A szeptember 21-i gyakorlat témája
Rövid összefoglaló

1. Adott egy urna, és abban 70 piros és 30 fehér golyó. Egymás után húzunk ebből az urnából 100 alkalommal. Minden húzás után a kihúzott golyót visszadobjuk egy másik azonos színű golyóval együtt. Arra vagyunk kivácsiak, hogy mi a valószínűsége annak, hogy a 30. húzás eredménye piros.

Készítsük el ennek a feladatnak egy valószínűségi modelljét. E feladat modelljét, illetve néhány ezzel kapcsolatos alapvető fogalmat részletesen megtárgyaltunk. Bevezettük ebben a modellben a ξ_j, $j = 1, \dots, 100$, valószínűségi változókat, melyekre $\xi_j = 1$, ha a j-ik húzás eredménye piros, és $\xi_j = 0$, ha a j-ik húzás eredménye fehér. Beláttuk, hogy a fenti jelölésekkel

$$P(\xi_1 = \varepsilon_1, \xi_2 = \varepsilon_2, \dots, \xi_{100} = \varepsilon_{100}) = P(\xi_{\pi(1)} = \varepsilon_1, \xi_{\pi(2)} = \varepsilon_2, \dots, \xi_{\pi(100)} = \varepsilon_{100}),$$

ahol $(\pi(1), \dots \pi(100))$ az $(1, \dots, 100)$ számok tetszőleges permutációja, ε_j, $1 \le j \le 100$ vagy nullla vagy egy. Ez az azonosság lehetővé tette a feladat egyszerű megoldását. Ugyanis, mint azt megmutattuk, e tényből következik, hogy a 30. és 1. húzás eredménye ugyanolyan valószínűséggel lesz piros. Ennek a feladat megoldása lehet az alapja az 1. kitűzött feladat megoldásának.

2. Egy estélyen megjelenik 100 házaspár. Egy táncmester, aki nem tudja, hogy kik házastársak és kik nem véletlen módon párba rendezi a férfiakat és nőket a tánc előtt. Mi a valószínűsége annak, hogy egyetlen házaspár sem táncol együtt? Mi ez a valószínűség n házaspár esetén, és mi ennek a valószínűségnek a határértéke, ha $n \to \infty$.

Ennek a feladatnak célja megmutatni, hogy az események halmazokkal való reprezentálása, az eseményalgebrával való számolás, illetve bizonyos alapvető valószínűségi (valójában leszámolással kapcsolatos kombinatorikai) azonosságok segítséget jelentenek bizonyos feladatok megoldásában.

Tekintsünk megfelelő valószínűségi mezőt, melyen ez a feladat megfogalmazható, és defini'aljuk a következő A_j eseményeket:

$$A_j = \text{a } j\text{-ik házaspár egütt táncol.} \quad 1 \le j \le 100.$$

Ekkor minket a $P(\overline{A_1 + \cdots + A_{100}}) = 1 - P(A_1 + \cdots + A_{100})$ valószínűség érdekel. Vegyük észre, hogy a $P(A_{j_1} \cdots A_{j_k}) = \dfrac{(100-k)!}{100!}$ azonosság igaz. Továbbá érvényes a következő az irodalomban szita-formulának nevezett eredmény:

$$P(A_1 + \cdots + A_n) = S_1 - S_2 + S_3 - \cdots + (-1)^{n+1} S_n,$$

ahol

$$S_k = \sum_{1 \le j_1 < \cdots < j_k \le n} P(A_{j_1} \cdots A_{j_k}).$$

Továbbá az S_1, $S_1 - S_2$, $S_1 - S_2 + S_3$, előjeles összegek váltakozva felső és alsó becslést adnak a $P(A_1 + \cdots + A_n)$ valószínűségre.

A szita-formula segítségével kapjuk, hogy n házaspár esetében

$$S_k = \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!}.$$

Innen adódik, hogy a minket érdeklő valószínűség

$$\begin{aligned} P(\overline{A_1 + \cdots + A_n}) &= 1 - P(A_1 + \cdots + A_n) \\ &= 1 - S_1 + S_2 + \cdots + (-1)^n S_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \quad \to \frac{1}{e} \text{ ha } n \to \infty. \end{aligned}$$

Tegyünk néhány megjegyzést a szita-formulával kapcsolatban:

Adva egy A esemény vezessük be az A^ε jelölést, ahol $\varepsilon = \pm 1$, $\varepsilon = 1$ estében $A^1 = A$, $\varepsilon = -1$ esetében $A^{-1} = \Omega \setminus A = \bar{A}$. Definiáljuk tetszőleges $j_s = 0$ vagy $j_s = 1$, $s = 1, \dots, n$ számokra a

$$A(j_1, \dots, j_n) = A^{-1^{j_1}} \cdots A^{-1^{j_n}}$$

halmazt. Vegyük észre, hogy

$$A_1 + \cdots + A_n = \sum_{(j_1,\dots,j_n)\colon\ (j_1,\dots,j_n) \neq (0,\dots,0)} A(j_1, \dots, j_n),$$

és

$$A_k = \sum_{(j_1,\dots,j_n)\colon\ j_k \neq 0} A(j_1, \dots, j_n), \quad 1 \le k \le n.$$

Továbbá az $A(j_1, \dots, j_n)$ események különböző $(j_1, \dots, j_n)$ paraméterek diszjunktak. Ilyen módon a szita-formula két oldalán szereplő kifejezés kifejezhető a $P(A(j_1, \dots, j_n))$ valószínűségek lineáris kombinációjának segítségével. A bizonyítás lényege az, hogy a kifejezés két oldalán ugyanaz a lineáris kombináció jelenik meg. A szita-formula elnevezés onnan származik, hogy egy rögzített $A(j_1, \dots, j_n)$ esemény $P(A(j_1, \dots, j_n))$ valószínűsége különböző számossággal és előjellel jelenik meg az $S_1, S_2, \dots$ kifejezésekben. Ezek összeszámlása "kiszitálása" szolgáltatja a szita-formula bizonyítását.

Az előadás végén felidéztük a függetlenség fogalmát. Mivel ez a fogalom alapvető fontosságú a valószínűségszámításban, erről egy részletesebb kiegészítést írtam, és e fogalom megbeszélését folytattuk a következő alkalommal is.