sqrt{\mathbf{E}(Y^2) - \mathbf{E}(Y)^2}}$$

Συντελεστής Συσχέτισης Pearson

Αν δύο μεταβλητές είναι ανεξάρτητες, τότε $\text{cov}(X, Y) = 0$ και $\rho_{X, Y} = 0$ επίσης. Το αντίθετο δεν ισχύει, δηλαδή μπορεί δύο μεταβλητές να είναι γραμμικά ασυσχέτιστες ($\rho_{X, Y} = 0$) και να είναι εξαρτημένες.

Άσκηση

Έστω, $X = -1, 0, 1$ με $f_X(-1) = f_X(0) = f_X(1) = 1/3$ και $Y = 1$, αν $X = 0$ και $Y = 0$, αν $X = -1, 1$. Προφανώς οι X, Y είναι εξαρτημένες. Να δείξετε ότι $\rho_{X, Y} = 0$.

Υπόδειξη

Μελετήστε το γινόμενο XY .

$$X \cdot Y = 0 \text{ ή } -1 \text{ ή } 1$$

$$\text{Για } X = -1, Y = 0 \text{ και } X \cdot Y = 0$$

$$\text{Για } X = 0, Y = 1 \text{ και } X \cdot Y = 0$$

$$\text{Για } X = 1, Y = 0 \text{ και } X \cdot Y = 0$$

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = 0.$$

$$E(X) = -1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0$$

$$E(Y) = \text{αδιόριστο} \dots$$

Συντελεστής Συσχέτισης Pearson

Άσκηση

Έστω ότι οι τ.μ. X_1, X_2 είναι ανεξάρτητες με αναμενόμενη τιμή 0 και διακύμανση σ^2 . Να βρεθεί ο συντελεστής συσχέτισης Pearson των τ.μ. $X_1 + 2X_2$ and $4X_1 - 3X_2$.

Λύση

$$(δ) G_Y(t) = E(z^t) = z^0 \cdot 0,6 + z^1 \cdot 0,4 = 0,4z + 0,6.$$

Άσκηση εμπέδωσης / επανάληψης

Στον πίνακα παρουσιάζεται η κοινή συνάρτηση μάζας πιθανότητας $f_{X,Y}$.

(α) Να αποδείξετε ότι $f_X(1) = 0,3$, $f_X(2) = 0,5$, $f_X(3) = 0,2$.

(β) Να σχεδιάσετε το διάγραμμα της αθροιστικής συνάρτησης κατανομής F_X .

(γ) Να αποδείξετε ότι $f_Y(0) = 0,6$, $f_Y(1) = 0,4$.

(δ) Δείξτε ότι $M_Y(t) = 0,4e^t + 0,6$, $G_Y(t) = 0,4z + 0,6$.

(ε) Δείξτε ότι $E(X) = 1,9$, $E(X^2) = 4,1$, $\text{Var } X = 0,49$, $\sigma_X = 0,7$.

(ζ) Δείξτε ότι $E(Y) = 0,4$, $E(Y^2) = 0,4$, $\text{Var } Y = 0,24$, $\sigma_Y = 0,49$.

(η) Δείξτε ότι $E(XY) = 0,8$ και $\text{Cov}(X, Y) = 0,04$.

(θ) Δείξτε ότι $\text{Corr}(X, Y) = 0,117$.

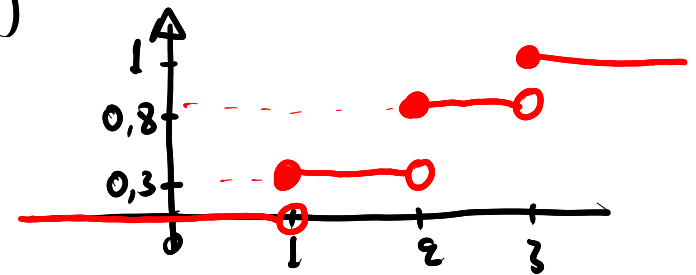
(ι) Δείξτε ότι οι τ.μ. X, Y είναι εξαρτημένες.

$$(δ) M_Y(t) = E(e^{tY}) = e^{t \cdot 0} \cdot 0,6 + e^{t \cdot 1} \cdot 0,4 = 0,4e^t + 0,6.$$

(α)

$f_{X,Y}$	X=1	X=2	X=3	
Y=0	0,2	0,3	0,1	0,6
Y=1	0,1	0,2	0,1	0,4
	0,3	0,5	0,2	

(β)



$$(ε) E(X) = 1 \cdot 0,3 + 2 \cdot 0,5 + 3 \cdot 0,2 = 1,9$$



$$E(X^2) = 1^2 \cdot 0,3 + 2^2 \cdot 0,5 + 3^2 \cdot 0,2 = 4,1$$

$$\text{Var}(X) = E(X-\mu)^2 = E(X^2) - [E(X)]^2 = 4,1 - 1,9^2 = 0,49$$

$$\sigma_x = \sqrt{\text{Var}(X)} = \sqrt{0,49} = 0,7$$

$$(μ) E(XY) = 1 \cdot 0 \cdot 0,2 + 1 \cdot 1 \cdot 0,1 + 2 \cdot 0 \cdot 0,3 + 2 \cdot 1 \cdot 0,2 + 3 \cdot 0 \cdot 0,1 + 3 \cdot 1 \cdot 0,1 = 0,8$$

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0,8 - 1,9 \cdot 0,4 = 0,04$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{0,04}{0,7 \cdot 0,49} = 0,117.$$

$$(θ) f_{X,Y}(1,0) = 0,2 \text{ και } f_X(1) \cdot f_Y(0) = 0,3 \cdot 0,6 = 0,18 \neq 0,2 \Rightarrow X, Y: \text{εξαρτημένοι}$$

Άσκηση εμπέδωσης / επανάληψης

(α) – (γ): Από τον πίνακα.

(δ) $M_Y(t) = E(e^{Yt}) = 0,4e^t + 0,6$, $G_Y(t) = E(z^Y) = 0,4z + 0,6$.

(ε)

$$E(X) = 1 \cdot f_X(1) + 2 \cdot f_X(2) + 3 \cdot f_X(3) = 1 \cdot 0,3 + 2 \cdot 0,5 + 3 \cdot 0,2 = 1,9$$

$$E(X^2) = 1^2 \cdot f_X(1) + 2^2 \cdot f_X(2) + 3^2 \cdot f_X(3) = 1 \cdot 0,3 + 4 \cdot 0,5 + 9 \cdot 0,2 = 4,1,$$

$$\text{Var } X = E(X^2) - E(X)^2 = 4,1 - 1,9^2 = 0,49, \sigma_X = 0,49^{1/2} = 0,7.$$

(ζ) ανάλογα με το (ε) $E(Y) = 0,4$, $E(Y^2) = 0,4$, $\text{Var } Y = 0,24$, $\sigma_Y = 0,49$.

$$\begin{aligned} \text{(η)} \quad E(XY) &= 1 \cdot 0 \cdot f_{XY}(1, 0) + 2 \cdot 0 \cdot f_{XY}(2, 0) + 3 \cdot 0 \cdot f_{XY}(3, 0) + 1 \cdot 1 \cdot f_{XY}(1, 1) + 2 \cdot 1 \cdot f_{XY}(2, 1) + 3 \cdot 1 \cdot f_{XY}(3, 1) = \\ &= 1 \cdot 0 \cdot 0,2 + 2 \cdot 0 \cdot 0,3 + 3 \cdot 0 \cdot 0,1 + 1 \cdot 1 \cdot 0,1 + 2 \cdot 1 \cdot 0,2 + 3 \cdot 1 \cdot 0,1 = 0,8. \end{aligned}$$

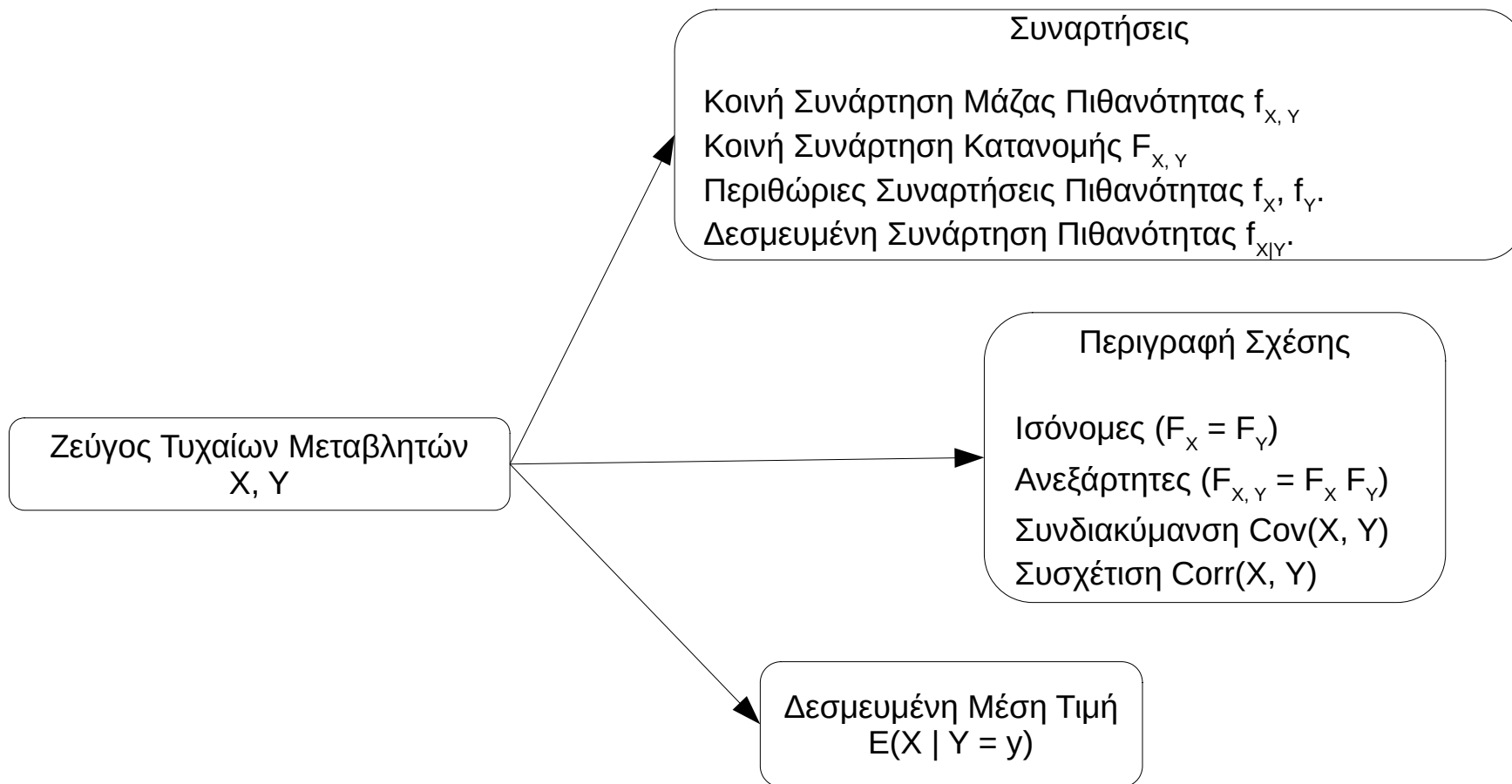
$$\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y) = 0,04.$$

$$\text{(θ)} \quad \text{Corr}(X, Y) = \text{Cov}(X, Y) / (\sigma_X \cdot \sigma_Y) = 0,04 / (0,7 \cdot 0,49) = 0,117.$$

(ι) $f_{XY}(1, 0) = 0,2 \neq 0,18 = f_X(1) \cdot f_Y(0)$ άρα οι X, Y είναι εξαρτημένες.

$f_{X,Y}$	$X=1$	$X=2$	$X=3$	f_Y
$Y=0$	0,2	0,3	0,1	0,6
$Y=1$	0,1	0,2	0,1	0,4
f_X	0,3	0,5	0,2	1

Γραφική Σύνοψη



Συνήθεις κατανομές διακριτών μεταβλητών

Συνήθεις κατανομές διακριτών μεταβλητών

κατανομή	συνάρτηση πιθανότητας	παράμετροι	μέση τιμή	διακύμανση
Bernoulli	$p(1 - p)$	$p \in (0, 1)$	p	$p(1 - p)$
Διωνυμική	$\binom{n}{k} p^k (1 - p)^{n-k}$	$p \in (0, 1), n \in \mathbb{N}$	np	$np(1 - p)$
Αρνητική διωνυμική	$\binom{r+k-1}{r-1} p^r (1 - p)^k$	$p \in (0, 1), r \in \mathbb{N}$	$r \frac{p}{1 - p}$	$r \frac{1 - p}{p^2}$
Γεωμετρική	$p(1 - p)^{n-1}$	$p \in (0, 1)$	$\frac{1}{p}$	$\frac{1 - p}{p^2}$
Υπεργεωμετρική	$\frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$	$n, N, K \in \mathbb{N}, n, K \leq N$	$\frac{nK}{N}$	$\frac{nK(N - K)(N - n)}{N^2(N - 1)}$
Poisson	$\frac{\lambda^k}{k!} e^{-\lambda}$	$\lambda \in \mathbb{R}_+$	λ	λ

Κατανομή Bernoulli

Κατανομή Bernoulli

Η **κατανομή Μπερνούλλι** (Bernoulli) είναι μια συνάρτηση κατανομής διακριτής τυχαίας μεταβλητής. Με αυτήν περιγράφεται ένα τυχαίο πείραμα με δύο πιθανά αποτελέσματα (επιτυχία - αποτυχία) και πιθανότητα επιτυχίας p .

Πιο συγκεκριμένα, αν X είναι η τ.μ. που παίρνει τιμές 0 (αποτυχία) ή 1 (επιτυχία) και $X \sim \text{Bernoulli}(p)$ (γράφουμε και $X \sim B(1, p)$) τότε:

$$P(X = 1) = p \text{ και } P(X = 0) = q = 1 - p.$$

Η κατανομή Bernoulli αξιοποιείται για την μοντελοποίηση πειραμάτων με δυο πιθανά αποτελέσματα.

Κατανομή Bernoulli

Άσκηση

Να βρεθεί η αναμενόμενη τιμή, η διακύμανση και η τυπική απόκλιση της $X \sim \text{Bernoulli}(p)$.

$$X=0 \text{ με } P(X=0)=1-p, \quad X=1 \text{ με } P(X=1)=p.$$

$$E X = 0 \cdot (1-p) + 1 \cdot p = p. \quad E(X^2) = 0^2 \cdot (1-p) + 1^2 \cdot p = p.$$

$$\text{Var}(X) = E(X^2) - (E X)^2 = p - p^2 = p(1-p).$$

$$\sigma = \sqrt{\text{Var}(X)} = \sqrt{p \cdot (1-p)}.$$

Bernoulli: Γεωμετρικά χαρακτηριστικά

$X \sim \text{Bernoulli}(p)$: $P(X = 1) = p$ και $P(X = 0) = q = 1 - p$.

Αναμενόμενη τιμή: $EX = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = p$.

Διακύμανση: $\text{Var}X = EX^2 - (EX)^2 = p - p^2 = p \cdot (1 - p) = p \cdot q$.

Τυπική απόκλιση: $\sigma = \sqrt{p \cdot (1 - p)}$

Ασυμμετρία: $\text{skew}(X) = \gamma = \frac{\mu_3}{\sigma^3} = \frac{1 - 2p}{\sqrt{p(1 - p)}}$

Κυρτότητα: $\text{kurt}(X) = \alpha = \frac{\mu_4}{\sigma^4} = \frac{1 - 3p(1 - p)}{p(1 - p)}$

Διωνυμική κατανομή

Διωνυμική κατανομή

Άσκηση

Έστω ένα πείραμα που εκτελείται 8 φορές, ανεξάρτητα η μία από την άλλη και ότι σε κάθε μία επανάληψη υπάρχει πιθανότητα επιτυχίας p και αποτυχίας $1 - p$. Αν X είναι το πλήθος των επιτυχιών, να βρεθεί η πιθανότητα $P(X = 3)$.

Λύση

$$X = 0, 1, \dots, 8$$

$$P(X=3) = \binom{8}{3} \cdot p^3 \cdot (1-p)^5$$

$$\underline{1} \quad \underline{1} \quad \underline{1} \quad \underline{0} \quad \underline{0} \quad \underline{0} \quad \underline{0} \quad \underline{0} \quad : p^3 \cdot (1-p)^5$$

$\binom{8}{3} = C(8,3)$: αριθμός τοποθετήσεων των 3 \uparrow στις 8 θέσεις
επιτυχία

Διωνυμική κατανομή

Γέννηση της Διωνυμικής Κατανομής

Έστω ένα πείραμα που εκτελείται n φορές, ανεξάρτητα η μία από την άλλη και ότι σε κάθε μία επανάληψη υπάρχει πιθανότητα επιτυχίας p και αποτυχίας $1 - p$. Αν X είναι το πλήθος των επιτυχιών, να βρεθεί η πιθανότητα $P(X = k)$, $k = 0, 1, 2, \dots, n$.

Λύση

Διωνυμική κατανομή

Έστω ότι ένα πείραμα επαναλαμβάνεται n φορές και κάθε φορά υπάρχει σταθερή πιθανότητα επιτυχίας p . Τότε, αν η τ.μ. X μετράει το πλήθος επιτυχιών στα n ανεξάρτητα πειράματα, είναι

$$P(X = \kappa) = \binom{n}{\kappa} p^{\kappa} (1 - p)^{n - \kappa}, \quad \kappa = 0, 1, 2, \dots, n.$$

Λέμε ότι η διακριτή τ.μ. X ακολουθεί τη **διωνυμική κατανομή** (binomial distribution) και γράφουμε $X \sim \mathbf{B}(n, p)$.

Σημείωση

Έστω ότι σε ένα σάκο υπάρχουν N αντικείμενα από τα οποία τα K έχουν μία ιδιότητα και προσδιορίζουν την “επιτυχία” στην επιλογή. Αν επιλέξουμε n αντικείμενα με επανατοποθέτηση και $X = \{\text{πλήθος επιτυχιών στις } n \text{ επιλογές}\}$ τότε $X \sim B(n, p)$, όπου $p = K/N$.

Διωνυμική: Γεωμετρικά χαρακτηριστικά

$$X \sim B(n, p): P(X = \kappa) = \binom{n}{\kappa} p^\kappa (1-p)^{n-\kappa}, \kappa = 0, 1, 2, \dots, n.$$

Αναμενόμενη τιμή: $\underline{EX = n \cdot p.}$ $E(X) = \sum_{\kappa=0}^n \binom{n}{\kappa} \cdot p^\kappa \cdot (1-p)^{n-\kappa} = n \cdot p.$

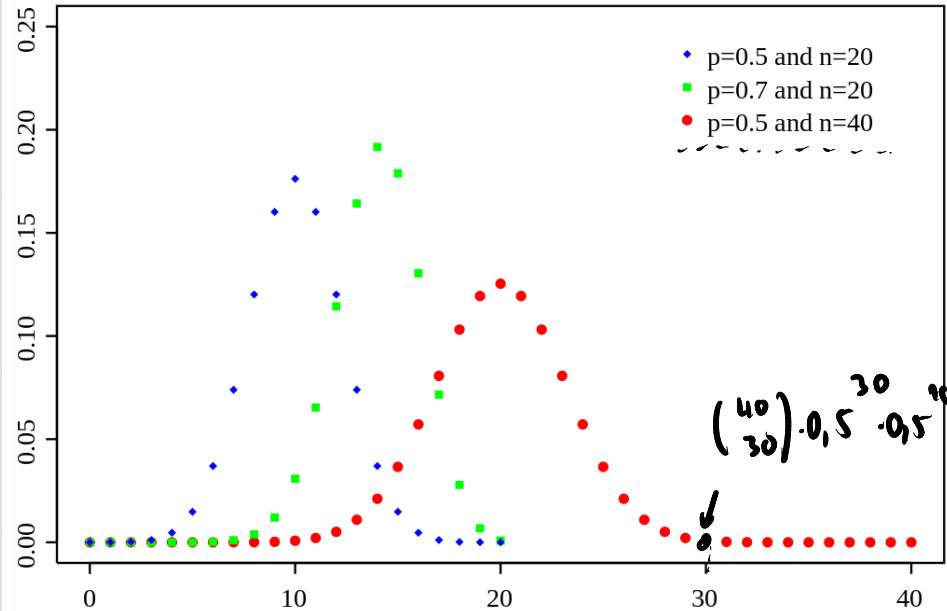
Διακύμανση: $\text{Var}X = n \cdot p \cdot q = n \cdot p \cdot (1-p).$

Τυπική απόκλιση: $\sigma = \sqrt{n \cdot p \cdot (1-p)}$

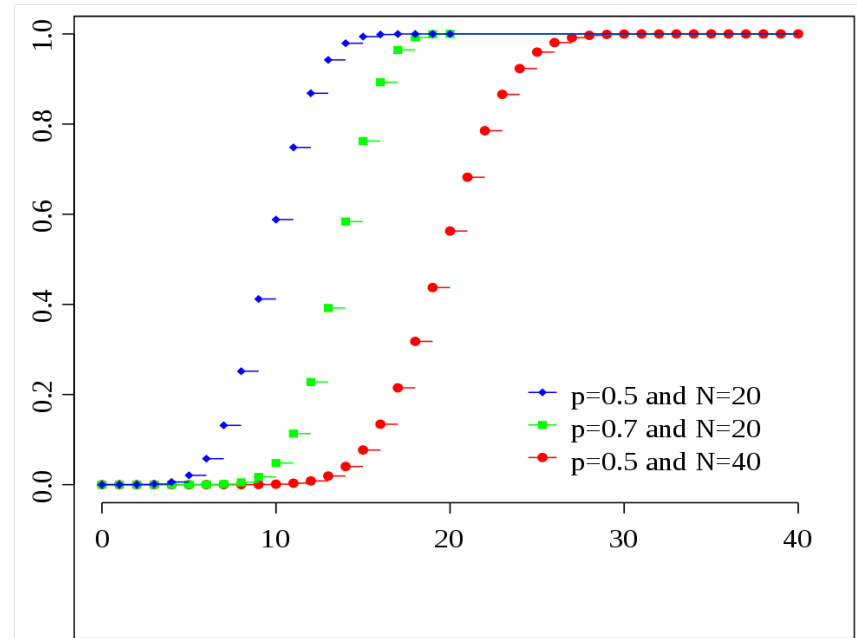
Ασυμμετρία: $\text{skew}(X) = \gamma = \frac{\mu_3}{\sigma^3} = \frac{1-2p}{\sqrt{np(1-p)}}$

Κυρτότητα: $\text{kurt}(X) = \alpha = \frac{\mu_4}{\sigma^4} = \frac{1-6p(1-p)}{np(1-p)}$

Διωνυμική κατανομή



Συνάρτηση πιθανότητας $P(X = i)$, $i = 0, \dots, n$



Συνάρτηση κατανομής $F(x) = P(X \leq x)$, $x \in \mathbb{R}$

Πιθανογεννήτρια συνάρτηση διωνυμικής

Παράδειγμα

Αν $X \sim B(n, p)$ τότε $P(X = \kappa) = \binom{n}{\kappa} p^\kappa (1 - p)^{n - \kappa}$ και

$$\begin{aligned} G_X(z) &= E(z^X) = \sum_{\kappa=0, 1, \dots} P(X = \kappa) \cdot z^\kappa \\ &= \sum_{\kappa=0, 1, \dots} \binom{n}{\kappa} p^\kappa (1 - p)^{n - \kappa} \cdot z^\kappa \\ &= \sum_{\kappa=0, 1, \dots} \binom{n}{\kappa} (zp)^\kappa (1 - p)^{n - \kappa} \\ &= (zp + 1 - p)^n \\ &= (1 + (z - 1)p)^n. \end{aligned}$$

Άρα, $G_X(z) = (1 + (z - 1)p)^n$ για **κάθε** $z \in \mathbf{C}$, ειδικότερα,

$G_X(x) = (1 + (x - 1)p)^n$, για **κάθε** $x \in \mathbf{R}$.

Ροπογεννήτρια συνάρτηση διωνυμικής

Άσκηση

Αν $X \sim B(n, p)$ να δείξετε ότι $M_X(t) = E(e^{tx}) = (q + pe^t)^n$.

Απόδειξη

Αναγνώριση διωνυμικής κατανομής

Ένα τυχαίο πείραμα είναι διωνυμικό πείραμα και υποστηρίζει τον ορισμό μίας διωνυμικής τυχαίας μεταβλητής όταν:

- έχει σταθερό αριθμό (n) δοκιμών.
- κάθε δοκιμή είναι ανεξάρτητη από τις άλλες.
- κάθε δοκιμή έχει μόνο δύο πιθανά αποτελέσματα, «επιτυχία» και «αποτυχία».
- η πιθανότητα επιτυχίας p για κάθε δοκιμή, είναι σταθερή και το συμπλήρωμα της είναι η πιθανότητα αποτυχίας $q = 1 - p$.

Ασκήσεις στη διωνυμική κατανομή

1. Αν $X \sim B(3, 0,5)$ να βρεθεί η συνάρτηση μάζας πιθανότητας της X .

Λύση

$$\binom{3}{0} = \frac{3!}{0! \cdot (3-0)!}$$

$n = 3$ ανεξ. εν. πειρ. $p = 0,5$: πιθαν. επιτυχ.

$$\binom{3}{1} = \frac{3!}{1! \cdot (3-1)!}$$

$X = \{\text{αριθμός επιτυχιών στις 3 εναν.}\} = 0, 1, 2, 3$

$$P(X=0) = \binom{3}{0} \cdot 0,5^0 \cdot (1-0,5)^3 = 1 \cdot 1 \cdot 0,5^3 = 0,5^3 = 0,125.$$

$$P(X=1) = \binom{3}{1} \cdot 0,5^1 \cdot (1-0,5)^2 = 3 \cdot 0,5 \cdot 0,5^2 = 0,375$$

$$P(X=2) = \binom{3}{2} \cdot 0,5^2 \cdot (1-0,5)^1 = 3 \cdot 0,5^2 \cdot 0,5 = 0,375$$

$$P(X=3) = \binom{3}{3} \cdot 0,5^3 \cdot (1-0,5)^0 = 1 \cdot 0,5^3 \cdot 1 = 0,125.$$

Ασκήσεις στη διωνυμική κατανομή

2. Γνωρίζουμε ότι το 60% των φοιτητών της σχολής ζει μόνιμα στην πόλη της Ξάνθης. Επιλέγουμε τυχαία 5 φοιτητές από τον κατάλογο όλων των φοιτητών που σπουδάζουν στη σχολή. Να βρεθεί η πιθανότητα:

(α) Οι 3 να ζουν μόνιμα στην πόλη.

(β) Τουλάχιστον 4 να ζουν μόνιμα στην πόλη.

(γ) Το πολύ 2 να ζουν μόνιμα στην πόλη.

Λύση

$$X = \{\text{πλήθος φοιτητών στους 5 που ζουν στην Ξάνθη}\} \sim B(5, 0.6)$$

$$(α) P(X=3) = \binom{5}{3} \cdot 0.6^3 \cdot 0.4^2 = \frac{5!}{3! \cdot 2!} \cdot 0.6^3 \cdot 0.4^2 = \frac{5 \cdot 4 \cdot 3!}{3! \cdot 2} \cdot 0.6^3 \cdot 0.4^2 = 0,3456.$$

$$(β) P(X \geq 4) = P(X=4) + P(X=5) = \binom{5}{4} \cdot 0.6^4 \cdot 0.4^1 + \binom{5}{5} \cdot 0.6^5 \cdot 0.4^0 = \dots$$

$$(γ) P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = \dots$$

Ασκήσεις στη διωνυμική κατανομή

3. Ένας παίκτης παίζει ρουλέτα και ποντάρει συνεχώς στο μαύρο χρώμα. Η πιθανότητα να έρθει μαύρο σε ένα οποιοδήποτε γύρισμα της ρουλέτας είναι ίση με $18/37$. Να βρεθεί η πιθανότητα να κερδίσει 4 φορές σε 10 γυρίσματα της ρουλέτας.

Λύση

$$X = \{\text{αριθμός νικών στις 10}\} \sim B\left(10, \frac{18}{37}\right)$$

$$P(X=4) = \binom{10}{4} \cdot \left(\frac{18}{37}\right)^4 \cdot \left(1 - \frac{18}{37}\right)^6 = \dots$$

$$\binom{10}{4} = \frac{10!}{4! \cdot 6!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot \cancel{6!}}{2 \cdot 3 \cdot 4 \cdot \cancel{6!}} = 210.$$

Ασκήσεις στη διωνυμική κατανομή

4. Μία εταιρεία λιπασμάτων ισχυρίζεται ότι το λίπασμα της αυξάνει την παραγωγή του 80% των χωραφιών που εφαρμόζεται. Ένας συνεταιρισμός, χρησιμοποίησε το λίπασμα της εταιρείας σε 7 χωράφια και παρατήρησαν αύξηση στην παραγωγή στα 3 από αυτά. Πόσο πιθανό ήταν να συμβεί αυτό, αν ο ισχυρισμός της εταιρείας είναι σωστός;

Λύση

Ασκήσεις στη διωνυμική κατανομή

5. Πόσα παιδιά πρέπει να αποκτήσει μια οικογένεια ώστε να έχει ένα τουλάχιστον αγόρι και ένα τουλάχιστον κορίτσι με πιθανότητα μεγαλύτερη από α) 95%, β) 99%

Λύση

Έστω n παιδιά. $X = \{\text{ηλικίους αγόριων σε } n\} \sim B(n, 0.5)$

$$\begin{aligned} p &= P(\text{τουλ. 1 Α + τουλ. 1 Κ σε } n) = P(0 < X < n) = 1 - P(X=0) - P(X=n) \\ &= 1 - \binom{n}{0} \cdot 0.5^0 \cdot 0.5^n - \binom{n}{n} \cdot 0.5^n \cdot 0.5^0 = 1 - 2 \cdot 0.5^n \end{aligned}$$

$$\begin{aligned} \text{α) Πρέπει } p > 0.95 &\Leftrightarrow 1 - 2 \cdot 0.5^n > 0.95 \Leftrightarrow 2 \cdot 0.5^n < 0.05 \Leftrightarrow \\ &\Leftrightarrow 0.5^n < 0.025 \Leftrightarrow n > \log_{0.5}(0.025) \end{aligned}$$

Ασκήσεις στη διωνυμική κατανομή

6. Ένα test πολλαπλής επιλογής αποτελείται από 20 ερωτήσεις. Για κάθε ερώτηση, υπάρχουν 4 πιθανές απαντήσεις και μια μόνο από τις οποίες είναι σωστή. Κάθε σωστή απάντηση παίρνει 1 μονάδα. Να βρεθεί ο ~~μέσος~~ αναμενόμενος βαθμός των φοιτητών που απαντούν στην τύχη.

Λύση

$p = 0.25$: πιθανότητα σωστής απάντησης.

$X = \{ \text{βαθμολογία} \} = \{ \text{αριθμός σωστών στις 20} \} \sim B(20, 0.25)$

$$E(X) = n \cdot p = 20 \cdot 0.25 = 5.$$

Ασκήσεις στη διωνυμική κατανομή

7. Έχουμε 4 tablet, τα οποία έχουν όλα ίδια πιθανότητα p να είναι ελαττωματικά. Γνωρίζουμε ότι η πιθανότητα να μην υπάρχει ελαττωματικό tablet είναι ίση με την πιθανότητα να είναι όλα ελαττωματικά. Αν $X = \{\text{πλήθος ελαττωματικών tablet}\}$, τότε:

- (α) Να βρεθεί η συνάρτηση πιθανότητας της τ.μ. X
- (β) Να βρεθεί η μέση τιμή και η τυπική απόκλιση της κατανομής.

Κατανομή Poisson

Κατανομή Poisson

Έστω ότι κάποιες αφίξεις συμβαίνουν **εντελώς τυχαία** στο χρόνο και **ανεξάρτητα μεταξύ τους** με ρυθμό λ αφίξεις / μονάδα χρόνου. Έστω, επίσης

$$X = \{\text{πλήθος αφίξεων στο διάστημα } (0, 1] \} = 0, 1, 2, \dots$$

Αποδεικνύεται ότι:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

Την κατανομή αυτή την ονομάζουμε **κατανομή Poisson** με παράμετρο λ και γράφουμε **$X \sim \text{Poisson}(\lambda)$** ή απλά **$X \sim P(\lambda)$** .

Η γέννηση της κατανομής Poisson ως όριο δοκιμασιών Bernoulli

Απόδειξη (1/2)

Επιλέγουμε $n \in \mathbb{N}$ και διαμερίζουμε το διάστημα $(0, 1]$ σε n υποδιαστήματα μήκους $\delta = 1/n > 0$:
 $(0, \delta], (\delta, 2\delta], (2\delta, 3\delta], \dots, (1 - \delta, 1],$

Σε κάθε ένα διάστημα αναμένονται $\lambda \cdot \delta = \lambda/n$ αφίξεις. Αν $X_i = 1$ όταν υπάρχει άφιξη στην i διαμέριση του $(0, 1]$ και $X_i = 0$ όταν δεν υπάρχει, τότε $EX_i = \lambda/n$, και

$$X_i \sim \text{Bernoulli}(p_\delta), \quad i = 1, 2, \dots, n, \quad \text{με } p_\delta = \lambda/n.$$

Για τις μεταβλητές $X_i, i = 1, 2, \dots, n$, σημειώνουμε ότι:

(α) Είναι στοχαστικά ανεξάρτητες (τυχαίες αφίξεις)

(β) $X_1 + X_2 + \dots + X_n = \{\text{σύνολο αφίξεων στο } (0, \delta] \cup (\delta, 2\delta] \cup (2\delta, 3\delta] \cup \dots \cup (1 - \delta, 1] \}$

(γ) $E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = n \cdot \lambda \cdot \delta = \lambda$

Αν $Y_n = X_1 + X_2 + \dots + X_n$ τότε $Y_n \sim B(n, p_\delta)$.

Σημείωση

Ο λόγος λ/n , για n αρκούντως μεγάλο είναι μικρότερο της μονάδας και μπορεί να θεωρηθεί πιθανότητα.

Η γέννηση της κατανομής Poisson ως όριο δοκιμασιών Bernoulli

Απόδειξη (2/2)

Καθώς $\delta = 1/n$, είναι $\lim_{n \rightarrow \infty} \delta = \lim_{n \rightarrow \infty} 1/n = 0$ και η διαμέριση $(0, \delta], (\delta, 2\delta], (2\delta, 3\delta], \dots, (1 - \delta, 1]$ προσεγγίζει την σημειακή ανάλυση του διαστήματος $(0, 1]$, επιτρέποντας με τον τρόπο αυτό την καταγραφή των αφίξεων σε οποιαδήποτε χρονική στιγμή του διαστήματος. Δηλαδή,

$$\lim_{n \rightarrow \infty} Y_n = \lim_{n \rightarrow \infty} (X_1 + X_2 + \dots + X_n) = \{\text{σύνολο αφίξεων στο } (0, 1]\} = X$$

Είναι

$$P(X = k) = \lim_{n \rightarrow \infty} P_{Y_n}(Y_n = k) = \lim_{n \rightarrow \infty} C(n, k) p_\delta^k (1 - p_\delta)^{n-k} = e^{-\lambda} \lambda^k / k!, \quad k = 0, 1, 2, \dots, (*)$$

δηλαδή, αποδείχθηκε ότι:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

(*) Η απόδειξη της τελευταίας ισότητας γίνεται στις επόμενες 4 διαφάνειες.

Απόδειξη του $\lim_{n \rightarrow \infty} C(n, k) p_{\delta}^k (1 - p_{\delta})^{n-k} = e^{-\lambda} \lambda^k / k!$ (1/4)

Από τον ορισμό του ο αριθμός e είναι ίσος με

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^n$$

Ένα χρήσιμο όριο που προκύπτει από τον ορισμό είναι το

$$e^{\lambda} = \lim_{n \rightarrow \infty} \left(1 + \frac{\lambda}{n} \right)^n, \lambda \in \mathbb{R}.$$

Απόδειξη του $\lim_{n \rightarrow \infty} C(n, k) p_{\delta}^k (1 - p_{\delta})^{n-k} = e^{-\lambda} \lambda^k / k!$ (2/4)

$$e^{\lambda} = \lim_{n \rightarrow \infty} \left(1 + \frac{\lambda}{n}\right)^n, \quad \lambda \in \mathbb{R}.$$

Απόδειξη

Θέτουμε $y_n = \left(1 + \frac{\lambda}{n}\right)^n$, άρα $\ln y_n = n \ln \left(1 + \frac{\lambda}{n}\right) = \frac{\ln \left(1 + \frac{\lambda}{n}\right)}{\frac{1}{n}}$

Είναι $\lim_{n \rightarrow \infty} \ln y_n = \lim_{n \rightarrow \infty} \frac{\ln \left(1 + \frac{\lambda}{n}\right)}{\frac{1}{n}} = \lim_{n \rightarrow \infty} \frac{\frac{1}{1 + \frac{\lambda}{n}} \left(-\frac{\lambda}{n^2}\right)}{-\frac{1}{n^2}} = \lambda.$

Άρα $\lim_{n \rightarrow \infty} \left(1 + \frac{\lambda}{n}\right)^n = \lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} e^{\ln y_n} = e^{\lambda}.$

Απόδειξη του $\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = e^{-\lambda} \lambda^k / k!$ (3/4)

Θεώρημα

Έστω $Y_n \sim B(n, p_n)$, $p_n = p(n)$ και έστω ότι υπάρχει $\lambda > 0$ τέτοιο ώστε $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda$. Τότε

$$\lim_{n \rightarrow \infty} P_{Y_n}(Y_n = k) = e^{-\lambda} \lambda^k / k!, \quad k = 0, 1, 2, \dots$$

Απόδειξη

Αν $Y_n \sim B(n, p_n)$, τότε $P_{Y_n}(X = k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k}$. Καθώς, $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda$, είναι $|n \cdot p_n - \lambda| < \varepsilon$ ή $|p_n - \lambda / n| < \varepsilon / n$, δηλαδή $\lim_{n \rightarrow \infty} (p_n - \lambda / n) = 0$. Είναι:

$$\begin{aligned} \lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} &\simeq \lim_{n \rightarrow \infty} \frac{n(n-1)(n-2) \dots (n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \lim_{n \rightarrow \infty} \frac{n^k + O(n^{k-1})}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \lim_{n \rightarrow \infty} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

Απόδειξη του $\lim_{n \rightarrow \infty} \binom{n}{k} p_{\delta}^k (1 - p_{\delta})^{n-k} = e^{-\lambda} \lambda^k / k!$ (4/4)

Θεώρημα

Έστω $Y_n \sim B(n, p_n)$, $p_n = p(n)$ και έστω ότι υπάρχει $\lambda > 0$ τέτοιο ώστε $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda$. Τότε

$$\lim_{n \rightarrow \infty} P_{Y_n}(Y_n = \kappa) = e^{-\lambda} \lambda^{\kappa} / \kappa!, \kappa = 0, 1, 2, \dots$$

Απόδειξη

Καθώς, $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$ και $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = 1$ (γιατί;), βρίσκουμε

$$\binom{n}{k} p^k (1 - p)^{n-k} \simeq \frac{\lambda^k e^{-\lambda}}{k!}.$$

Κατανομή Poisson

Συνοψίζοντας:

$X \sim \text{Poisson}(\lambda)$, αν για κάποιο $\lambda > 0$, $P(X = \kappa) = e^{-\lambda} \lambda^\kappa / \kappa!$, $\kappa = 0, 1, 2, \dots$

Η κατανομή Poisson εφαρμόζεται σε φαινόμενα με διακριτή συχνότητα και σταθερή πιθανότητα να συμβούν στη μονάδα μέτρησης χρόνου ή χώρου.

Η παράμετρος λ συμβολίζει την αναμενόμενη τιμή της X στην μονάδα μέτρησης.

Κατανομή Poisson

Παραδείγματα τ.μ. που ακολουθούν την Poisson κατανομή:

- οι τηλεφωνικές κλήσεις που φτάνουν σε ένα σύστημα (X : πλήθος κλήσεων, λ : μέσο πλήθος κλήσεων).
- τα φωτόνια που φτάνουν σε ένα τηλεσκόπιο (X : πλήθος φωτονίων, λ : μέσο πλήθος φωτονίων που φτάνουν).
- ο αριθμός των μεταλλάξεων σε ένα κλώνο του DNA ανά μονάδα μήκους (X : πλήθος μεταλλάξεων, λ : μέσο πλήθος μεταλλάξεων που παρατηρούνται).
- Ο αριθμός των γκολ σε αγώνες ποδοσφαίρου (X : πλήθος γκολ, λ : μέσο πλήθος γκολ σε ανάλογους αγώνες).

Κατανομή Poisson

Άσκηση

Να αποδείξετε ότι αν $X \sim \text{Poisson}(\lambda)$ τότε $E(X) = \lambda$.

Απόδειξη

$$\begin{aligned} E(X) &= \sum_{k=0}^{+\infty} k \cdot P(X=k) = \sum_{k=0}^{+\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \cdot \lambda \\ &= e^{-\lambda} \cdot e^{\lambda} \cdot \lambda = \lambda. \end{aligned}$$

Poisson: Γεωμετρικά χαρακτηριστικά

$$Y \sim \text{Poisson}(\lambda): P(X = \kappa) = e^{-\lambda} \lambda^{\kappa} / \kappa!, \kappa = 0, 1, 2, \dots$$

Αναμενόμενη τιμή: $EX = \lambda.$

Διακύμανση: $\text{Var}X = \lambda.$

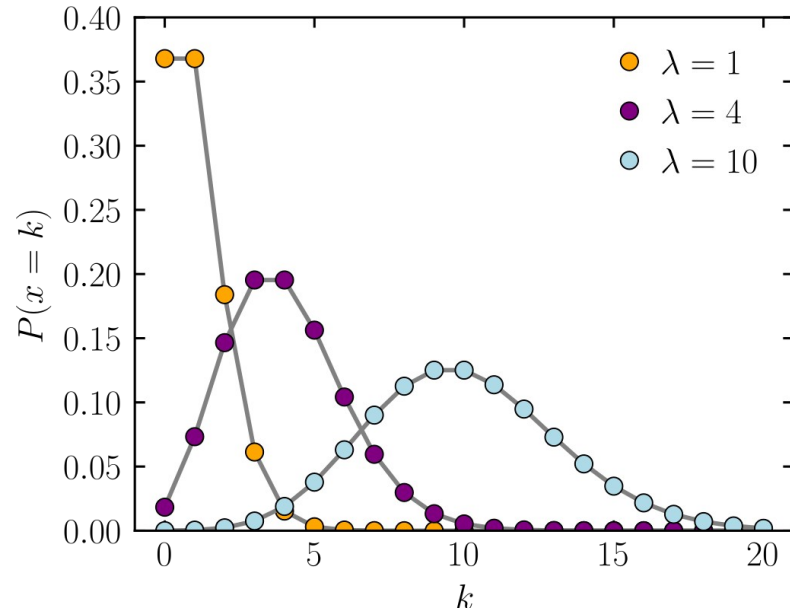
Τυπική απόκλιση: $\sigma = \sqrt{\lambda}$

Ασυμμετρία: $\text{skew}(X) = \gamma = \frac{\mu_3}{\sigma^3} = \frac{1}{\sqrt{\lambda}}$

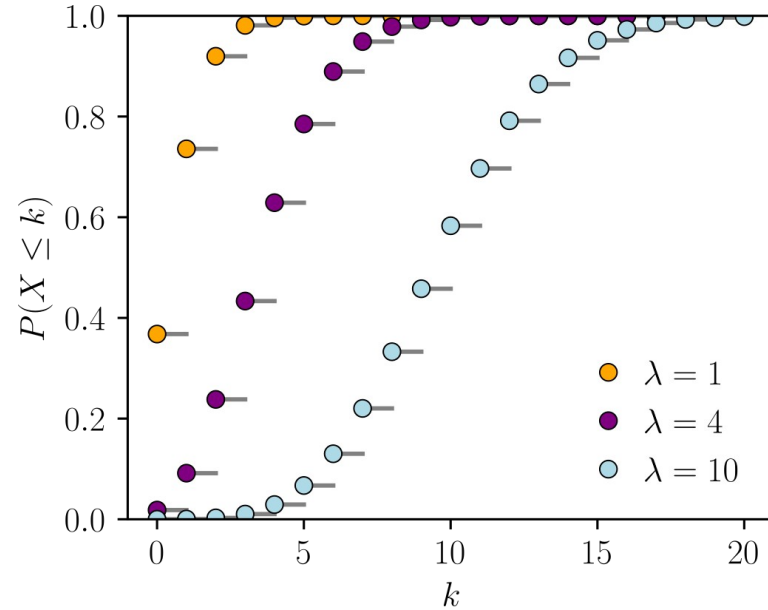
Κυρτότητα: $\text{kurt}(X) = \alpha = \frac{\mu_4}{\sigma^4} = \frac{1}{\lambda}$

Σημείωση: Από τον τρόπο ορισμού της η Poisson έχει πάντα θετική ασυμμετρία (ουρά προς τα δεξιά)

Κατανομή Poisson

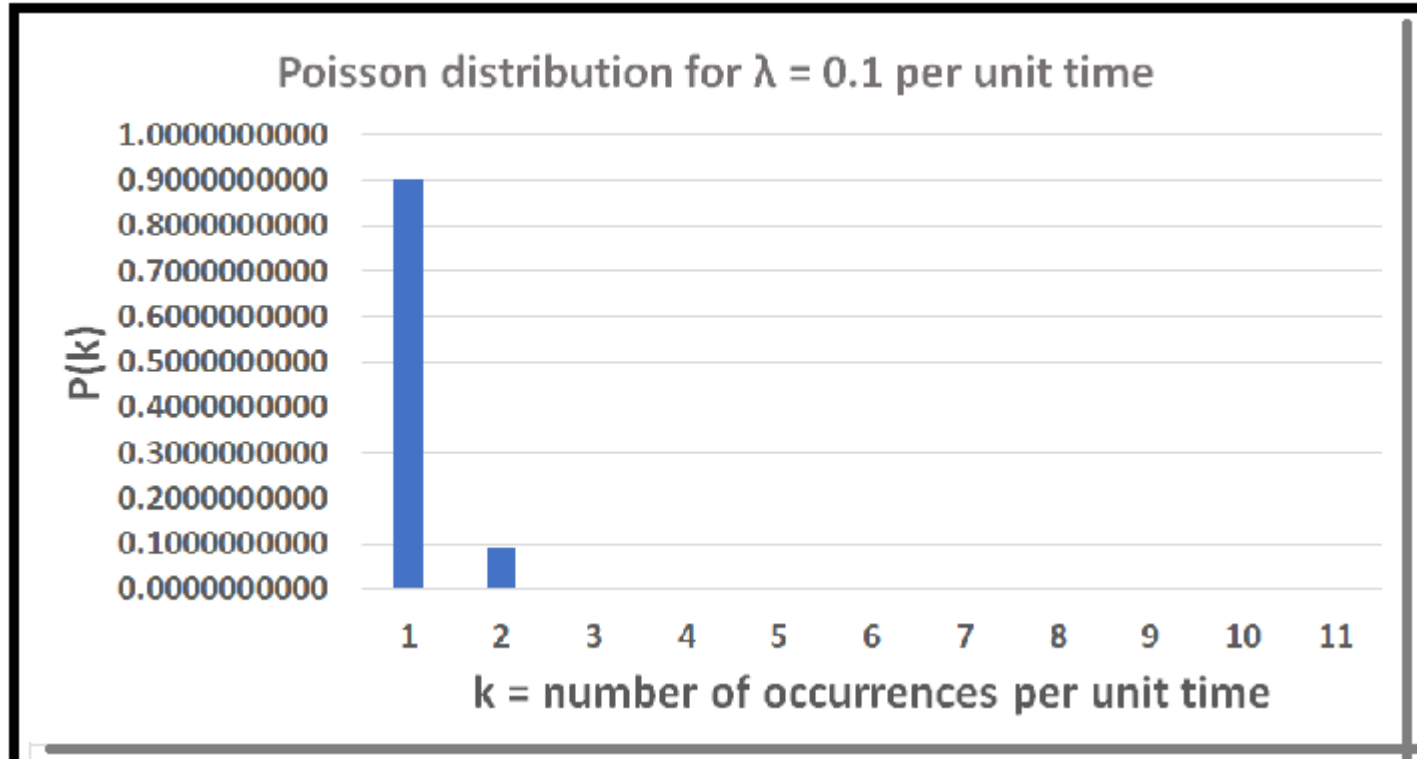


Συνάρτηση πιθανότητας $P(X = \kappa) = e^{-\lambda} \lambda^\kappa / \kappa!$, $\kappa = 0, 1, 2, \dots$

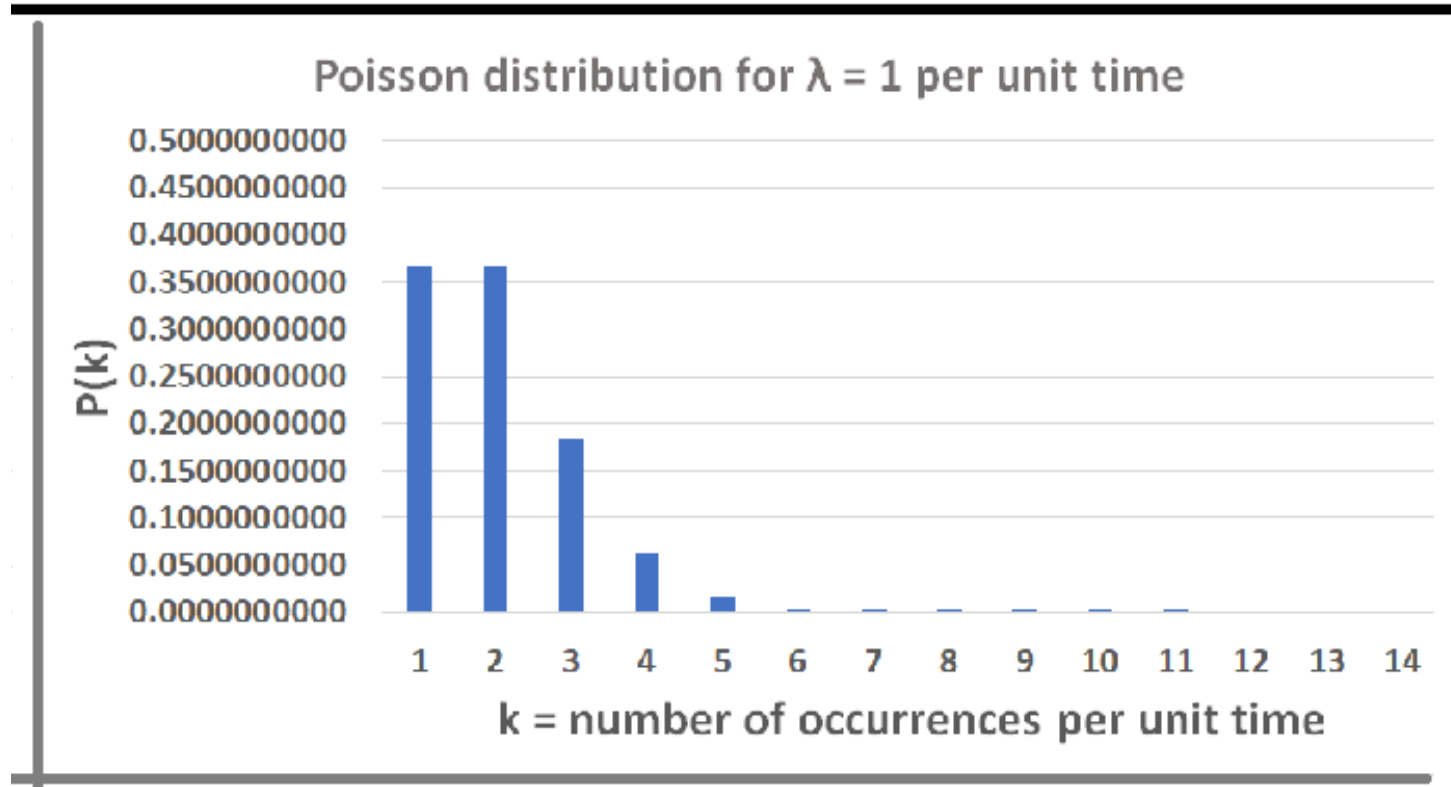


Συνάρτηση κατανομής $F(x) = P(X \leq x)$, $x \in \mathbb{R}$

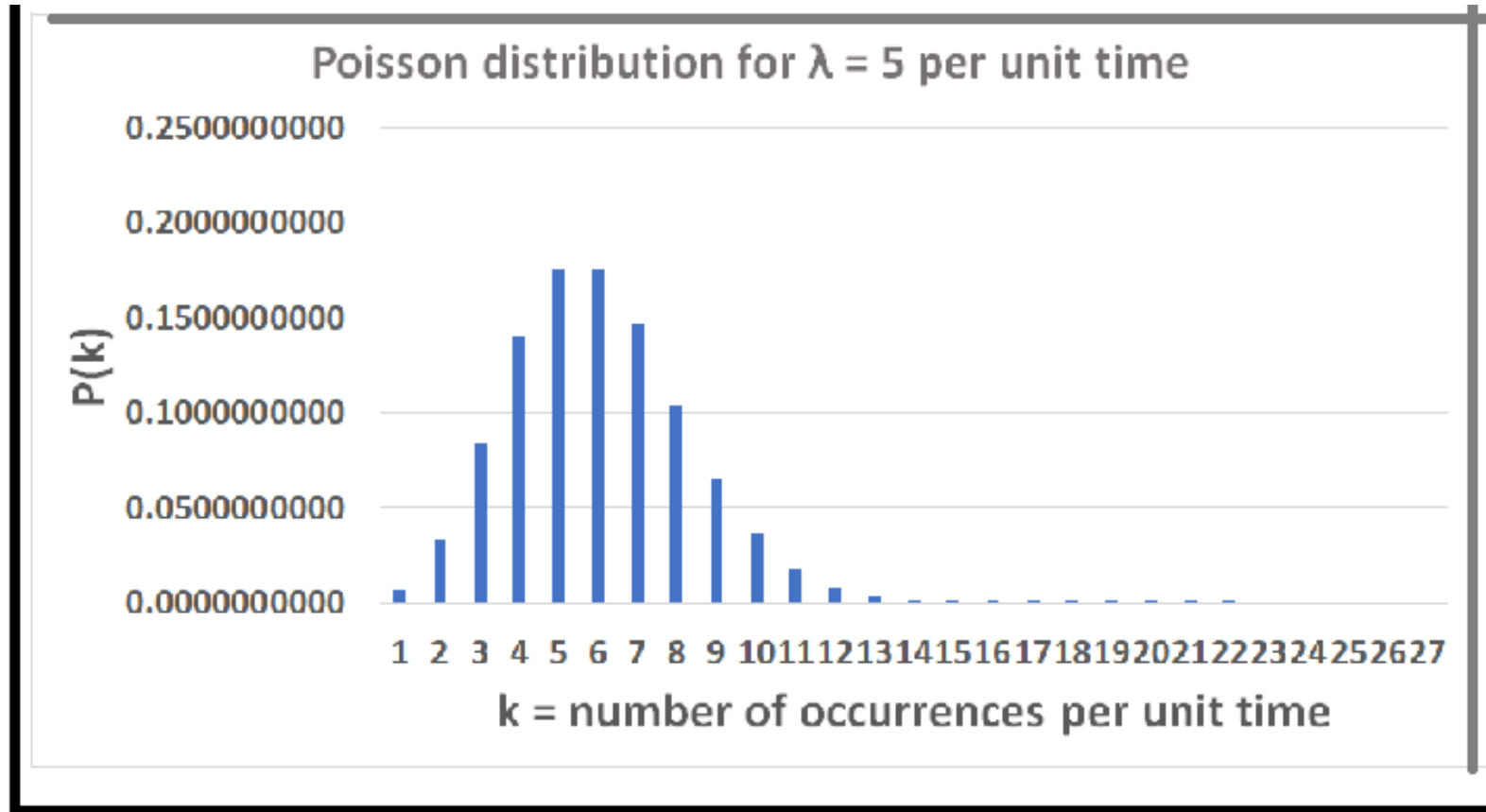
$$\lambda = 0,1: P(X = \kappa) = e^{-0,1} \cdot 0,1^\kappa / \kappa!, \kappa = 0, 1, \dots$$



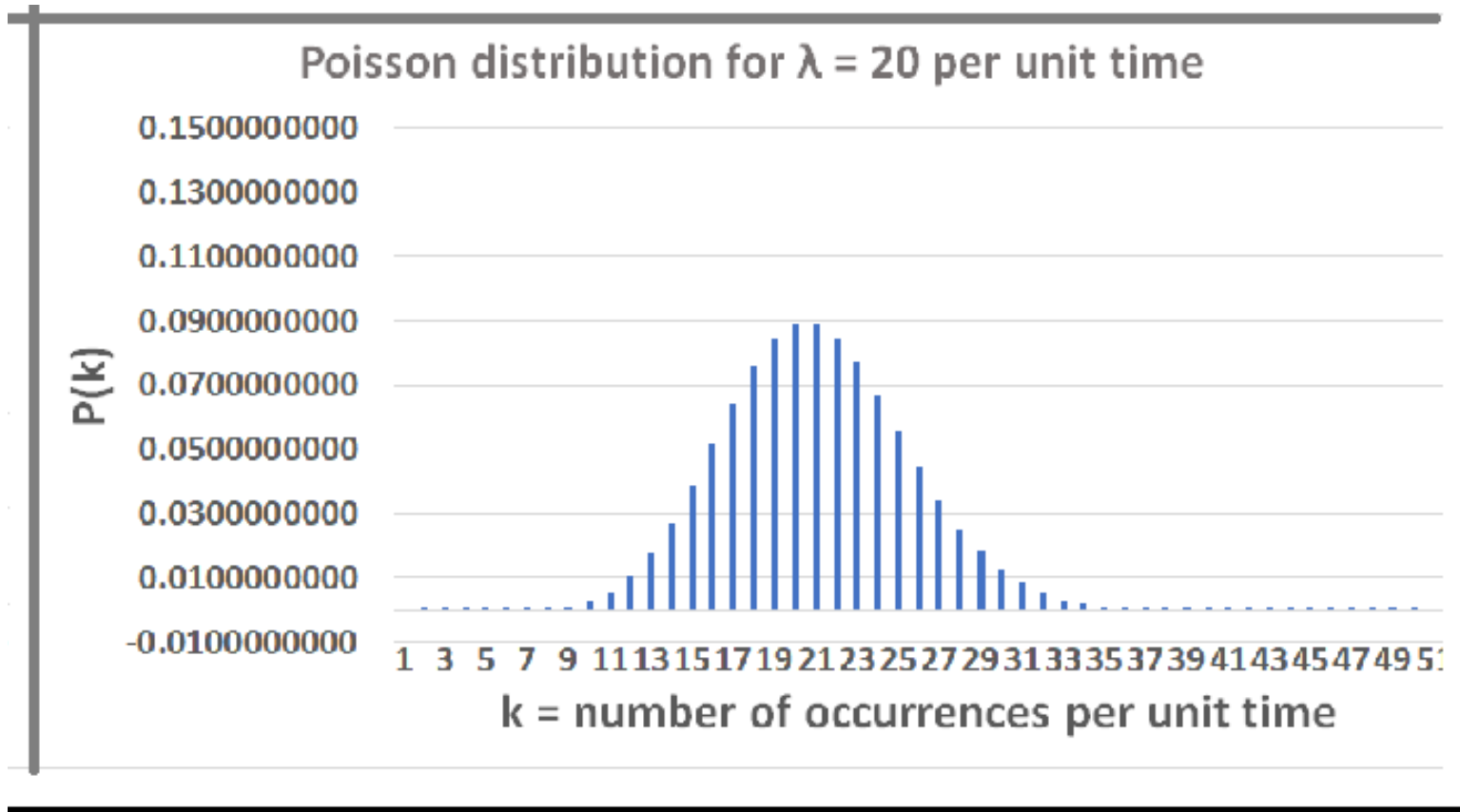
$$\lambda = 1: P(X = \kappa) = e^{-1} / \kappa!, \kappa = 0, 1, \dots$$



$$\lambda = 5: P(X = \kappa) = e^{-5} \cdot 5^{\kappa} / \kappa!, \kappa = 0, 1, \dots$$



$$\lambda = 20: P(X = \kappa) = e^{-20} \cdot 20^\kappa / \kappa!, \kappa = 0, 1, \dots$$



Παράδειγμα 1

Ο αριθμός των σωματιδίων που εκπέμπει μια ραδιενεργός πηγή ανά δευτερόλεπτο ακολουθεί την κατανομή Poisson με μέσο $\lambda = 5$. Να υπολογιστούν οι πιθανότητες: $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ και $P(X \geq 3)$.

Λύση

Παράδειγμα 1

Ο αριθμός των σωματιδίων που εκπέμπει μια ραδιενεργός πηγή ανά δευτερόλεπτο ακολουθεί την κατανομή Poisson με μέσο $\lambda = 5$. Να υπολογιστούν οι πιθανότητες: $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ και $P(X \geq 3)$.

Λύση

Είναι $X \sim \text{Poisson}(5)$ άρα $P(X = \kappa) = e^{-5} \cdot 5^\kappa / \kappa!$. Υπολογίζουμε:

- $P(X = 0) = e^{-5} \cdot 5^0 / 0! = e^{-5} = 0,0067 = 0,67\%$.
- $P(X = 1) = e^{-5} \cdot 5^1 / 1! = 5 \cdot e^{-5} = 0,0337 = 3,37\%$.
- $P(X = 2) = e^{-5} \cdot 5^2 / 2! = 25 \cdot e^{-5} / 2 = 0,0842 = 8,42\%$.
- $P(X \geq 3) = 1 - P(X < 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)] = 0,8753 = 87,5\%$.

Παράδειγμα 2

Έστω ότι η τ.μ. X ακολουθεί την κατανομή Poisson με παράμετρο λ . Αν ισχύει $P(X \leq 1) = 4 \cdot P(X = 2)$, να προσδιοριστεί η παράμετρος λ και να υπολογιστούν οι πιθανότητες $P(X < 3)$, $P(X \geq 1 | X < 3)$.

Λύση

Παράδειγμα 2

Έστω ότι η τ.μ. X ακολουθεί την κατανομή Poisson με παράμετρο λ . Αν ισχύει $P(X \leq 1) = 4 \cdot P(X = 2)$, να προσδιοριστεί η παράμετρος λ και να υπολογιστούν οι πιθανότητες $P(X < 3)$, $P(X \geq 1 | X < 3)$.

Λύση

Είναι $X \sim \text{Poisson}(\lambda)$ άρα $P(X = \kappa) = e^{-\lambda} \cdot \lambda^\kappa / \kappa!$. Τώρα, $P(X \leq 1) = 4 \cdot P(X = 2) \leftrightarrow$

$$\leftrightarrow P(X = 0) + P(X = 1) = 4 \cdot P(X = 2) \leftrightarrow e^{-\lambda} + e^{-\lambda} \cdot \lambda = 4 e^{-\lambda} \cdot \lambda^2 / 2 \leftrightarrow 2\lambda^2 - \lambda - 1 = 0 \leftrightarrow \lambda = -1/2 \text{ ή } \lambda = 1.$$

Καθώς, το λ αποτελεί ρυθμό εμφάνισης γεγονότων, πρέπει να είναι θετικός αριθμός. Άρα $\lambda = 1$.

$$P(X < 3) = P(X=0) + P(X=1) + P(X=2) = e^{-1} \cdot 1^0 / 0! + e^{-1} \cdot 1^1 / 1! + e^{-1} \cdot 1^2 / 2! = e^{-1} \cdot 5 / 2 = 0,9196 \approx 92\%$$

$$P(X \geq 1 | X < 3) = P(X \geq 1, X < 3) / P(X < 3) = P(X = 1 \text{ ή } X = 2) / P(X < 3)$$

$$= [P(X = 1) + P(X = 2)] / P(X < 3) = 3/5 = 0,6 = 60\%.$$

Πιθανογεννήτρια συνάρτηση Poisson

Παράδειγμα

Αν $X \sim \text{Poisson}(\lambda)$ τότε $P(X = \kappa) = \lambda^\kappa \cdot e^{-\lambda} / \kappa!$, $\kappa = 0, 1, 2, \dots$, και

$$\begin{aligned} G_X(z) &= E(z^X) = \sum_{\kappa=0, 1, \dots} P(X = \kappa) \cdot z^\kappa \\ &= \sum_{\kappa=0, 1, \dots} \lambda^\kappa \cdot e^{-\lambda} / \kappa! \cdot z^\kappa \\ &= e^{-\lambda} \cdot \sum_{\kappa=0, 1, \dots} (z\lambda)^\kappa / \kappa! \\ &= e^{-\lambda} \cdot e^{z\lambda} \\ &= e^{(z-1)\lambda}. \end{aligned}$$

Άρα, $G_X(z) = e^{(z-1)\lambda}$ για **κάθε** $z \in \mathbf{C}$, ειδικότερα, $G_X(x) = e^{(x-1)\lambda}$, για **κάθε** $x \in \mathbf{R}$.

Ροπογεννήτρια συνάρτηση Poisson

Άσκηση

Να δείξετε ότι αν $X \sim \text{Poisson}(\lambda)$ τότε $M_X(t) = E(e^{tX}) = e^{\lambda(e^t - 1)}$

Απόδειξη

Άθροισμα ανεξάρτητων Poisson μεταβλητών

Πρόταση

Το άθροισμα δύο **ανεξάρτητων** τυχαίων μεταβλητών $X_1 \sim \text{Poisson}(\lambda_1)$ και $X_2 \sim \text{Poisson}(\lambda_2)$ είναι ομοίως κατανομή Poisson με παράμετρο $\lambda_1 + \lambda_2$, δηλαδή, αν $X \sim X_1 + X_2$, και X_1, X_2 ανεξάρτητες τυχαίες μεταβλητές τότε $X \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

Απόδειξη

Αρκεί να δείξουμε ότι $P(X = z) = e^{-\lambda} \cdot \lambda^z / z!$, για κάθε $z \in \mathbb{N}$ και $\lambda = \lambda_1 + \lambda_2$. Πράγματι, είναι

$$\begin{aligned} P(X = z) &= \sum_{j=0, 1, \dots, z} P(X_1 = j, X_2 = z - j) \\ &= \sum_{j=0, 1, \dots, z} P(X_1 = j) \cdot P(X_2 = z - j) && \text{(οι } X_1, X_2, \text{ είναι ανεξάρτητες τ.μ.)} \\ &= \sum_{j=0, 1, \dots, z} e^{-\lambda_1} \cdot \lambda_1^j / j! \cdot e^{-\lambda_2} \cdot \lambda_2^{z-j} / (z-j)! \\ &= \sum_{j=0, 1, \dots, z} e^{-(\lambda_1 + \lambda_2)} \cdot \lambda_1^j \cdot \lambda_2^{z-j} / [j! \cdot (z-j)!] \\ &= e^{-(\lambda_1 + \lambda_2)} \cdot \sum_{j=0, 1, \dots, z} (z \text{ ανά } j) \cdot \lambda_1^j \cdot \lambda_2^{z-j} / z! \\ &= e^{-(\lambda_1 + \lambda_2)} \cdot (\lambda_1 + \lambda_2)^z \\ &= e^{-\lambda} \cdot \lambda^z / z! \end{aligned}$$

Άθροισμα ανεξάρτητων Poisson μεταβλητών

Πρόταση

Το άθροισμα δύο **ανεξάρτητων** τυχαίων μεταβλητών $X_1 \sim \text{Poisson}(\lambda_1)$ και $X_2 \sim \text{Poisson}(\lambda_2)$ είναι ομοίως κατανομή Poisson με παράμετρο $\lambda_1 + \lambda_2$, δηλαδή, αν $X \sim X_1 + X_2$, και X_1, X_2 ανεξάρτητες τυχαίες μεταβλητές τότε $X \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

Παρατηρήσεις

1. Η υπόθεση της (στοχαστικής) ανεξαρτησίας των X_1, X_2 είναι ουσιαστική. Αν οι X_1, X_2 δεν είναι ανεξάρτητες τότε $\text{Cov}(X_1, X_2) > 0$ και κατά συνέπεια θα είναι

$$\text{Var}(X_1 + X_2) = \text{Var}(X_1) + 2\text{Cov}(X_1, X_2) + \text{Var}(X_2) > \text{Var}(X_1) + \text{Var}(X_2) = E(X_1 + X_2),$$

δηλαδή, η τ.μ. $X_1 + X_2$ δεν μπορεί να ακολουθεί την κατανομή Poisson στην οποία $\mu = \sigma^2$.

2. Η παραπάνω πρόταση δεν αποκλείει το ενδεχόμενο το άθροισμα δύο **εξαρτημένων** Poisson μεταβλητών να είναι Poisson μεταβλητή. Ένα παράδειγμα αναφέρεται στη δημοσίευση [Jacod, J. (1975). Two Dependent Poisson Processes Whose Sum Is Still a Poisson Process. Journal of Applied Probability, 12(1), 170–172. <https://doi.org/10.2307/3212423>]

Ασκήσεις στην κατανομή Poisson

1. (α) Αν $X \sim \text{Poisson}(2)$ να βρεθεί η $P(X = 2)$. (β) Αν $X \sim \text{Poisson}(0,5)$ να βρεθεί η $P(X = 1)$

Λύση

$$(α) X \sim \text{Poisson}(2) \Rightarrow P(X=k) = e^{-2} \cdot \frac{2^k}{k!}$$

$$P(X=2) = e^{-2} \cdot \frac{2^2}{2!} = 2e^{-2}$$

$$(β) X \sim \text{Poisson}(0,5) \Rightarrow P(X=k) = e^{-0,5} \frac{0,5^k}{k!}$$

$$P(X=1) = e^{-0,5} \frac{0,5^1}{1!} = 0,5 \cdot e^{-0,5}$$

$$\binom{10}{4} = \frac{10!}{4!6!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6!}{4 \cdot 3 \cdot 2 \cdot 6!} = 210.$$

Ασκήσεις στην κατανομή Poisson

2. Οι αφίξεις οχημάτων σε ένα πρατήριο καυσίμων αποτελούν διαδικασία Poisson με αναμενόμενο πλήθος 2 οχήματα κάθε 1 λεπτό. Το 80% των οχημάτων είναι Ι.Χ. αυτοκίνητα και το 20% φορτηγά.

α) Ποια η πιθανότητα να μην εμφανιστεί κανένα όχημα στη διάρκεια ενός πεντάλεπτου;

β) Έστω ότι έρχονται 10 οχήματα στο πρατήριο. Ποια η πιθανότητα από τα 10 οχήματα τα 4 να είναι φορτηγά;

γ) Ποια η πιθανότητα σε ένα πεντάλεπτο να έχουμε ακριβώς 4 Ι.Χ. και 2 φορτηγά;

δ) Δεδομένου ότι δεν έχει εμφανιστεί φορτηγό τα τελευταία 5 λεπτά, υπολογίστε την πιθανότητα ότι δεν θα εμφανιστεί κανένα φορτηγό και στα επόμενα 5 λεπτά.

Λύση

$$\lambda = 2 \text{ αφίξεις / λεπτό} = 10 \text{ αφίξεις / 5 λεπτά.}$$

(α) $X = \{ \text{αφίξεις οχημάτων σε 5 λεπτά} \} \sim \text{Poisson}(10)$

$$P(X=0) = e^{-10} \cdot \frac{10^0}{0!} = e^{-10}.$$

$$(β) P(X=6) \cdot P(Y_1=2).$$

$$Y_1 = \{ \text{φορτηγά σε 6 οχήματα} \} \sim B(6, 0.2)$$

$$(β) Y = \{ \text{φορτηγά σε 10 οχήματα} \} \sim B(10, 0.2) \Rightarrow P(Y=4) = \binom{10}{4} 0.2^4 \cdot 0.8^6 = \dots$$

Ασκήσεις στην κατανομή Poisson

3. Το πλήθος σωματιδίων που εκπέμπονται από μία ραδιενεργή πηγή είναι τυχαία μεταβλητή που ακολουθεί την κατανομή Poisson. Αν η πιθανότητα εκπομπής κανενός σωματιδίου στη μονάδα του χρόνου είναι $1/3$ τότε να βρεθεί η πιθανότητα εκπομπής 2 ή περισσότερων σωματιδίων.

Λύση

Ασκήσεις στην κατανομή Poisson

4. Η πιθανότητα ένα άτομο να παρουσιάσει σοβαρές παρενέργειες σε ένα εμβόλιο είναι 0,002. Εάν κάνουν το εμβόλιο 2000 άτομα, ποια είναι η πιθανότητα να έχουν σοβαρές παρενέργειες (α) ακριβώς 2 άτομα (β) τουλάχιστον 3 άτομα.

Λύση

Ασκήσεις στην κατανομή Poisson

5. Γνωρίζουμε ότι σε μία περιοχή υπάρχουν 4 φωλιές πουλιών ανά στρέμμα. Να βρεθεί η πιθανότητα σε ένα χωράφι 5 στρεμμάτων να υπάρχουν 8 φωλιές πουλιών.

Λύση

$$\lambda = 4 \text{ φωλιές / στρέμμα} = 20 \text{ φωλιές / 5 στρέμματα}$$

$$X = \{\text{αριθμός φωλιών στα 5 στρέμματα}\} \sim \text{Poisson}(20)$$

$$P(X=8) = e^{-20} \frac{20^8}{8!} = \dots$$

Ασκήσεις στην κατανομή Poisson

6. Σε ένα βιβλίο 400 σελίδων υπάρχουν 200 τυπογραφικά λάθη.

(α) Ποιο είναι το μέσο πλήθος λαθών ανά σελίδα;

(β) Ποια είναι η πιθανότητα σε μία σελίδα να μην υπάρχουν λάθη;

(γ) Ποιο είναι το πλήθος σελίδων του βιβλίου που περιμένουμε να μην έχει τυπογραφικά λάθη;

Λύση

Ασκήσεις στην κατανομή Poisson

7. Μία φαρμακευτική αποθήκη εξυπηρετεί 10 φαρμακεία στο νομό της Ξάνθης. Το μέσο πλήθος των συσκευασιών ενός φαρμάκου που πωλείται από κάθε ένα από τα φαρμακεία δίνεται στον παρακάτω πίνακα:

Φαρμακείο	1	2	3	4	5	6	7	8	9	10
Πλήθος	2,1	3,5	7	1	0,4	5	2,6	1	3	4,4

Οι πωλήσεις κάθε ενός φαρμακείου είναι ανεξάρτητες. Ποια είναι η πιθανότητα, μία ημέρα να ζητηθούν από την αποθήκη 40 συσκευασίες από το φάρμακο;

Λύση