A Valószínűségszámítás I. előadássorozat negyedik előadása.
2007. február 27.

Először mutatok egy olyan feladatot, amelyet viszonylag könnyen meg tudunk oldani az előző előadáson megfogalmazott valószínűségszámítási szita formula segítségével.

Egy estélyen megjelenik n 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 ennek a valószínűségnek a határértéke, ha $n \to \infty$?

Megoldás: Definiáljuk a következő A_j eseményeket:

$$A_j = \text{a } j\text{-ik házaspár együtt táncol}, \quad 1 \le j \le n.$$

Ekkor minket a $P(\overline{A_1 \cup \dots \cup A_n}) = 1 - P(A_1 \cup \dots \cup A_n)$ valószínűség érdekel. Számoljuk ki a $P(A_1 \cup \dots \cup A_n)$ valószínűséget a valószínűségszámítási szita formula segítségével. Ennek érdekében vegyük észre, hogy a $P(A_{j_1} \cap \dots \cap A_{j_k}) = \frac{(n-k)!}{n!}$ azonosság érvényes minden lehetséges $1 \le j_1 < \dots < j_k \le n$ szám-k-asra. Ugyanis az összes lehetséges párbaállítások száma $n!$, míg az olyan párbaállítások száma, amelyben a j_1-ik, j_2-ik, ..., j_k-ik házaspár egy párba kerül $(n-k)!$. Ezért a valószínűségszámítási szita formulában bevezetett S_k mennyiség értéke $S_k = \binom{n}{k}\frac{(n-k)!}{n!} = \frac{1}{k!}$ minden $1 \le k \le n$ számra.
Innen adódik, hogy a minket érdeklő valószínűség n házaspár esetén

$$\begin{aligned} P(\overline{A_1 \cup \dots \cup A_n}) &= 1 - P(A_1 \cup \dots \cup A_n) \\ &= 1 - S_1 + S_2 + \dots + (-1)^n S_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!}, \end{aligned}$$

ahonnan

$$P(\overline{A_1 \cup \dots \cup A_n}) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \to \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = \frac{1}{e} \qquad \text{ha } n \to \infty.$$

Tehát nagy n számra annak a valószínűsége, hogy egy házaspár sem fog együtt táncolni közelítőleg $\frac{1}{e}$.

Egy másik feladat, ahol a szita formula jól alkalmazható.

Feladat:

Egy cornflake gyártó cég minden dobozba betesz egy kupont, és összesen 10 különböző kupont használ. Mekkora annak a valószínűsége, hogy mind a 10 kupont megkapja egy olyan vásárló, aki 25 doboz cornflake-et vesz?

Megoldás: Jelölje A_j, $1 \le j \le 10$ azt az eseményt, hogy a j-ik kupont megkapta a vásárló, és $\bar{A}_j$ ennek az eseménynek a komplementerét. Ekkor a $P\left(\bigcap_{j=1}^{10} A_j\right)$

valószínűséget kell kiszámolnunk. Viszont $P\left(\bigcap_{j=1}^{10} A_j\right) = 1 - P\left(\bigcup_{j=1}^{10} \bar{A}_j\right)$, és $P\left(\bigcup_{j=1}^{10} \bar{A}_j\right) = \sum_{k=1}^{10} (-1)^k S_k$, ahol $S_k = \sum P(\bar{A}_{j_1} \cap \bar{A}_{j_2} \cap \cdots \cap \bar{A}_{j_k})$, és a $\{j_1, \dots, j_k\}$ indexhalmaz a fenti szummában az $\{1, \dots, 10\}$ halmaz összes k elemű részhalmazából áll.

Viszont $P(\bar{A}_{j_1} \cap \bar{A}_{j_2} \cap \cdots \cap \bar{A}_{j_k}) = \left(\frac{10-k}{10}\right)^{25}$, annak a valószínűsége, hogy a lehetséges 10 kuponból mind a 25 vásárlásnál a $j_1, \dots\ j_k$ indexű kuponoktól különböző $10-k$ kupon valamelyikét kapjuk. Ezért $S_k = \binom{10}{k}\left(\frac{10-k}{10}\right)^{25}$ minden $1 \le k \le 10$ indexre $k = 10$-re $S_k = S_{10} = 0$, ahonnan $P\left(\bigcap_{j=1}^{10} A_j\right) = \sum_{k=0}^{9} (-1)^k \binom{10}{k}\left(1 - \frac{k}{10}\right)^{25}$.

A valószínűségszámítási szita formula bizonyítását a kiegészítésben fogom tárgyalni. Itt azt mutatom meg, hogy a kombinatorikus szita formula egyszerűen levezethető a valószínűségszámítási szita formulából.

Ennek érdekében tekintsünk bizonyos $A_1, \dots, A_n$ véges halmazokat, és jelöljük e halmazok unióját $A = A_\cup \cdots \cup A_n$-nel. Legyen $A = \{x_1, \dots, x_N\}$ egy N elemű halmaz, és vezessük be azt a valószínűségi mezőt, amelyben az A halmaz elemeit választhatjuk ki egyenletes eloszlással. Részletesebben kifejtve, a következő $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt definiáljuk: $\Omega = A$, $\mathcal{A}$ az A halmaz összes részhalmazából álló σ-algebra, $P(\{x_j\}) = \frac{1}{N}$ minden $1 \le j \le N$ számra, ahonnan $P(B) = \frac{|B|}{N}$ minden $B \subset A$ halmazra. Megmutatom, hogy a kombinatorikus szita formulát megkaphatjuk a valószínűségi szita formula segítségével, ha azt a most definiált $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn az $A_1, \dots, A_n$ halmazokra alkalmazzuk.

Jelöljük a kombinatorikus szita formulában definiált S_j, $1 \le j \le n$, összegek megfelelőit a valószínűségi szita formulában $\bar{S}_j$-vel. Ekkor nyílván $S_j = N\bar{S}_j$. Továbbá, mivel $P(A) = 1$, a valószínűségi szita formula azt adja, hogy $S_1 - S_2 + S_3 - \cdots + (-1)^{n+1} S_n = N(\bar{S}_1 - \bar{S}_2 + \bar{S}_3 - \cdots + (-1)^{n+1}\bar{S}_n) = NP(A) = N$, és mivel $|A| = N$, ezt kellett bizonyítanunk. A kombinatorikus szita formulában megfogalmazott egyenlőtlenségek hasonlóan vezethetők le ezen állítások megfelelőiből a valószínűségszámítási szita formulában.

Megfordítva, a valószínűségszámítási szita formula is egyszerűen levezethető a kombinatorikus szita formulából. De ehhez szükségünk van a valószínűségi változók és azok várható értékének a fogalmára és néhány e fogalmakkal kapcsolatos fontos eredményre. Ezért e levezetés ismertetését az előadás végére halasztom.

Következő témánk a valószínűségszámítás egyik fontos eredménye, a Borel–Cantelli lemma. Először informálisan ismertetem, hogy milyen probléma vizsgálatában jelent meg ez az eredmény.

A kérdés a következő: Mikor mondhatjuk azt, hogy bizonyos események közül végtelen sok bekövetkezik: a.) egy valószínűséggel, b.) pozitív valószínűséggel? Például mikor mondhatjuk egy (vagy pozitív) valószínűséggel azt, hogy végtelen sok olyan nap van, amelyikben valami jó következik? Ez azt jelenti, hogy bármely nap bízhatunk abban, hogy nem múlt el az összes olyan nap, amelyben valami jó történik, érdemes még élni.

A Borel–Cantelli lemmát megfogalmazásának és bizonyításának érdekében érdemes a halmazelmélet nyelvén megfogalmazni azt a tényt, hogy bízonyos események közül végtelen sok következik be. Ez azt jelenti, hogy ha adva van végtelen sok $A_1, A_2, \dots$ esemény (halmaz) egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, akkor azt az eseményt, hogy ezek közül végtelen sok következik be definiálni kívánjuk az $A_1, A_2, \dots$ események halmazelméleti függvényeként, azaz (megszámlálható) unió, metszet és komplementerképzés segítségével.

Ezt a feladatot megoldottuk korábban, lásd a 2. előadás 7. feladatát. Eszerint a keresett esemény $\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k$ alakban adható meg, és ennek a valószínűségére vagyunk kiváncsiak. Erről a kérdésről ad információt a Borel–Cantelli lemma.

Borel–Cantelli lemma. *Legyen adva végtelen sok $A_1, A_2, \cdots$ esemény egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. A követkaző két állítás igaz:*

a.) Ha $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$, akkor

$$P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = 0,$$

azaz ebben az esetben egy valószínűséggel csak véges sok A_n esemény következik be.

b.) Ha $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$, és az A_n, $n = 1, 2, \dots,$ események függetlenek, akkor

$$P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = 1,$$

azaz ebben az esetben egy valószínűséggel végtelen sok sok A_n esemény következik be.

Teszek néhány megjegyzést ezzel az eredménnyel kapcsolatban. Ha a $P(A_n)$ valószínűségek viszonylag kicsik, ami jelen esetben azt jelenti, hogy az összegük konvergens, akkor csak véges sok A_n esemény következik be 1 valószínűséggel. A másik irányú állításban nemcsak azt követeltük meg, hogy az események valószínűségei legyenek viszonylag nagyok, összegük legyen divergens, hanem azt is, hogy az egyes A_n események legyenek függetlenek. Ebben az esetben viszont erősebb állítást fogalmaztam meg. Nemcsak azt állítottam, hogy ebben az esetben annak a valószínűsége hogy végtelen sok A_n esemény következik be pozitív, hanem azt is, hogy ez a valószínűség 1. Ha tehát az A_n események függetlenek, akkor csak két lehetőség fordulhat elő. Vagy nulla valószínűséggel következik be végtelen sok A_n esemény (ha a valószínűségek összege konvergens) vagy pedig egy valószínűséggel (ha a valószínűségek összege divergens). Közbülső lehetőség nincs.

A b) esetben megfogalmazott eredményben az A_n események függetlensége nagyon fontos feltétel. Ezt a feltételt lehet gyengíteni, de teljesen elhagyni nem lehet. Ezt mutatja az alábbi két példa:

Legyen az $(\Omega, \mathcal{A}, P)$ valószínűségi mező a $[0,1]$ intervallum, rajta a Borel mérhető halmazok σ-algebrája, és a P valószínűségi mérték a Lebesque mérték.

1. példa. Legyen $A_n = \left[0, \frac{1}{2}\right]$ minden $n = 1, 2, \dots,$ számra. Ekkor $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$, $\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k = \left[0, \frac{1}{2}\right]$, ezért $P\left(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k\right) = \frac{1}{2}$.

2. példa. Legyen $A_n = \left(0, \frac{1}{n}\right)$, $n = 1, 2, \dots$. Ekkor $\sum\limits_{n=1}^{\infty} P(A_n) = \sum\limits_{j=1}^{\infty} \frac{1}{n} = \infty$, $\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k = \emptyset$, ezért $P\left(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k\right) = 0$.

Az első példában egy olyan esetet láttunk, amelyben annak a valószínűsége, hogy végtelen sok A_n következik be, $\frac{1}{2}$, tehát sem nem nulla sem nem 1. A második példában egy valószínűséggel csak véges sok A_n esemény következett be, noha a $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$ reláció teljesült. Természetesen egyik példában sem voltak a tekintett A_n események függetlenek. Ezek a példák mutatják, hogy a Borel–Cantelli tétel b) részében szereplő függetlenség feltétel lényeges. Az gyengíthető ugyan, de teljesen el nem hagyható.

Rátérek a Borel–Cantelli lemma bizonyítására. A bizonyításhoz szükségünk van a következő lemmára, amelynek állítását tartalmazza a 2. előadás 14. oldalán megfogalmazott Tétel A.

Lemma a valószínűségi mérték folytonosságáról. *Legyen adva $B_1 \subset B_2 \subset \cdots$, $B_n \in \mathcal{A}$, $n = 1, 2, \dots,$ eseményeknek növekvő sorozata egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Legyen $B = \bigcup\limits_{n=1}^{\infty} B_n$. Ekkor létezik a $\lim\limits_{n\to\infty} P(B_n)$ határérték, és $\lim\limits_{n\to\infty} P(B_n) = P(B)$.*

Legyen adva $C_1 \supset C_2 \supset \cdots$, $C_n \in \mathcal{A}$. $n = 1, 2, \dots,$ eseményeknek csökkenő sorozata egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Legyen $C = \bigcap\limits_{n=1}^{\infty} C_n$. Ekkor létezik a $\lim\limits_{n\to\infty} P(C_n)$ határérték, és $\lim\limits_{n\to\infty} P(C_n) = P(C)$.

A Borel–Cantelli lemma bizonyítása. Az a.) rész bizonyítása:

$$P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) \le P\left(\bigcup_{k=n}^{\infty} A_k\right) \le \sum_{k=n}^{\infty} P\left(A_k\right) \quad \text{minden } n \text{ számra.}$$

Mivel $\sum\limits_{k=1}^{\infty} P\left(A_k\right) < \infty$, minden $\varepsilon > 0$ számhoz létezik olyan $n = n(\varepsilon)$ szám, hogy $\sum\limits_{k=n}^{\infty} P\left(A_k\right) < \varepsilon$, ahonnan az előző reláció szerint $P\left(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k\right) \le \varepsilon$. Innen következik az a.) rész állítása.

A b.) rész bizonyítása: Az előbbi lemma szerint a valószínűségi mérték folytonosságáról kapjuk, hogy az A_n halmazok függetlensége miatt

$$\begin{aligned} P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\right) &= \lim_{N\to\infty} P\left(\bigcup_{k=N}^{\infty} A_k\right) = \lim_{N\to\infty}\left[\lim_{M\to\infty} P\left(\bigcup_{k=N}^{M} A_k\right)\right] \\ &= \lim_{N\to\infty}\left[\lim_{M\to\infty}\left(1-P\left(\bigcap_{k=N}^{M}(\Omega\setminus A_k)\right)\right)\right] \\ &= \lim_{N\to\infty}\left[\lim_{M\to\infty}\left(1-\prod_{k=N}^{M}(1-P(A_k))\right)\right]. \end{aligned}$$

Ezért elég belátni, hogy minden rögzített N számra $\lim\limits_{M\to\infty}\prod\limits_{k=N}^{M}(1-P(A_k))=0$, ha $\sum\limits_{n=1}^{\infty}P(A_n)=\infty$. Viszont az $1-x\le e^{-x}$ egyenlőtlenség alapján $\prod\limits_{k=N}^{M}(1-P(A_k))\le \exp\left\{-\sum\limits_{k=N}^{M}P(A_k)\right\}$, ahonnan $\lim\limits_{M\to\infty}\prod\limits_{k=N}^{M}(1-P(A_k))=0$, ha $\sum\limits_{n=1}^{\infty}P(A_n)=\infty$.

A felhasznált $1-x\le e^{-x}$ egyenlőtlenség például a következő módon látható: Az $f(x)=e^x$ függvény konvex, $f(0)=1$, $f'(0)=1$. Az $f(x)=e^x$ függvény konvexitása miatt e függvény $(0,f(0))$ ponton átmenő érintője minden valós x-re a az $f(x)$ függvény alatt van, azaz $e^x\ge x+1$. Az x szám helyett $-x$-et írva megkapjuk a kívánt állítást.

Megjegyzés. Az analízisben bizonyítják és használják a következő eredményt: Adva x_n valós számoknak olyan sorozata, amelyre $0\le x_n<1$ minden $n=1,2,\dots,$ számra $\prod\limits_{k=1}^{\infty}(1-x_k)>0$, ha $\sum\limits_{k=1}^{\infty}x_k<\infty$, és $\prod\limits_{k=1}^{\infty}(1-x_k)=0$, ha $\sum\limits_{k=1}^{\infty}x_k=\infty$. Bár nekünk erre nem lesz szükségünk, az előadás kiegészítésében megadom ennek a formulának előbb heurisztikus indoklását majd precíz bizonyítását, mivel ez tanulságos.

Diszkrét eloszlású valószínűségi változók és azok függetlensége

Először bevezettem a valószínűségi változó fogalmát.

Valószínűségi változó fogalma. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező. Egy ezen a téren értelmezett valós értékű (mérhető) $\xi(\omega)$ függvényt valószínűségi változónak fogunk nevezni. Azaz ξ az $\Omega \to R^1$ leképezés. Az a megkötés, hogy a függvény legyen mérhető azt jelenti, hogy minden x valós számra az $\{\omega\colon\ \xi(\omega) < x\} \in \mathcal{A}$ feltétel teljesül.*

Időnként hasznos nemcsak valós értékű valószínűségi változóról beszélni, hanem kissé általánosabban olyan valószínűségi változókat definiálni, amelyek értékeiket egy általános téren veszik fel. (Például bizonyos esetekben érdemes olyan valószínűségi változókat tekinteni, amelyek értéke komplex szám, vagy nem egyetlen szám, hanem egy szám-n-es. Az ilyen valószínűségi változókat vektor értékűnek szokták hívni.) Annak érdekében, hogy ezt megtehessük tekintsünk valamilyen X halmazt, és e halmaz kitüntetett részhalmazainak $\mathcal{X}$ osztályát, amelyek σ-algebrát alkotnak. Egy az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált X térbeli értékeket felvevő (mérhető) függvényt X-tér értékű valószínűségi változónak nevezünk. Az, hogy ez a függvény mérhető azt jelenti, hogy minden $Y \in \mathcal{X}$ halmazra az $\{\omega\colon\ \xi(\omega) \in Y\} \in \mathcal{A}$ feltétel teljesül.

Megjegyzés. Láttuk, hogy a valószínűségi változók (szép, azaz mérhető) függvényeket jelentenek. Az egyetlen különbség a korábban használt függvények és a most bevezetett valószínűségi változók között abban áll, hogy most egy általánosabb téren, egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező Ω halmazán és nem feltétlenül a számegyenesen értelmezett függvényeket tekintünk. Ez egyben azt is jelenti, hogy valószínűségi változókkal ugyanúgy számolhatunk, mint ahogy azt eddig függvényekkel tettük.

Felmerülhet a kérdés, hogy miért vezettük be azt a mesterkéltnek tűnő feltételt, hogy a valószínűségi változó legyen *mérhető* függvény. Azért, mert természetes kívánság az, hogy legyen értelme annak valószínűségéről beszélni, hogy az $\{\omega\colon\ \xi(\omega) < x\}$ esemény bekövetkezett. Viszont csak $\mathcal{A}$-beli halmazoknak a valószínűségéről tudunk beszélni. Természetes ennél is többet kivánni, nevezetesen azt, hogy a számegyenes minden ,,szép" B halmazára az $\{\omega\colon\ \xi(\omega) \in B\}$ halmaz legyen a $\mathcal{A}$ σ-algebrában, azaz legyen értelme ennek az eseménynek a valószínűségéről is beszélni. A mértékelméletben belátták, hogy amennyiben a számegyenes Borel mérhető részhalmazait tekintjük szép halmazoknak, akkor a mérhetőség általunk megkövetelt feltételeiből következik ez az erösebb tulajdonság is.

A fenti érvelés azt mutatja, hogy a mérhetőség megkövetelése természetes. Másrészt további eredmények azt is mutatják, hogy minden ,,definiálható" függvény egyben mérhető is, ezért ennek a tulajdonságnak a megkövetelése nem jelent igazi megszorítást, a továbbiakban megfeledkezhetünk róla. Az általános terekben értelmezett mérhetőségről hasonlót mondhatunk. Érdemes megjegyezni, hogy bár formálisan a valós és általános térbeli függvény mérhetőségét másképp definiáltuk ez a különbség csak látszólagos. Vezessük be a számegyenesen a Borel mérhető halmazok $\mathcal{B}$ σ-algebráját, és társítjuk a számegyenest ezzel a σ-algebrával. Ekkor a valós értékű mérhető függvényeknek az általános esetbeli definiciója az $(X, \mathcal{X}) = (R, \mathcal{B})$ választással megegyezik a valós értékű mérhető függvények eredeti definiciójával.

Diszkrét eloszlású valószínűségi változók és azok eloszlása. *Egy ξ valószínűségi változót egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn diszkrétnek vagy diszkrét eloszlásúnak hívunk, ha megadható egy véges vagy megszámlálhatóan végtelen számosságú $x_1, x_2, \cdots$, halmaz úgy, hogy az $\{\omega\colon\ \xi(\omega) = x_n\}$, $n = 1, 2, \dots$, halmazok teljes eseményrendszert alkotnak. Egy diszkrét eloszlású valószínűségi változó eloszlását meghatározzák azok az $x_1, x_2, \dots$, értékek amelyeket az felvesz és a $p_n = P(\xi = x_n)$ valószínűségek. (Jegyezzük meg, hogy $\sum\limits_n P(\xi = x_n) = 1$.)*

Legyenek $\xi_1, \dots, \xi_k$ diszkrét eloszlású valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Ezek együttes eloszlását meghatározzák a $P(\xi_1 = x_{j_1}, \xi_2 = x_{j_2}, \dots, \xi_k = x_{j_k})$, $j_s = 1, 2, \dots$, $1 \le s \le k$, valószínűségek.

Diszkrét eloszlású valószínűségi változók függetlensége. *Legyenek $\xi_1, \dots, \xi_k$ diszkrét eloszlású valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Vegyenek fel ezek a valószínűségi változók valamely $x_1, x_2, \dots$ értékeket. Azt mondjuk, hogy a ξ_1, ξ_2, ... , ξ_k valószínűségi változók függetlenek, ha*

$$P\left(\xi_1 = x_{j_1}, \xi_2 = x_{j_2}, \dots, \xi_k = x_{j_k}\right) = P\left(\xi_1 = x_{j_1}\right) P\left(\xi_2 = x_{j_2}\right) \cdots P\left(\xi_k = x_{j_k}\right)$$

minden $j_s = 1, 2, \dots$, $1 \le s \le k$, indexre.

1. Megjegyzés. A fenti definicióban feltettük, hogy a diszkrét eloszlású valószínűségi változók amelyeknek az együttes eloszlását vagy függetlenségét definiáltuk ugyanazokat az értékeket veszik fel. Ez azonban nem jelent megszorítást. Ugyanis, ha egyesítjük ezen valószínűségi változók értékkészletét, akkor újra legfeljebb megszámlálható számosságú halmazt kapunk, és mondhatjuk, hogy az egyes valószínűségi változók ezen halmazból veszik fel az értékeitket. Lehetséges, hogy ily módon sok nulla valószínűségi eseményt is bevezettünk, de ez nem okoz problémát.

2. Megjegyzés. Annak, hogy valószínűségi változóknak az eloszlását definiáltuk mélyebb oka van. A valószínűségszámításban olyan problémákkal foglalkozunk, amelyek csak attól függnek, hogy különböző eseményeknek mennyi a valószínűségük, különböző valószínűségi változók milyen értéket milyen valószínűséggel vesznek fel. Arra vagyunk kiváncsiak, hogy ezen ismereteknek milyen következményei vannak. Viszont a valószínűségszámíásban vizsgált kérdésekre adott válasz nem függ attól, hogy milyen véletlen hatások váltották ki a megfigyelt véletlen eredményeket, azaz milyen az a valószínűségi mező, amelyben a vizsgált események és valószínűségi változók megjelentek. Ez az oka annak, hogy minket igazából nem maguk a valószínűségi változók, hanem azok (együttes) eloszlása érdekel.

Megfogalmazok egy viszonylag egyszerű, de hasznos állítást.

Lemma független diszkrét eloszlású valószínűségi változók tulajdonságairól. *Legyenek adva $\xi_1, \dots, \xi_k$ független, diszkrét eloszlású valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Legyenek $A_1 \subset \{x_1, x_2, \dots\}, A_2 \subset \{x_1, x_2, \dots\}, \dots, A_k \subset \{x_1, x_2, \dots\}$ tetszőleges halmazok. Ekkor*

$$P\left(\xi_1 \in A_1, \xi_2 \in A_2, \dots, \xi_k \in A_k\right) = P\left(\xi_1 \in A_1\right) P\left(\xi_2 \in A_2\right) \cdots P\left(\xi_k \in A_k\right).$$

Speciálisan, ekkor tetszőleges $\{j_1,\dots,j_s\}\subset\{1,\dots,k\}$ indexhalmazra a $\xi_{j_1},\dots,\xi_{j_s}$ valószínűségi változók függetlenek.

A lemma bizonyítása. A függetlenség miatt

$$
\begin{aligned}
&P(\xi_1\in A_1,\xi_2\in A_2,\dots,\xi_k\in A_k)\\
&\quad=\sum_{x_1\in A_1}\sum_{x_2\in A_2}\cdots\sum_{x_k\in A_k}P(\xi_1=x_1,\xi_2=x_2,\dots,\xi_k=x_k)\\
&\quad=\sum_{x_1\in A_1}\sum_{x_2\in A_2}\cdots\sum_{x_k\in A_k}P(\xi_1=x_1)P(\xi_2=x_2)\cdots P(\xi_k=x_k)\,.
\end{aligned}
$$

Másrészt

$$
\begin{aligned}
&P(\xi_1\in A_1)\,P(\xi_2\in A_2)\cdots P(\xi_k\in A_k)\\
&\quad=\sum_{x_1\in A_1}P(\xi_1=x_1)\sum_{x_2\in A_2}P(\xi_2=x_2)\cdots\sum_{x_k\in A_k}P(\xi_k=x_k)\,.
\end{aligned}
$$

Elvégezve a beszorzásokat megkapjuk a lemma első állítását.

A lemma második állítását megkapjuk az első állítás speciális eseteként $A_{j_u}=\{x_{j_u}\}$, $1\le u\le s$ és $A_j=\{x_1,x_2,\dots\}$, $u\in\{1,\dots,k\}\setminus\{j_1,\dots,j_s\}$ választással.

Megjegyzés. Független, diszkrét eloszlású valószínűségi változók eredeti definiciója erősen kihasználta azt, hogy diszkrét eloszlású valószínűségi változókról beszélünk. Például, ha olyan $\xi_1,\dots,\xi_n$ valószínűségi változókról beszélünk, amelyek egy pontnak az egységintervallumra történő egyenletes eloszlású véletlen ledobását írják le, akkor a tekintett valószínűségi változók teljesítik a $P(\xi_j=x)=0$ relációt. Ilyen esetben a diszkrét eloszlások függetlenségét megadó definició mindössze a $0=0$ semmitmondó feltételt írja elő. Ezért az általános esetben más, tartalmasabb követelményt kell találnunk valószínűségi változók függetlenségének a jellemzésére. Az előző lemma állítása sugall egy ilyen definiciót. Bár bizonyos kényelmi okok miatt kissé más definiciót fogunk megadni, ki fog derülni, hogy az ekvivalens az e lemma eredménye által sugallt definicióval.

Érdemes megérteni események és diszkrét eloszlású valószínűségi változók függetlensége közötti kapcsolatot. Ennek érdekében bevezetem egy halmaz indikátorfüggvényének a fogalmát. Lehet, hogy ezt az egyszerű és természetes fogalmat analízis tanulmányaikban is bevezették, de ott valószínűleg a karakterisztikus függvény elnevezést használták. Ez az elnevezés természetes. Mi mégsem ezt használjuk, mert a valószínűségszámításban ez a kifejezés már foglalt; egy egészen más fogalmat illetnek ezzel a névvel.

Halmaz indikátorfüggvényének a definiciója. *Legyen adva egy $A\in\mathcal{A}$ esemény egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Az A halmaz indikátorfüggvényén azt a $\chi_A(\omega)$ valószínűségi változót értjük, amelyre $\chi_A(\omega)=1$, ha $\omega\in A$, és $\chi_A(\omega)=0$, ha $\omega\notin A$.*

A következő egyszerű lemma érvényes.

Lemma események és azok indikátorfüggvényének kapcsolatáról. *Legyenek* A_1, $\dots$, A_k *események egy* $(\Omega, \mathcal{A}, P)$ *valószínűségi mezőn. Az* $A_1,\dots,A_k$ *események akkor és csak akkor függetlenek, ha azok* $\chi_{A_1}(\omega),\dots,\chi_{A_k}(\omega)$ *indikátorfüggvényei függetlenek.*

A lemma bizonyítása. Ha a $\chi_{A_1}(\omega),\dots,\chi_{A_k}(\omega)$ indikátorfüggvények függetlenek, akkor ezek tetszőleges részhalmaza is független az előző lemma eredménye szerint. Ezért minden $\{j_1,\dots,j_s\}\subset\{1,\dots,k\}$ indexhalmazra

$$\begin{aligned} P\left(A_{j_1}\cap\cdots\cap A_{j_s}\right) &= P\left(\chi_{A_{j_1}}=1,\dots,\chi_{A_{j_s}}=1\right) \\ &= P\left(\chi_{A_{j_1}}=1\right)\cdots P\left(\chi_{A_{j_s}}=1\right) = P\left(A_{j_1}\right)\cdots P\left(A_{j_s}\right). \end{aligned}$$

Ha az $A_1,\dots,A_k$ események függetlenek, akkor egyes A_j eseményeket azok komplementerével helyettesítve ismét független eseményeket kapunk. Felírva minden ilyen relációt, megkapjuk az összes olyan azonosságot, amelyet a $\chi_{A_1}(\omega),\dots,\chi_{A_k}(\omega)$ indikátorfüggvényeknek teljesíteniük kell, ha azok függetlenek.

Ugyancsak hasznos bizonyos esetekben a következő lemma állítása.

Lemma. *Legyenek* $\xi_1,\dots,\xi_k,\xi_{k+1},\dots,\xi_{k+l}$ *diszkrét értékű, független valószínűségi változók,* $f(x_1,\dots,x_k)$ *egy* k *változós függvény, és definiáljuk az* $\eta=f(\xi_1,\dots,\xi_k)$ *valószínűségi változót. Ekkor* $\eta,\xi_{k+1},\dots,\xi_{k+l})$ *független valószínűségi változók.*

Bizonyítás: Rögzítsünk egy y számot, és vezessük be a következő $B=B(y)$ halmazt. $B=\{(x_1,\dots,x_k)\colon\ x_j\in X,\ 1\le j\le k,\ f(x_1,\dots,x_k)=y\}$. Ekkor

$$\begin{aligned} &P(\eta=y,x_{k+1}=x_{k+1},\dots,x_{k+l}=x_{k+l}) \\ &\quad=\sum_{(x_1,\dots,x_k)\in B} P(\xi_1=x_1,\dots,\xi_k=x_k,x_{k+1}=x_{k+1},\dots,x_{k+l}=x_{k+l}) \\ &\quad=\sum_{(x_1,\dots,x_k)\in B} P(\xi_1=x_1)\cdots P(\xi_k=x_k)P(x_{k+1}=x_{k+1})\cdots P(x_{k+l}=x_{k+l}) \\ &\quad=P(\eta=y)P(x_{k+1}=x_{k+1})\cdots P(x_{k+l}=x_{k+l}), \end{aligned}$$

és ezt kellett belátni.

Tárgyalom a következő feladatot.

Feladat.

Véletlenül meghívunk 30 embert. Tegyük fel, hogy az egyes embereknek egymástól függetlenül van születésnapjuk, és minden ember esetében $\frac{1}{365}$ annak a valószínűsége, hogy az év valamely napján született. Mi annak a valószínűsége, hogy van két ember a társaságban, akiknek ugyanaznap van a születésnapjuk?

Általánosabban, van n urna, amelyekbe bedobunk egymástól függetlenül k golyót úgy, hogy mindegyik golyó egyforma valószínűséggel esik az egyes urnákba. Mi annak a valószínűsége, hogy van olyan urna amelybe legalább két golyó esik? Érdekel minket továbbá ennek a valószínűségnek a viselkedése, ha mind az n mind a k szám nagy, és a $k = k(n)$ számnak megfelelő a nagyságrendje. Lássuk be, hogy a fenti valószínűségnek van határértéke, ha $n \to \infty$, $\frac{k}{\sqrt{n}} \to \alpha$ valamilyen $0 \le \alpha < \infty$ számmal, és határozzuk meg ezt a határértéket.

Megoldás: Jelölje ξ_j azt a valószínűségi változót, hogy a j-ik embernek az év hanyadik napján van a születésnapja. Ekkor a ξ_j, $1 \le j \le 30$, valószínűségi változók függetlenek, $P(\xi_j = l) = \frac{1}{365}$, $1 \le j \le 30$, $1 \le l \le 365$, és $P(\xi_j \ne \xi'_j \text{ ha } j \ne j')$ annak a valószínűsége, hogy mindenkinek különböző nap van a születésnapja. Ez a valószínűség viszont

$$\frac{365 \cdot 364 \cdots (365 - 30 + 1)}{365^k} = \prod_{j=1}^{29} \left(1 - \frac{j}{365}\right),$$

mert annak valószínűsége, hogy az első ember születésnapja az l_1-ik, a másodiké az l_2-ik és így tovább a k-ik ember születésnapja az l_k-ik napon van $\frac{1}{365^k}$, tetszőleges $1 \le l_j \le 365$, $1 \le j \le 30$ számok esetén, és ezeket a számokat $\prod\limits_{j=0}^{k-1}(365 - j)$ módon választhatjuk úgy, hogy mindegyik l_j szám különböző legyen. Így annak a valószínűsége, hogy van két ember akinek ugyanazon a napon van a születésnapja $1 - \prod\limits_{j=1}^{29} \left(1 - \frac{j}{365}\right)$.

Hasonlóan, annak valószínűsége, hogy ha k golyót dobunk n urnába az adott módon, akkor van olyan urna, amelyikbe legalább két golyó esik $1 - \prod\limits_{j=1}^{k-1} \left(1 - \frac{j}{n}\right)$.

Adjunk jó közelítést a $\log \prod\limits_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) = \sum\limits_{j=1}^{k-1} \log\left(1 - \frac{j}{n}\right)$ kifejezésre, ha $n \to \infty$, $\frac{k(n)}{\sqrt{n}} \to \alpha$. Heurisztikus érvelés szerint mivel $\log\left(1 - \frac{j}{n}\right) \sim -\frac{j}{n}$ a $\log(1 + x)$ függvény Taylor sorfejtése szerint, ezért $\sum\limits_{j=1}^{k-1} \log\left(1 - \frac{j}{n}\right) \sim -\sum\limits_{j=1}^{k-1} \frac{j}{n} = -\frac{(k-1)k}{2n}$, ahonnan $\log \prod\limits_{j=1}^{k(n)-1} \left(1 - \frac{j}{n}\right) \to -\frac{\alpha^2}{2}$, ha $n \to \infty$, és $\frac{k(n)}{\sqrt{n}} \to \alpha$. Ez a számolás precizzé tehető, ha felhasználjuk például azt az egyenlőtlenséget, amely szerint

$$\left|\log\left(1 - \frac{j}{n}\right) + \frac{j}{n}\right| \le \frac{2j^2}{n^2} \le \frac{\text{const.}}{n}, \quad \text{ha } n \text{ elég nagy és } \frac{j}{\sqrt{n}} \le \alpha + 1,$$

ami szintén következik a $\log(1 + x)$ Taylor sorfejtéséből. Innen azt kapjuk, hogy $1 - \prod\limits_{j=1}^{k(n)-1} \left(1 - \frac{j}{n}\right) \to 1 - e^{-\alpha^2/2}$, ha $n \to \infty$, és $\frac{k(n)}{\sqrt{n}} \to \alpha$.

Következő témánk a valószínűségszámítás egyik legfontosabb fogalmához, a várható értékhez kapcsolódik. Először, a jobb megértés érdekében csak diszkrét eloszlású valószínűségi változó várható értékét vezetem be. Tárgyalom annak legfontosabb tulajdonságait. Nagyon fontos, hogy megértsük jól azokat a feladatokat, amelyek arról szólnak, hogy hogyan tudjuk valószínűségi változók várható értékét effektíve kiszámítani.

Diszkrét eloszlású valószínűségi változó várható értéke. *Legyen ξ diszkrét eloszlású valószínűségi változó, amely $x_1, x_2, \dots$ értékeket vesz fel $p_k = P(\xi = x_k)$ valószínűséggel, $k = 1, 2, \dots,$ $\sum\limits_{k=1}^{\infty} P(\xi = x_k) = 1$. A ξ valószínűségi változó várható értéke az*

$$E\xi = \sum_{k=1}^{\infty} x_k P(\xi = x_k)$$

összeg, feltéve hogy ez az összeg abszolut konvergens. Ha ez az összeg nem abszolut konvergens, akkor nem definiáljuk az $E\xi$ várható értéket.

1. Megjegyzés. Most és a továbbiakban az egyszerűbb és egységesebb jelölés érdekében feltesszük, hogy a diszkrét eloszlású ξ valószínűségi változó (megszámlálhatóan) végtelen sok értéket vesz fel. Ez nem jelent megszorítást, mert ha ξ csak véges sok értéket vesz fel, akkor is feltehetjük, hogy ezenkívül még felvesz végtelen sok értéket nulla valószínűséggel.

2. Megjegyzés. A várható értéket az angol *expectation* szó kezdőbetűjével az E betűvel jelöljük. Manapság ez a legelterjedtebb jelölés, bár használják az M betűt is a német *mathematischer Mittelwert* szó alapján.

Emlékeztető. Azt mondjuk, hogy a $\sum\limits_{n=1}^{\infty} y_n$ sor abszolut konvergens, ha nemcsak a $\sum\limits_{n=1}^{\infty} y_n$, hanem a $\sum\limits_{n=1}^{\infty} |y_n|$ sor is konvergens, azaz $\sum\limits_{n=1}^{\infty} |y_n| < \infty$. Ha csak a $\sum\limits_{n=1}^{\infty} y_n$ sor konvergens, és a $\sum\limits_{n=1}^{\infty} |y_n|$ sor divergens, akkor feltételesen konvergens sorról beszélünk.

Idézzünk fel néhány eredményt, amelyek megmagyarázzák miért követeltük meg a várható érték definiciójában, hogy a $\sum\limits_{k=1}^{\infty} x_k P(\xi = x_k)$ sor ne csak konvergens, hanem abszolut konvergens is legyen.

Ha a $\sum\limits_{n=1}^{\infty} y_n$ sor abszolut konvergens, akkor ennek tetszőleges átrendezése azaz tetszőleges olyan $\sum\limits_{n=1}^{\infty} y_{k(n)}$ összeg, amelyre a $k(n) = j$ egyenletnek pontosan egy megoldása van minden $1 \le j < \infty$ egész számra szintén konvergens, és ugyanaz a határértéke mint az eredeti sorrendben. Ez a tulajdonság nem abszolut konvergens sorozatokra már nem érvényes. Sőt, ebben az esetben a következő negatív jellegű állítás érvényes.

Legyen $\sum\limits_{n=1}^{\infty} x_n$ egy konvergens, de nem abszolut konvergens összeg. Ekkor tetszőleges $-\infty \le \alpha \le \infty$ számra létezik e sor tagjainak olyan (az α számtól függő) $k(n)$, $n = 1, 2, \dots$ átrendezése, amelyre $\lim\limits_{N\to\infty} \sum\limits_{n=1}^{N} x_{k(n)} = \alpha$.

Miért fontos az a megkötés, hogy a várható értéket definiáló összeg nemcsak konvergens, hanem abszolut konvergens is? Vegyük például azt az esetet, amikor a ξ valószínűségi változó értékei az összes racionális szám halmaza. A racionális számok, mivel megszámlálhatóan sokan vannak felsorolhatóak mint egy $x_1, x_2, \dots,$ sorozat, de ennek a felsorolásnak nincs egy természetes sorrendje. Ezért a várható értéket definiáló összeg csak akkor értelmes, ha annak értéke nem függ attól, hogy a ξ valószínűségi változó lehetséges x_k értékeit milyen sorrendben soroljuk fel.

Az abszolut konvergens sorok fent említett tulajdonsága szerepelt az analízis tananyagban. A bizonyítás megismétlése helyett lássunk inkább egy példát, amely megmutatja, hogy a nem abszolut konvergens sorok másképpen viselkednek, mint az abszolut konvergensek. Tekintsük az $S = 1 - 1 + \frac{1}{2} - \frac{1}{2} + \frac{1}{3} - \frac{1}{3} + \frac{1}{4} - \frac{1}{4} + \cdots$ sort. Ez a sor konvergens, határértéke zéró, ugyanakkor nem abszolut konvergens, mert $\sum\limits_{k=1}^{\infty} \frac{1}{k} = \infty$. Tekintsük e sor következő átrendezését: $S = 1 + \sum\limits_{k=1}^{\infty} \left(\frac{1}{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}-1} - \frac{1}{k} \right)$. Ez az átrendezett sor divergens, mert a k-ik tagra $\frac{1}{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}-1} - \frac{1}{k} \ge \frac{2^k}{2^{k+1}} - \frac{1}{k} = \frac{1}{2} - \frac{1}{k}$. Röviden, kissé nagyvonalúan, azt mondhatjuk, hogy abszolut konvergens sorozatokkal ugyanolyan szabadon számolhatunk (átrendezhetjük őket, abszolut konvergens sorok szorzataiban elvégezhetjük a tagonkénti szorzást), mint véges összegek esetében, míg nem abszolut konvergens sorok esetében ezt nem tehetjük meg.

Számítsuk ki a várható értéket egy viszonylag egyszerű feladatban.

Egy szabályos pénzdarabot feldobunk 100-szor egymás után. Számoljuk ki a fejdobások számának a várható értékét.

Megoldás: Annak valószínűsége, hogy pontosan k fejdobás történik 100 dobás esetében $\binom{100}{k}2^{-100}$, $0 \le k \le 100$. Ezért $E\xi = \sum\limits_{k=0}^{100} k\binom{100}{k}2^{-100}$ a keresett várható érték. Ezt az összeget ki tudjuk számítani a következő észrevétel segítségével: $k\binom{n}{k} = n\binom{n-1}{k-1}$, ha $1 \le k \le n$. Innen $n = 100$ választással

$$\begin{aligned} E\xi &= \sum_{k=0}^{100} k\binom{100}{k}2^{-100} = 100\sum_{k=0}^{99}\binom{99}{k}2^{-100} = 50\sum_{k=0}^{99}\binom{99}{k}\left(\frac{1}{2}\right)^{99} \\ &= 50 \cdot \left(\frac{1}{2}+\frac{1}{2}\right)^{99} = 50. \end{aligned}$$

Heurisztikusan a következő módon érvelnénk: Egy pénzfeldobás esetén a fejek számának a várható értéke $\frac{1}{2}$, száz pénzfeldobás esetén a pénzdobások számának várható értéke $100 \cdot \frac{1}{2} = 50$.

A következő fontos eredmény, amelynek bizonyítása a következő előadás anyaga lesz megmagyarázza, hogy a fenti heurisztika miért adott helyes eredményt, és ennek segítségével hogyan lehet kiszámítani sok esetben egyszerűen valószínűségi változók várható értékét.

Tétel. *Legyenek ξ_1, ξ_2 (diszkrét eloszlású) valószínűségi változók, amelyeknek létezik várható értékük. Ekkor a $\xi_1 + \xi_2$ összegnek is létezik várható értéke, és*

$$E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2.$$

Következmény. *Legyenek adva $\xi_1, \xi_2, \dots, \xi_k$ (diszkrét eloszlású) valószínűségi változók, amelyeknek létezik várható értékük, és legyenek $c_1, \dots, c_k$ valós számok. Ekkor a $c_1\xi_1 + c_2\xi_2 + \cdots + c_k\xi_k$ valószínűségi változónak is létezik várható értéke, és*

$$E\left(c_1\xi_1 + c_2\xi_2 + \cdots + c_k\xi_k\right) = c_1E\xi_1 + c_2E\xi_2 + \cdots + c_kE\xi_k.$$

Megjegyzés. Érdemes hangsúlyozni, hogy a fenti eredményben nem tételeztük fel, hogy a tekintett valószínűségi változók függetlenek. Ez az észrevétel bizonyos alkalmazásokban hasznos. Később tanulni fogjuk, hogy a most megfogalmazott eredmény nemcsak diszkrét eloszlású valószínűségi változók összegeire érvényes.

Lássuk, hogy hogyan tudjuk kiszámítani az előző feladatban tekintett várható értéket a fenti eredmény segítségével.

Vezessük be a következő valószínűségi változókat: $\xi_j = 1$, ha a j-ik dobás eredménye fej, $\xi_j = 0$, ha a j-ik dobás eredménye írás. Azaz a ξ_j valószínűségi változó azt számolja, hogy mennyi a j-ik dobás eredményének hozadéka a fej-dobások számához. Ekkor minket az $E\left(\sum\limits_{j=1}^{100} \xi_j\right)$ várható érték érdekel. Viszont a fenti eredmény alapján $E\left(\sum\limits_{j=1}^{100} \xi_j\right) = \sum\limits_{j=1}^{100} E\xi_j$, és $E\xi_j = 1 \cdot P(\xi_j = 1) + 0 \cdot P(\xi_j = 0) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}$ minden $1 \le j \le 100$ indexre, ahonnan $E\left(\sum\limits_{j=1}^{100} \xi_j\right) = 100 \cdot \frac{1}{2} = 50$.

Feladat:

Egy szabályos dobókockát feldobunk 100-szor egymás után. Számoljuk ki a dobáseredmények összegének a várható értékét.

Megoldás: Vezessük be a következő ξ_j, $1 \le j \le 100$ valószínűségi változókat: $\xi_j = 1$, ha a j-ik dobás 1, $\xi_j = 2$, ha a j-ik dobás 2, $\xi_j = 3$, ha a j-ik dobás 3, $\xi_j = 4$, ha a j-ik dobás 4, $\xi_j = 5$, ha a j-ik dobás 5, és $\xi_j = 6$, ha a j-ik dobás 6. Ekkor a vizsgált összeg $S = \sum\limits_{j=1}^{100} \xi_j$, ahonnan $ES = \sum\limits_{j=1}^{100} E\xi_j$. Mivel $E\xi_j = \frac{1}{6}(1 + 2 + \cdots + 6) = 3.5$ minden $1 \le j \le 100$ indexre, innen $ES = 100 \cdot 3.5 = 350$.

Végül két állítást bizonyítok, ahol felhasználom a most tanult fogalmakat, és azt, hogy a várható érték additív. Egyrészt megmutatom, hogy hogyan lehet a valószínűségszámítási szita formulát a kombinatorikus szita formulából levezetni. Másrészt megfogalmazom és bebizonyítom az első előadásban kimondott bizonyos húzássorozatok valószínűségéről szóló lemma egy természetes és hasznos általánosítását.

Annak érdekében, hogy a valószínűségi szita formulát igazoljuk a kombinatorikai szita formula segítségével tekintsük azt az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt, ahol az A_1, $\dots$, A_n események definiálva vannak, és definiáljuk minden $\omega \in \Omega$ elemi eseményre és $B \in \mathcal{A}$ halmazra azt a $\bar{B} = \bar{B}(\omega)$ (1 vagy nulla elemet tartalmazó halmazt), amely az ω elemből áll, ha $\omega \in B$ és az üres halmazzal egyenlő, ha $\omega \notin B$. Megmutatom, hogy a valószínűségszámítási szita formula levezethető a kombinatorikus szita formulából, ha azt az $\bar{A}_1(\omega), \dots, \bar{A}_n(\omega)$ halmazok uniójára alkalmazzuk minden $\omega \in \Omega$ elemi eseményre, majd ezeket az azonosságokat ‘átlagoljuk’ az $\omega \in \Omega$ elemi eseményekre a P valószínűségi mérték szerint. E procedura során felhasználjuk a várható érték néhány fontos tulajdonságát. Nevezetesen arra van szükségünk, hogy események valószínűsége és indikátorfüggvényük várható értéke egyenlő, és a várható érték additív. Vegyük észre azt is, hogy $|\bar{B}(\omega)| = I_B(\omega)$ minden $\omega \in \Omega$ elemi eseményre és $B \in \mathcal{A}$ halmazra.

Vezessük be a valószínűségi szita formulában definiált S_k összegeknek megfelelő $\bar{S}_k(\omega) = \sum\limits_{1 \le j_1 < \cdots < j_k \le n} |\bar{A}_{j_1}(\omega) \cap \cdots \cap \bar{A}_{j_k}(\omega)| = \sum\limits_{1 \le j_1 < \cdots < j_k \le n} I_{A_{j_1} \cap \cdots \cap A_{j_k}}(\omega)$, $1 \le k \le n$, valószínűségi változókat. A várható érték additív tulajdonsága és események valószínűsége és indikátor függvényük várható értékének egyenlősége miatt $E\bar{S}_k(\omega) = S_k$,

$1 \le k \le n$. Továbbá a kombinatorikus szita formula alapján $|I_{A_1 \cup \dots \cup A_n}(\omega)| = |\bar{A}_1)(\omega) \cup \dots \cup \bar{A}_n(\omega)| = \bar{S}_1(\omega) - \bar{S}_2(\omega) + \bar{S}_3(\omega) - \dots + (-1)^{n+1}\bar{S}_n(\omega)$. Ebben az azonosságban várható értéket véve és újból felhasználva a várható érték additívitását illetve események valószínűsége és indikátor függvényük várható értékének az egyenlőségét megkapjuk a valószínűségi diagram formula azonosságát. Az ebben az eredményben felírt egyenlőtlenségek hasonlóan bizonyíthatóak a kombinatorikus diagram formula egyenlőtlenségeinek a segítségével.

Annak érdekében, hogy megfogalmazhassam az első előadásban kimondott bizonyos húzássorozatok valószínűségéről szóló lemma egy hasznos általánosítását először bevezetem a következő fogalmat.

Felcserélhető valószínűségi változók véges sorozatának a definiciója. *Legyenek $\xi_1, \dots, \xi_n$ diszkrét eloszlású valószínűségi változók valamely $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyek értékeiket valamely $X = \{x_1, x_2, \dots\}$ véges vagy megszámlálhatóan végtelen halmazon veszik fel. Jelölje $\Pi(n)$ az $\{1, 2, \dots, n\}$ halmaz összes permutációjából álló halmazt. Azt mondjuk, hogy a $\xi_1, \dots, \xi_n$ valószínűségi változók felcserélhetőek, ha akárhogy választunk ki egy $\pi = (\pi(1), \dots, \pi(n)) \in \Pi(n)$ permutációt és $x_l \in X$, $1 \le l \le n$, elemeket az X halmazból, teljesül a*

$$P(\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_n = x_n) = P(\xi_{\pi(1)} = x_1, \xi_{\pi(2)} = x_2, \dots, \xi_{\pi(n)} = x_n)$$

azonosság.

Értsük meg, hogy a felcserélhető valószínűségi változók szemléletes tartalma a következő. Felcserélhető valószínűségi változók esetében a $P(\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_n = x_n)$ valószínűség csak attól függ, hogy a tekintett $\xi_1, \dots, \xi_n$ valószínűségi változók milyen $x \in X$ értékeket vesznek fel és milyen multiplicitással, de nem függ attól, hogy milyen sorrendben veszik fel ezeket az értékeket. Vegyük észre, hogy egy ilyen tulajdonságot használtunk fel a visszatevéses urnamodellek vizsgálatában is. Érvényes a következő egyszerű, de hasznos összefüggés.

Lemma felcserélhető valószínűségi változók véges dimenziós eloszlásairól. *Legyenek $\xi_1, \dots, \xi_n$ diszkrét eloszlású, felcserélhető valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyek értékeiket valamely $X = \{x_1, x_2, \dots\}$ véges vagy megszámlálható halmazon veszik fel. Tekintsünk valamilyen rögzített $1 \le k \le n$ számot, az X tér valamely $x_l \in X$, $1 \le l \le k$, elemeinek sorozatát és az $\{1, \dots, n\}$ halmaz egy k elemű $\{j_1, \dots, j_k\} \subset \{1, \dots, n\}$ részhalmazát. Érvényes a*

$$P(\xi_1 = x_1, \dots, \xi_k = x_k) = P(\xi_{j_1} = x_1, \dots, \xi_{j_k} = x_k)$$

azonosság.

Bizonyítás. Rögzítsünk egy olyan $\pi \in \Pi(n)$ permutációt, amelyre $\pi(1) = j_1$, $\pi(2) = j_2$, ..., $\pi(k) = j_k$. Felírhatjuk a

$$\begin{aligned} &P(\xi_1 = x_1, \dots, \xi_k = x_k) \\ &\qquad = \sum_{x_l\colon x_l \in X, \text{ ha } k+1 \le l \le n} P(\xi_1 = x_1, \dots, \xi_k = x_k, x_{k+1} = x_{k+1}, \dots, \xi_n = x_n) \end{aligned}$$

azonosságot. Másrészt a valószínűségi változók felcserélhetősége miatt

$$\begin{aligned}
&\sum_{x_l\colon x_l\in X,\ \text{ha}\ k+1\le l\le n} P(\xi_1=x_1,\dots,\xi_k=x_k,x_{k+1}=x_{k+1},\dots,\xi_n=x_n)\\
&\quad=\sum_{x_l\colon x_l\in X,\ \text{ha}\ k+1\le l\le n} P(\xi_{\pi(1)}=x_1,\dots,\xi_{\pi(k)}=x_k,x_{\pi(k+1)}=x_{k+1},\dots,\xi_{\pi(n)}=x_n)\\
&\quad=P(\xi_{\pi(1)}=x_1,\dots,\xi_{\pi(k)}=x_k)=P(\xi_{j_1}=x_1,\dots,\xi_{j_k}=x_k),
\end{aligned}$$

és ezekből az azonosságokból következik a lemma állítása.

Házi feladat:

Rögzítsünk valamely $r\ge 1$ és $s\ge 0$ egész számokat. Legyen adva egy urna abban valahány golyó, amelyek mindegyike az $1,\dots,r$ számokkal jelzett színek valamelyikét veszi fel. Húzzunk ki n golyót az urnából egymás után (minden urnában levő golyót egyforma valószínűséggel húzzuk) úgy, hogy azután, hogy kihúztunk egy golyót visszadobunk az urnába s darab a kihúzott golyó színével azonos színű golyót. Jelölje ξ_j a j-ik kihúzott golyó színét. Lássuk be, hogy $\xi_1,\dots,\xi_n$ felcserélhető valószínűségi változók.

Kiegészítés 1.

Az előadásban megemlítettem a következő eredményt:

Állítás. *Legyen x_n valós számoknak olyan sorozata, amelyre $0\le x_n<1$ minden $n=1,2,\dots,$ számra. Ekkor $\prod\limits_{k=1}^{\infty}(1-x_k)>0$, ha $\sum\limits_{k=1}^{\infty}x_k<\infty$, és $\prod\limits_{k=1}^{\infty}(1-x_k)=0$, ha $\sum\limits_{k=1}^{\infty}x_k=\infty$.*

Indoklás. Heurisztikusan a következző módon érvelhetünk: Logaritmust véve kapjuk, hogy $\prod\limits_{k=1}^{\infty}(1-x_k)>0$ akkor és csak akkor, ha $\sum\limits_{k=1}^{\infty}-\log(1-x_k)<\infty$. Viszont azt, hogy az utóbbi összeg konvergens-e vagy divergens az dönti el, hogy kis x_k számokra a $-\log(1-x_k)$ összeadandók milyen kicsik. Mivel $-\log(1-x)=x+\frac{x^2}{2}+\frac{x^3}{3}+\cdots$, a $-\log(1-x_k)$ mennyiségnek természetes jó közelítése az x_k kifejezés, (a Taylor sor első tagja), és ez azt sugallja, hogy a $\sum\limits_{k=1}^{\infty}-\log(1-x_k)$ illetve $\sum\limits_{k=1}^{\infty}x_k$ végtelen sorok egyszerre konvergensek vagy divergensek.

Némi munkával a fenti érvelés precízzé tehető. Először azt kell meggondolni, hogy a feladat állításában a $\prod\limits_{k=1}^{\infty}(1-x_k)>0$ tulajdonság a $\sum\limits_{k=1}^{\infty}-\log(1-x_k)<\infty$ egyenlőtlenséggel, a $\prod\limits_{k=1}^{\infty}(1-x_k)=0$ tulajdonság pedig a $\sum\limits_{k=1}^{\infty}-\log(1-x_k)=\infty$ relációval ekvivalens. Ehhez azt kell felhasználni, hogy esetünkben $0<1-x_k<1$, és $-\log(1-x_k) > 0$. Ezután vegyük észre, hogy a $-\sum\limits_{k=1}^{\infty}\log(1-x_k)$ és $\sum\limits_{k=1}^{\infty}x_k$ sorok mindegyike divergens, ha rögzítve egy kis pozitív számot például az $\frac{1}{10}$ számot az x_k számsorozatnak végtelen sok olyan tagja van, amelyek nagyobbak mint ez az $\frac{1}{10}$ szám. Továbbá, égy végtelen összeg konvergenciáját vagy divergenciáját nem befolyásolja véges sok tagjának az értéke. Ezért az összegekben szereplő $x_k \ge \frac{1}{10}$ tagokat elhagyva elég csak azzal az esettel foglalkozni, amikor $x_k \le \frac{1}{10}$ minden k-ra. Viszont ekkor $\frac{1}{2}x_k < -\log(1+x_k) < 2x_k$ minden $k = 1, 2, \dots$ számra, ahonnan $\frac{1}{2}\sum\limits_{k=1}^{n}x_k < \sum\limits_{k=1}^{n} -\log(1-x_k) < 2\sum\limits_{k=1}^{n}x_k$ minden n-re, és innen következik, hogy a $\sum\limits_{k=1}^{\infty}x_k$ és $\sum\limits_{k=1}^{\infty} -\log(1-x_k)$ sorok egyszerre konvergálnak vagy divergálnak.

Kiegészítés 2. *A valószínűségszámítási szita formula és annak egy általánosítása.*

A valószínűségszámítási szita formula egy olyan bizonyítását ismertetem, amely egy önmagában is érdekes, és más esetekben is alkalmazható eredményen alapul. Ezt az alábbi lemmában fogalmazom meg.

Lemma. *Legyenek adva valamely $A_1, \dots, A_n$ események egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, és definiáljuk ezek felhasználásával véges sok unió, metszet és komplementer segítségével bizonyos $B_j = f_j(A_1, \dots, A_n)$, $1 \le j \le k$, eseményeket. Rögzítsünk valamely $c_1, \dots, c_k$ valós számokat. A*

$$\sum_{j=1}^{k} c_j P(B_j) = \sum_{j=1}^{k} c_j P\left(f_j(A_1, \dots, A_n)\right) \ge 0$$

egyenlőtlenség teljesül tetszőleges valószínűségi mezőn definiált tetszőleges $A_1, \dots, A_n$ halmazokra, ha speciálisan teljesül abban a (2^n számú) speciális esetben, amikor mindegyik A_j halmaz vagy az Ω biztos vagy az $\emptyset$ üres esemény.

A lemma bizonyítása. Jelölje $A_j^{-1} = \Omega \setminus A_j$ az A_j halmaz komplementerét. Vezessük be az $A_j^1 = A_j$ jelölést, és jelöljön $(k_1, \dots, k_n)$, $k_j = \pm 1$, $1 \le j \le n$, valamely n hosszúságú ± 1 sorozatot. Írjuk fel mindegyik $B_j = f_j(A_1, \dots, A_n)$ eseményt konjunktív normálforma alakban. Felhasználva, hogy a konjunktív normálformában diszjunkt halmazok uniója jelenik meg, a vizsgálandó egyenlőtlenség felírható

$$\sum_{(k_1,\dots,k_n)\colon\, k_j=\pm 1,\ j=1,\dots,n} d(k_1, \dots, k_n) P\left(A_1^{k_1} \cap \dots \cap A_n^{k_n}\right) \ge 0 \tag{A}$$

alakban alkalmas $d(k_1, \dots, k_n)$ együtthatókkal. Az (A) egyenlőtlenség nyílván érvényes minden $A_1, \dots, A_n$ halmazrendszerre, ha $d(k_1, \dots, k_n) \ge 0$ minden $(k_1, \dots, k_n)$ argumentumra. Ezért elég megmutatni, hogy amennyiben az (A) egyenlőtlenség teljesül minden olyan speciális esetben, amikor az A_j, $1 \le j \le n$, halmazok mindegyike vagy az Ω biztos vagy az $\emptyset$ üres esemény, akkor $d(k_1,\dots,k_n)\ge 0$ az összes $(k_1,\dots,k_n)$ argumentumra.

Ezt megmutatandó, rögzítsünk egy n hosszúságú $(k_1,\dots,k_n)$ ±1 sorozatot, és definiáljuk ennek segítségével $A_1,\dots,A_n$ halmazoknak azt a sorozatát, amelyre $A_j=\Omega$, ha $k_j=1$, és $A_j=\emptyset$, ha $k_j=-1$, $1\le j\le n$. Ezzel a választással az (A) formula baloldalán szereplő kifejezés $d(k_1,\dots,k_n)$-nel egyenlő, ezért ez csak úgy lehet nem negatív, ha $d(k_1,\dots,k_n)\ge 0$. Mivel ezt az érvet minden lehetséges $(k_1,\dots,k_n)$ ±1 sorozatra alkalmazhatjuk, innen következik a lemma állítása.

A valószínűségszámítási szita formula bizonyítása a lemma segítségével. A lemma alapján elegendő belátni a valószínűségszámítási szita formulát abban a speciális esetben, ha mindegyik A_j esemény a biztos vagy az üres esemény. Ekkor ugyanis a szita formula bal és jobboldalán szereplő kifejezések különbségéről tudjuk, hogy egyrészt nagyobb vagy egyenlő mint nulla, másrészt kisebb egyenlő, mint nulla. Tekintsük azokat az eseteket, amikor az A_j események között r darab biztos és $n-r$ üres esemény van, $0\le r\le n$. Feltehetjük, hogy $r\ge 1$, mert $r=0$ esetén, amikor mindegyik A_j esemény az üres halmaz, az azonosság mindkét oldala nullával egyenlő. Ha $1\le r\le n$, akkor $S_l=\binom{r}{l}$ az $1\le l\le r$ esetben, és $S_l=0$, ha $l\ge r$. Ezért az $1\le r\le n$ esetben a szita formula jobboldalán álló kifejezés $\sum\limits_{l=1}^{r}(-1)^{l+1}S_l=\sum\limits_{l=1}^{r}(-1)^{l+1}\binom{r}{l}=1-\sum\limits_{l=0}^{r}(-1)^{l}\binom{r}{l}=1-(1-1)^r=1$, a baloldali kifejezés pedig szintén $P(A_1\cup\dots\cup A_n)=P(\Omega)=1$.

Hasonló meggondolások alapján a szita formulában megfogalmazott egyenlőtlenségek igazolához elég megmutatni azt, hogy $1-\sum\limits_{l=1}^{s}(-1)^{l+1}\binom{r}{l}\le 0$, azaz $\sum\limits_{l=0}^{s}(-1)^{l}\binom{r}{l}\le 0$, ha $1\le s\le r$, és s páratlan szám, és $\sum\limits_{l=0}^{s}(-1)^{l}\binom{r}{l}\ge 0$, ha $1\le s\le r$, és s páros szám. (A bizonyítandó állítás ezen redukciójához úgy jutunk a fenti lemma segítségével, hogy r-val jelöljük azon A_j halmazok számát, amelyekre $A_j=\Omega$, és felírjuk, hogy mit jelent a lemma szerint bizonyítandó állítás ebben az esetben. Az $r=0$ esetben az ellenőrizendő egyenlőtlenségek nyilvánvalóan teljesülnek, mert ekkor minden tekintett kifejezés nullával egyenlő.)

A bizonyítandó egyenlőtlenségeket először az $s\le\left[\frac r2\right]$ esetben látjuk be, ahol $[x]$ jelöli az x szám egész részét. Felhasználva az $\binom{r}{0}\le\binom{r}{1}\le\dots\le\binom{r}{[\frac r2]}$ egyenlőtlenségeket azt kapjuk ebben az esetben, hogy

$$\sum_{l=0}^{2s'+1}(-1)^l\binom{r}{l}=\sum_{l=0}^{s'}\left(\binom{r}{2l}-\binom{r}{2l+1}\right)\le 0,\ \text{ha } s=2s'+1,$$

és

$$\sum_{l=0}^{2s'}(-1)^l\binom{r}{l}=\binom{r}{0}+\sum_{l=1}^{s'}\left(\binom{r}{2l}-\binom{r}{2l-1}\right)\ge 0,\ \text{ha } s=2s'.$$

Az $\left[\frac r2\right]<s<r$ eset visszavezethető az $s\le\left[\frac r2\right]$ esetre a $\sum\limits_{l=0}^{s}(-1)^l\binom{r}{l}=-\sum\limits_{l=s+1}^{r}(-1)^l\binom{r}{l}=$

$-\sum\limits_{l=s+1}^{r}(-1)^l\binom{r}{r-l} = (-1)^{r+1}\sum\limits_{l=0}^{r-s-1}(-1)^l\binom{r}{l}$ azonosság segítségével, mivel ezen azonosság jobboldalának az előjelét ismerjük a $r-s-1 \le \left[\frac{r}{2}\right]$ egyenlőtlenség miatt. Végül az $s=r$ esetben $\sum\limits_{l=0}^{s}(-1)^l\binom{r}{l} = 0$, ezért a kívánt egyenlőtlenség ebben az esetben is teljesül.

Ismertetem a szita formula egy hasonlóan bizonyítható általánosítását.

A szita formula egy általánosítása. *Legyenek adva tetszőleges $A_1,\dots,A_n$ események egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, és vezessük be az $N=N(\omega)$ valószínűségi változót, amely egyenő azon j indexek számával, amelyekre az $\omega\in A_j$ reláció teljesül. Ekkor*

$$P(N=r)=\sum_{i=0}^{n-r}(-1)^i\binom{i+r}{i}S_{i+r}, \quad \text{minden } 0\le r\le n \text{ számra,}$$

ahol

$$S_0=1, \quad S_k=\sum_{1\le i_1<\cdots<i_k\le n}P\left(A_{i_1}\cap\cdots\cap A_{i_k}\right), \quad 1\le k\le n.$$

Továbbá ezen összeg részletösszegei felváltva alsó illetve felső becslést adnak a tekintett valószínűségekre, azaz

$$\sum_{i=0}^{s}(-1)^i\binom{i+r}{i}S_{i+r}\ge P(N=r) \quad \text{ha } 0\le s\le n-r \text{ és } s \text{ páros szám.}$$

és

$$\sum_{i=0}^{s}(-1)^i\binom{i+r}{i}S_{i+r}\le P(N=r) \quad \text{ha } 0\le s\le n-r \text{ és } s \text{ páratlan szám.}$$

Megjegyzés. Mivel $P(N=0)=1-P(A_1\cup\cdots\cup A_n)$ és $S_0=1$ ez az eredmény az $r=0$ esetben megegyezik a szita formulával.

A szita formula általánosításának a bizonyítása. A szita formula bizonyításához hasonlóan ebben az esetben is redukálhatjuk a bizonyítandó állítás igazolását arra a speciális esetre, amikor a tekintett A_j események között k darab Ω biztos és $n-k$ darab üres esemény van, $0\le k\le n$. Sőt, azt is feltehetjük, hogy $k>r$. Ugyanis $k<r$ esetén a felírt azonosság illetve egyenlőtlenségek mind a két oldala 0-val, míg $k=r$ esetén 1-gyel egyenlő. Ugyanis $S_{r+i}=0$, ha $r+i>k$, ami minden $i\ge 0$-ra teljesül, ha $k<r$, ha $k<r$. Ezért ebben az esetben a tekintett azonosság és egyenlőtlenségek jobboldalán 0 áll. A $k=r$ esetben $S_{r+i}=0$, ha $i\ge 1$, és $S_{r+0}=1$. Másrészt $P(N=r)=0$, ha $k\ne r$, és $P(N=r)=k$, ha $k=r$. Ezekből az állításokból következik az állítás fenti redukciójának a jogossága.

Másrészt a tekintett speciális esetben (amikor pontosan k darab biztos esemény van, és a többi $n-k$ esemény pedig üres) $S_{r+i} = \binom{k}{r+i}$, és $\binom{i+r}{i}\binom{k}{i+r} = \binom{k}{r}\binom{k-r}{i}$. Innen a bizonyítandó azonosság jobboldalán álló kifejezés

$$\begin{aligned}\sum_{i=0}^{n-r}(-1)^i\binom{i+r}{i}S_{r+i} &= \sum_{i=0}^{n-r}(-1)^i\binom{i+r}{i}\binom{k}{r+i} = \sum_{i=0}^{n-r}(-1)^i\binom{k}{r}\binom{k-r}{i}\\ &= \binom{k}{r}\sum_{i=0}^{k-r}(-1)^i\binom{k-r}{i} = \binom{k}{r}\sum_{i=0}^{k-r}(1-1)^{k-r} = 0,\end{aligned}$$

ha $k > r$. Innen következik, hogy a felírt azonosság valóban érvényes. Az egyenlőtlenségek bizonyítása hasonló, csak ebben az esetben a számolás végén a $\sum\limits_{i=0}^{k-r}(-1)^i\binom{k-r}{i} = 0$ azonosság helyett a szita formula bizonyítása során már igazolt a $\sum\limits_{i=0}^{s}(-1)^i\binom{k-r}{i} \ge 0$, ha $0 \le s \le k-r$ és s páros szám, és $\sum\limits_{i=0}^{s}(-1)^i\binom{k-r}{i} \le 0$, ha $0 \le s \le k-r$ és s páratlan szám egyenlőtlenségeket kell használni.

Bebizonyítok egy hasonló azonosságot, amelyben a $P(N \ge r)$ valószínűséget számoljuk ki az S_k mennyiségek segítségével, azaz annak a valószínűségét, hogy legalább r A_j esemény következett be. Megmutatom, hogy ez a formula egyszerűen levezethető az előző eredményben bizonyított a $P(N = r)$ valószínűséget kifejező azonosságból.

Formula a $P(N \ge r)$ valószínűség kifejezésére. *A szita formula előbb megfogalmazott általánosításában bevezetett jelöléseket alkalmazva felírhatjuk a*

$$P(N \ge r) = \sum_{i=0}^{n-r}(-1)^i\binom{i+r-1}{i}S_{i+r}, \quad \textit{minden } 1 \le r \le n \textit{ számra},$$

azonosságot, ahol

$$S_0 = 1, \quad S_k = \sum_{1 \le i_1 < \cdots < i_k \le n} P\left(A_{i_1} \cap \cdots \cap A_{i_k}\right), \quad 1 \le k \le n.$$

A formula bizonyítása. A képlet $r = n$-re érvényes, mert $P(N \ge n) = S_n$, és az azonosság jobboldalán álló kifejezés is ezzel egyenlő az $r = n$ esetben. (Ekkor az összeg csak az $i = 0$ tagot tartalmazza.) Az azonosságot r szerinti backward indukcióval és a $P(N = r)$ kifejezésre már bizonyított azonosság segítségével bizonyítjuk be. Ugyanis, ha az azonosságot tudjuk $r+1$-re, akkor felírhatjuk a

$$\begin{aligned}P(N \ge r) &= P(N = r) + P(N \ge r+1)\\ &= \sum_{i=0}^{n-r}(-1)^i\binom{i+r}{i}S_{i+r} + \sum_{i=0}^{n-r-1}(-1)^i\binom{i+r}{i}S_{i+r+1}\end{aligned}$$

azonosságot. Mivel $\sum\limits_{i=0}^{n-r-1}(-1)^i\binom{i+r}{i}S_{i+r+1} = \sum\limits_{i=1}^{n-r}(-1)^{i-1}\binom{i+r-1}{i-1}S_{i+r}$, innen azt kapjuk, hogy

$$\begin{aligned}P(N\ge r) &= \sum_{i=1}^{n-r}(-1)^i\left(\binom{i+r}{i}-\binom{i+r-1}{i-1}\right)S_{i+r}+S_r\\ &= \sum_{i=1}^{n-r}(-1)^i\binom{i+r-1}{i}S_{i+r}+S_r,\end{aligned}$$

mivel $\binom{i+r}{i} = \binom{i-r-1}{i-1} + \binom{i+r-1}{i}$, és ezt az azonosságot kellett bizonyítani.

Feladat:

Tekintsük az előadás (és a jegyzet) elején megfogalmazott példát, és számoljuk ki az általánosított szita formula segítségével annak valószínűségét, hogy pontosan r, $r\ge 0$, házaspár táncol együtt. Mi e valószínűség határértéke, ha $n\to\infty$?

Megoldás: Az eredeti feladat megoldása során kiszámoltuk, hogy $S_k = \frac{1}{k!}$. Ezért az általánosított szita formula alapján annak a valószínűsége, hogy az n házaspárból álló társaságban az együtt táncoló házaspárok N_n száma pontosan r

$$P(N_n=r) = \sum_{i=0}^{n-r}(-1)^i\binom{i+r}{i}S_{i+r} = \sum_{i=0}^{n-r}(-1)^i\binom{i+r}{i}\frac{1}{(i+r)!} = \frac{1}{r!}\sum_{i=0}^{n-r}(-1)^i\frac{1}{i!}.$$

Innen

$$\lim_{n\to\infty}P(N=r) = \frac{1}{r!}\sum_{i=0}^{\infty}(-1)^i\frac{1}{i!} = \frac{1}{r!}\frac{1}{e}.$$