ΠΙΘΑΝΟΤΗΤΕΣ Ι. ΙΑΝΟΥΑΡΙΟΣ 2010

ΛΥΣΕΙΣ ΘΕΜΑΤΩΝ. ΟΜΑΔΑ Α

Θ1. Η ζητούμενη πιθανότητα ισούται με

$$\mathbf{P}(\text{όλα τα σφαιρίδια είναι κόκκινα}) + \mathbf{P}(\text{όλα τα σφαιρίδια είναι μπλέ}) = \frac{(6)_4}{(14)_4} + \frac{(8)_4}{(14)_4}.$$

Σχόλιο: Εδώ πήραμε ως δειγματικό χώρο, έστω Ω_1, το σύνολο των διατεταγμένων τετράδων απο το σύνολο S των σφαιριδίων, και αυτός είναι ο πραγματικός δειγματικός χώρος γιατί τα σφαιρίδια εξάγονται διαδοχικά, και εκ των πραγμάτων είναι διαφορετικά φυσικά αντικείμενα. Εναλλακτικά, θα μπορούσαμε να πάρουμε ως δειγματικό χώρο Ω_2 το σύνολο όλων των υποσυνόλων του S με τέσσερα στοιχεία. Και αυτό είναι αποδεκτό γιατί απο τη φύση του πειράματος όλα τα στοιχεία του Ω_2 είναι ισοπίθανα. Με αυτή τη θεώρηση, στον υπολογισμό των πιθανοτήτων χρησιμοποιούμε συνδυασμούς αντί διατάξεων ή χρησιμοποιούμε την υπεργεωμετρική κατανομή.

Θ2. Θεωρούμε τα ενδεχόμενα

$$A_k : \text{επιλέγουμε κίβδηλο νόμισμα,}$$
$$A_a : \text{επιλέγουμε αμερόληπτο νόμισμα,}$$

και B το ενδεχόμενο το επιλεγμένο νόμισμα να δίνει στις τρείς ρίψεις "κεφαλή". Η ζητούμενη πιθανότητα είναι

$$\mathbf{P}(A_k \mid B) = \frac{\mathbf{P}(B \cap A_k)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B \cap A_k)}{\mathbf{P}(B \cap A_k) + \mathbf{P}(B \cap A_a)} = \frac{\mathbf{P}(B \mid A_k)\,\mathbf{P}(A_k)}{\mathbf{P}(B \mid A_k)\,\mathbf{P}(A_k) + \mathbf{P}(B \mid A_a)\,\mathbf{P}(A_a)}$$
$$= \frac{\left(\frac{4}{5}\right)^3 \frac{1}{5}}{\left(\frac{4}{5}\right)^3 \frac{1}{5} + \left(\frac{1}{2}\right)^3 \frac{4}{5}} = \frac{\frac{64}{125}}{\frac{64}{125} + \frac{1}{2}} = \frac{128}{253}.$$

Θ3. (**i**) Πρέπει το ολοκλήρωμα της f στο $\mathbb{R}$ να ισούται με 1. Έχουμε

$$\int_{-\infty}^{\infty} f(x)\,dx = c\int_0^3 (3-x)\,dx = c\left(9 - \frac{9}{2}\right) = 9c/2.$$

Άρα $c = 2/9$.

(**ii**) Προφανώς $F(x) = 0$ για $x \le 0$ και $F(x) = 1$ για $x \ge 3$, ενώ για $x \in (0,3)$ έχουμε

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \frac{2}{9}\int_0^x (3-t)\,dt = \frac{2x}{3} - \frac{x^2}{9}.$$

(**iii**) Οι δύο πρώτες ροπές της X είναι

$$\mathbf{E}(X) = \int_{-\infty}^{\infty} x f(x) = \frac{2}{9}\int_0^3 (3x - x^2)\,dx = ... = 1,$$
$$\mathbf{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x) = \frac{2}{9}\int_0^3 (3x^2 - x^3)\,dx = ... = 3/2.$$

Άρα η μέση τιμή είναι 1 και η διασπορά

$$V(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = 1/2.$$

(**iv**) Βρίσκουμε τη συνάρτηση κατανομής F_Y της Y.

$$F_Y(y) = \mathbf{P}(\log X \le y) = \mathbf{P}(X \le e^y) = F(e^y).$$

Απο τη μορφή της F έπεται ότι η F_Y είναι συνεχής παντού και παραγωγίσιμη με συνεχή παράγωγο στο συμπλήρωμα ενός πεπερασμένου συνόλου (στο $\mathbb{R}\setminus\{\log 3\}$), και άρα η Y είναι συνεχής τυχαία μεταβλητή με πυκνότητα

$$f_Y(y) := F_Y'(y) = \begin{cases} f(e^y)e^y = \frac{2}{9}(3-e^y)e^y & \text{αν } y \in (-\infty, \log 3), \\ 0 & \text{αν } y \in (\log 3, +\infty). \end{cases}$$

Για την $f_Y(\log 3)$ μπορούμε να επιλέξουμε οποιαδήποτε τιμή στο $[0,\infty)$, π.χ., $f_Y(\log 3) = 0$.

Θ4. Έστω X_A, X_B οι αριθμοί των εξαρτημάτων τύπου A και B αντίστοιχα που λειτουργούν. Απο τα δεδομένα του προβλήματος έπεται ότι η X_A ακολουθεί τη διωνυμική κατανομή με παραμέτρους n_1, p_1, ενώ η X_B ακολουθεί τη διωνυμική κατανομή με παραμέτρους n_2, p_2. Η ζητούμενη πιθανότητα είναι

$$\begin{aligned}\mathbf{P}(X_A \ge 2 \text{ και } X_B \ge 4) &= \mathbf{P}(X_A \ge 2)\,\mathbf{P}(X_B \ge 4) \\ &= \left(\sum_{k=2}^{n_1}\binom{n_1}{k}p_1^k(1-p_1)^{n_1-k}\right)\left(\sum_{k=4}^{n_2}\binom{n_2}{k}p_2^k(1-p_2)^{n_2-k}\right).\end{aligned}$$

Χρησημοποιήσαμε την ανεξαρτησία των X_A, X_B.

Θ5. [Όπως και στο Θέμα 2, έχουμε ένα πείραμα που γίνεται σε δύο βήματα, και η έκβαση του πρώτου επηρεάζει αυτήν του δεύτερου. Επομένως, πρέπει να διακρίνουμε περιπτώσεις για το τι συμβαίνει στο πρώτο βήμα.]

Προφανώς η N παίρνει τιμές στο $\mathbb{N}\setminus\{0\}$, και για k σε αυτό το σύνολο, $\mathbf{P}(N=k) = (1-p_1)^{k-1}p_1$. Δηλαδη η N ακολουθεί τη μετατοπισμένη γεωμετρική με παράμετρο p_1.

Μετά την πραγματοποίηση του πρώτου πειράματος, το N έχει πάρει μια τιμή, έστω k. Δεδομένου αυτού, η X ακολουθεί τη διωνυμική κατανομή με παραμέτρους k, p_2. Άρα

$$\begin{aligned}\mathbf{P}(X=1) &= \sum_{k=1}^{+\infty}\mathbf{P}(X=1 \text{ και } N=k) = \sum_{k=1}^{+\infty}\mathbf{P}(X=1\,|\,N=k)\,\mathbf{P}(N=k) \\ &= \sum_{k=1}^{+\infty}\binom{k}{1}p_2(1-p_2)^{k-1}\,(1-p_1)^{k-1}p_1 = p_1p_2\sum_{k=0}^{+\infty}k((1-p_1)(1-p_2))^k.\end{aligned}$$

Όμως για $x \in (-1,1)$ ισχύει

$$\sum_{k=0}^{+\infty}kx^k = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}.$$

Άρα

$$\mathbf{P}(X=1) = \frac{p_1p_2}{(p_1+p_2-p_1p_2)^2}.$$

Θ6. **(α)** Έχουμε

$$\begin{aligned}\rho(X+Y, X-Y) :&= \frac{\operatorname{Cov}(X+Y, X-Y)}{\sqrt{V(X+Y)V(X-Y)}} = \frac{\operatorname{Cov}(X,X)-\operatorname{Cov}(X,Y)+\operatorname{Cov}(Y,X)-\operatorname{Cov}(Y,Y)}{\sqrt{(V(X)+V(Y))(V(X)+V(-Y))}} \\ &= \frac{V(X)-V(Y)}{V(X)+V(Y)}.\end{aligned}$$

Στη δεύτερη ισότητα χρησιμοποιήσαμε τη διγραμμικότητα της Cov και την ανεξαρτησία των X, Y, ενώ στην τρίτη το ότι $\mathrm{Cov}(X,Y) = 0$ (επειδή οι X, Y είναι ανεξάρτητες), $\mathrm{Cov}(X,X) = V(X), \mathrm{Cov}(Y,Y) = V(Y), V(-Y) = V(Y)$.

(β) Η X έχει διασπορά $V(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = 1$, και όμοια, $V(Y) = 1$. Άν οι X, Y ήταν ανεξάρτητες, τότε $\mathrm{Cov}(X,Y) = 0$ και $V(X+Y) = V(X) + V(Y) + 2\mathrm{Cov}(X,Y) = 1 + 1 + 0 = 2$, ενώ μας έχει δοθεί ότι $V(X+Y) = 3$. Άρα δεν είναι ανεξάρτητες.

Θ7. (**i**) Πρέπει $\int_{\mathbb{R}^2} f(x,y)\,dx\,dy = 1$. Το ολοκλήρωμα ισούται με

$$\int_0^1 \int_{x^3}^{x} c\,dy\,dx = c\int_0^1 (x - x^3)\,dx = c/4.$$

Άρα $c = 4$.

(**ii**) Η συνάρτηση πυκνότητας-πιθανότητας της X είναι

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \mathbf{1}_{x\in(0,1)} \int_{x^3}^{x} 4\,dy = 4\,(x - x^3)\,\mathbf{1}_{x\in(0,1)}.$$

Για τη συνάρτηση πυκνότητας-πιθανότητας της Y παρατηρούμε ότι για $y \in (0,1)$ σταθερό, $f(x,y) \neq 0$ αρκιβώς όταν $y < x < y^{1/3}$, και μάλιστα τότε η f ισούται με 4. Οπότε

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx = \mathbf{1}_{y\in(0,1)} \int_{y}^{y^{1/3}} 4\,dx = 4\,(y^{1/3} - y)\,\mathbf{1}_{y\in(0,1)}.$$

Θ8. **(α)** Έστω X μια τυχαία μεταβλητή με κατανομή Poisson με παράμετρο λ. Για $t \in \mathbb{R}$ έχουμε

$$\Phi_X(t) = \mathbf{E}(t^X) = \sum_{k=0}^{+\infty} \mathbf{P}(X = k)\,t^k = \sum_{k=0}^{+\infty} e^{-\lambda}\frac{\lambda^k}{k!}\,t^k = e^{-\lambda}\sum_{k=0}^{+\infty}\frac{(\lambda t)^k}{k!} = e^{\lambda(t-1)}.$$

(β) Η πιθανογεννήτρια μιας κατανομής με τιμές στο $\mathbb{N}$ είναι πάντοτε πεπερασμένη στο $[-1,1]$ και χαρακτηρίζει την κατανομή. Το άθροισμα $X_1 \cdots + X_\nu$ έχει για κάθε $t \in \mathbb{R}$ πιθανογεννήτρια

$$\Phi_{X_1+\cdots+X_\nu}(t) := \mathbf{E}(t^{X_1}\cdots t^{X_\nu}) = \Phi_{X_1}(t)\cdots\Phi_{X_\nu}(t) = e^{(\lambda_1+\cdots+\lambda_\nu)(t-1)}, \tag{1}$$

όπου στη δεύτερη ισότητα χρησιμοποιήσαμε την ανεξαρτησία των $X_1, \ldots, X_\nu$, και στην τρίτη το ερώτημα (α). Όμως, πάλι απο το (α), η (**;;**) είναι η πιθανογεννήτρια της κατανομής Poisson με παράμετρο $\lambda_1 + \cdots + \lambda_\nu$. Έπεται ότι η $X_1 + \cdots + X_\nu$ ακολουθεί την κατανομή Poisson με παράμετρο $\lambda_1 + \cdots + \lambda_\nu$.

Θ9. Έστω $(X_i)_{i\geq 1}$ μια ακολουθία ανεξάρτητων ισόνομων τυχαίων μεταβλητών με κατανομή ομοιόμορφη στο $\{1,2,3,4\}$, και $S_n := X_1 + \cdots + X_n$ για κάθε $n \geq 1$. Οι δύο πρώτες ροπές της X_1 είναι

$$\mathbf{E}(X_1) = \frac{1}{4}(1+2+3+4) = \frac{5}{2},$$
$$\mathbf{E}(X_1^2) = \frac{1}{4}(1^2+2^2+3^2+4^2) = \frac{15}{2}.$$

Άρα η X_1 έχει μέση τιμή $5/2$ και πεπερασμένη διασπορά $\sigma^2 = \mathbf{E}(X_1^2) - (\mathbf{E}(X_1))^2 = 5/4$. Το κεντρικό οριακό θεώρημα λέει οτι η ακολουθία των τυχαίων μεταβλητών

$$\left(\frac{S_n - 5n/2}{\sqrt{5/4}\sqrt{n}}\right)_{n\geq 1}$$

συγκλίνει κατα κατανομή στην τυπική κανονική κατανομή, $N(0,1)$.

Για $n = 80$, έχουμε $5n/2 = 200$ και $\sqrt{(5/4)\,n} = 10$, οπότε

$$\mathbf{P}(190 \le S_{80} \le 220) = \mathbf{P}\left(-1 \le \frac{S_{80} - 200}{10} \le 2\right) \approx \Phi(2) - \Phi(-1).$$

Η προσέγγιση προκύπτει απο το κεντρικό οριακό θεώρημα. Τέλος, επειδή η $N(0,1)$ είναι συνεχής τυχαία μεταβλητή με πυκνότητα άρτια συνάρτηση, έχουμε $\Phi(-1) + \Phi(1) = 1$. Άρα η ζητούμενη προσέγγιση είναι $\Phi(2) + \Phi(1) - 1 \approx 0.8186$.