A Valószínűségszámítás I. előadássorozat negyedik előadása.
2002 február 19.

Összefoglaló:

Rátértünk a Borel–Cantelli lemma bizonyítására. A bizonyításhoz szükségünk van a következő lemmára.

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* $B = \bigcup_{n=1}^{\infty} B_n$. *Ekkor létezik a* $\lim_{n\to\infty} P(B_n)$ *határérték, és* $\lim_{n\to\infty} P(B_n) = P(B)$.

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* $C = \bigcap_{n=1}^{\infty} C_n$. *Ekkor létezik a* $\lim_{n\to\infty} P(C_n)$ *határérték, és* $\lim_{n\to\infty} P(C_n) = P(C)$.

Megjegyzés: Az itt megfogalmazott tulajdonság (formálisan) kissé általánosabb mint a Fazekas könyv 2.5 Állításában (23. oldal) szereplő hasonló néven nevezett tulajdonság, illetve a második előadás Tétel A állításának első fele. Ennek bizonyítása a második előadás kiegészítésében megtalálható.) Ezt az állítást tartalmazza a Fazekas könyv 29. oldalán szereplő 10. feladat is.

Bizonyítás: Legyen $D_n = B_{n+1} \setminus B_n$. Ekkor a valószínűségi mérték σ-additivitása miatt

$$P(B) = P(B_n) + \sum_{k=n}^{\infty} P(D_k), \tag{$*$}$$

ahonnan egyrészt $P(B_n) \le P(B)$ minden n-re, másrészt $\sum_{k=1}^{\infty} P(D_k) < \infty$. Az utóbbi állításból következik, hogy minden $\varepsilon > 0$ számhoz létezik olyan $n_0 = n_0(\varepsilon)$ szám, hogy $\sum_{k=n}^{\infty} P(D_k) < \varepsilon$, ha $n \ge n_0$. Innen és a $(*)$ formulából következik, hogy $P(B_n) \le P(B) \le P(B_n) + \varepsilon$, ha $n \ge n_0$, tehát igaz a Lemma első állítása.

Ha a második állítás feltételei teljesülnek akkor az $E_n = \Omega \setminus C_n$, $n = 1, 2, \dots$ halmazok növekvő sorozatot alkotnak, melyekre teljesül az $\bigcup_{n=1}^{\infty} E_n = \Omega \setminus \left(\bigcap_{n=1}^{\infty} C_n \right) = E$ reláció. Ezért a Lemma már bebizonyított része szerint $\lim_{n\to\infty} P(E_n) = P(E)$, és $\lim_{n\to\infty} P(C_n) = 1 - \lim_{n\to\infty} P(E_n) = 1 - P(E) = P(C)$.

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

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

Mivel $\sum\limits_{k=1}^{\infty} P(A_k) < \infty$, minden $\varepsilon > 0$ számhoz létezik olyan $n = n(\varepsilon)$ szám, hogy $\sum\limits_{k=n}^{\infty} P(A_k) < \varepsilon$, 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, melyre $0 \le x_n < 1$ 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.

Korábbi feladatokban láttuk hogyan lehet egyszerűen kiszámítani a $A_1 \cup \dots \cup A_n$ alakú események valószínűségét, ha az A_j, $1 \le j \le n$, események függetlenek. Ezek a számolások természetesen kihasználták a tekintett események függetlenségét. A következőben azt tárgyaljuk meg, hogy amennyiben nincs feltétlenül függetlenség a tekintett események között, de ki tudjuk számolni az $A_{j_1} \cap \dots \cap A_{j_s}$ alakú események valószínűségét, akkor egy úgynevezett szita formula segítségével ki tudjuk számítani az $A_1 \cup \dots \cup A_n$ események valószínűségét is. Továbbá megmutatjuk, hogy ez lehetővé teszi érdekes feladatok megoldását. A következő feladatot fogjuk tárgyalni:

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 $P(\overline{A_1 \cup \cdots \cup A_n}) = 1 - P(A_1 \cup \cdots \cup A_n)$ valószínűség érdekel. Vegyük észre, hogy a $P(A_{j_1} \cap \cdots \cap A_{j_k}) = \dfrac{(n-k)!}{n!}$ azonosság igaz. Ugyanis az összes lehetséges párbaállítások száma $n!$, míg az olyan párbaállítások száma, melyben a j_1-ik, j_2-ik, ..., j_k-ik házaspár egy párba kerül $(n-k)!$. Továbbá érvényes a következő, az irodalomban szita-formulának nevezett eredmény, melyet külön fogunk tárgyalni.

Szita formula. *Legyenek $A_1, \dots, A_n$ tetszőleges események egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ekkor*

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

ahol

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

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 n házaspár esetén

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

és ezért

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

A szita formula bizonyítása: 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 $r_s = 1$ vagy $r_s = -1$, $s = 1, \dots, n$ számokra az

$$A(r_1, \dots, r_n) = A^{r_1} \cap \cdots \cap A^{r_n}$$

eseményeket. A bizonyítás gondolata abból áll, hogy a bizonyítandó azonosságnak mind a két oldalát kifejezzük az $A(r_1,\dots,r_n)=A^{r_1}\cap\cdots\cap A^{r_n}$ események $P(A(r_1,\dots,r_n))$ valószínűségeinek lineáris kombinációjaként, és ellenőrizzük, hogy a két kifejezés megegyezik. Vegyük észre, hogy

$$A_1\cup\cdots\cup A_n=\bigcup_{(r_1,\dots,r_n):\ (r_1,\dots,r_n)\neq(-1,\dots,-1)} A(r_1,\dots,r_n),$$

és tetszőleges $A_{j_1}\cap\cdots\cap A_{j_s}$, $1\le j_1<j_2<\cdots<j_s\le n$ alakú halmazra

$$A_{j_1}\cap\cdots\cap A_{j_s}=\bigcup_{(r_1,\dots,r_n):\ r_{j_u}=1,\ 1\le u\le s} A(r_1,\dots,r_n).$$

Továbbá az $A(r_1,\dots,r_n)$ események különböző $(r_1,\dots,r_n)$ paraméterek esetében diszjunktak. Ezért

$$P(A_1\cup\cdots\cup A_n)=\sum_{(r_1,\dots,r_n):\ (r_1,\dots,r_n)\neq(-1,\dots,-1)} P(A(r_1,\dots,r_n)),$$

és tetszőleges $A_{j_1}\cap\cdots\cap A_{j_s}$ alakú halmazra

$$P(A_{j_1}\cap\cdots\cap A_{j_s})=\sum_{(r_1,\dots,r_n):\ r_{j_u}=1,\ 1\le u\le s} P(A(r_1,\dots,r_n)).$$

Ilyen módon a az $A(r_1,\dots,r_n)$ halmazok $P(A(r_1,\dots,r_n))$ valószínűségeinek a lineáris kombinációjaként fejezhető ki a szita-formula két oldalán szereplő kifejezés. Azt kell belátni, hogy a $P(A(r_1,\dots,r_n))$ szám a szitaformula két oldalán szereplő kifejezésben ugyanazzal az együtthatóval szerepel. Az azonosság baloldalán ez az együttható 1, ha $\{r_1,\dots,r_n\}\neq\{-1,\dots,-1\}$, és nulla, ha $\{r_1,\dots,r_n\}=\{-1,\dots,-1\}$. Ha az $(r_1,\dots,r_s)$ halmaz l darab 1 és $n-l$ darab -1 jegyet tartalmaz, akkor ennek valószínűsége a szitaformula jobboldalán ezt $\binom{l}{1}-\binom{l}{2}+\cdots\pm\binom{l}{l}=\sum\limits_{k=1}^{l}(-1)^{k+1}\binom{l}{k}$ együtthatóval szerepel. Ugyanis az S_k összegben ez a kifejezés $\binom{l}{k}$ együtthatóval szerepel. Miért?

Innen a vizsgált együttható $\sum\limits_{k=1}^{l}(-1)^{k+1}\binom{l}{k}=1-\sum\limits_{k=0}^{l}(-1)^k\binom{l}{k}=1-(1-1)^k=1$, ha az $A(r_1,\dots,r_n)$ halmaz $r_1,\dots,r_n$ indexei között $l\neq 0$ 1-es van. A maradék esetben pedig az együttható nulla. Ez azt jelenti, hogy a szita formula két oldal]án szereplő együthatók megegyeznek.

Nem kötelező házi feladat: Lássuk be, hogy a szitaformula jelöléseit használva

$$S_1-S_2=\sum_{j=1}^{n}P(A_j)-\sum_{1\le j<k\le n}P(A_j\cap A_k)\le P(A_1\cup\cdots\cup A_n)\le S_1=\sum_{j=1}^{n}P(A_j)$$

és általában

$$\sum_{k=1}^{2l}(-1)^{k+1}S_k\le P(A_1\cup\cdots\cup A_n)\le\sum_{k=1}^{2l-1}(-1)^{k+1}S_k$$

minden l indexre. (Legyen $S_k=0$, ha $k>n$.)

Diszkrét eloszlású valószínűségi változók és azok függetlensége

Először bevezettük 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, melyek értékeiket egy általános téren veszik fel. (Például bizonyos esetekben érdemes olyan valószínűségi változókat tekinteni, melyek 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, melyek 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.

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 kívánni, 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 hasonlókat mondhatunk. Érdemes még 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 valószínűségi változók és azok eloszlása. *Egy ξ valószínűségi változót egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn diszkrétnek hívunk, ha megadható egy véges vagy megszámlálhatóan végtelen számosságú $x_1, x_2, \cdots,$ halmaz úgy, hogy az $\{\omega, \xi(\omega) = x_n\}$, $n = 1, 2, \dots,$ halmazok teljes eseményrendszert alkotnak. Egy diszkrét valószínűségi változó eloszlását meghatározzák azok az $x_1, x_2, \dots,$ értékek melyeket az felvesz és a*

$p_n = P(\xi = x_n)$ valószínűségek. (Jegyezzük meg, hogy $\sum_n P(\xi = x_n) = 1$.)

Legyenek $\xi_1, \dots, \xi_k$ diszkrét valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyek 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 valószínűségi változók függetlensége. *Legyenek $\xi_1, \dots, \xi_k$ diszkrét valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Azt mondjuk, hogy ezek az 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 valószínűségi változók melyeknek 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, melyek 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, melyben a vizsgált események és valószínűségi változók megjelentek. Ez az oka annak, hogy minket igazából nem maguk a valószínűségi változók, hanem azok (együttes) eloszlása érdekel.

Peladat:

Legyenek $\xi_1, \dots, \xi_k$ független, diszkrét 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.

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, melyre $\chi_A(\omega) = 1$, ha $\omega \in A$, és $\chi_A(\omega) = 0$, ha $\omega \notin A$.*

Ezután megfogalmaztuk a következő gyakorlaton tárgyalandó feladatokat:

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.

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 melybe 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 $P(\xi_j \neq \xi'_j$ ha $j \neq j')$ 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 valószínűlége, 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 $\log \prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right)=\sum_{j=1}^{k-1}\log\left(1-\frac{j}{n}\right)$ 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 $\sum_{j=1}^{k-1}\log\left(1-\frac{j}{n}\right)\sim-\sum_{j=1}^{k-1}\frac{j}{n}=-\frac{(k-1)k}{2n}$, ahonnan $\log\prod_{j=1}^{k(n)-1}\left(1-\frac{j}{n}\right)\to-\frac{\alpha^2}{2}$, ha $n\to\infty$, és $\frac{k(n)}{\sqrt{n}}\to\alpha$. Ez a számolás precizzé tehető, ha felhasználjuk például azt az egyenlőtlenséget, mely szerint

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

ami szintén következik a $\log(1+x)$ Taylor sorfejtéséből. Miért? Innen kapjuk, hogy $1-\prod_{j=1}^{k(n)-1}\left(1-\frac{j}{n}\right)\to1-e^{-\alpha^2/2}$, ha $n\to\infty$, és $\frac{k(n)}{\sqrt{n}}\to\alpha$.

A következőben a valószínűségszámítás egyik legfontosabb fogalmát, a várható értéket definiáltuk, és tárgyaltuk annak legfontosabb tulajdonságait. Egyelőre csak diszkrét valószínűségi változó várható értékével foglalkoztunk.

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

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

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

Idézzünk fel néhány eredményt, melyek 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.

Miért fontos ez a tulajdonság? Vegyük például azt az esetet, ha a ξ valószínűségi változó az összes racionális számot veszi fel. Akkor ezek a racionális számok mivel megszámlálhatóan sokan vannak felsorolhatóak mint egy $x_1, x_2, \dots,$ sorozat, de nincs a felsorolásban egy természetes sorrend. Ezért a várható értéket definiáló összeg csak akkor értelmes, ha az nem függ attól, hogy a ξ valószínűségi változó lehetséges x_k értékeit milyen sorrendben soroljuk fel.

Az abszolut konvergens sorok fenn említett tulajdonsága szerepelt az analízis tananyagban. A bizonyítás megismétlése helyett lássunk inkább egy példát, mely megmutatja, hogy a nem abszolut konvergens sorok másképpen viselkednek. 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.

Ezek után tekintsük a várható érték kiszámítását egy viszonylag egyszerű esetben.

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

$$E\xi = \sum_{k=0}^{100} k \binom{100}{k} 2^{-100} = 100 \sum_{k=0}^{99} \binom{99}{k} 2^{-100} = 50 \sum_{k=0}^{99} \binom{99}{k} \left(\frac{1}{2} \right)^{99}$$

$$= 50 \cdot \left(\frac{1}{2} + \frac{1}{2}\right)^{99} = 50.$$

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, melynek 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) valószínűségi változók, melyeknek 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 $\xi_1, \xi_2, \dots, \xi_k$ (diszkrét) valószínűségi változók, melyeknek létezik várható értékük, és legyenek $c_1, \dots, c_k$ valós számok. Ekkor a $c_1\xi_1 + c_2\xi_2 + \dots + c_k\xi_k$ valószínűségi változónak is létezik várható értéke, és*

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

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 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.

Kiegészítés:

Az előadásban megemlítettük a következő eredményt:

Állítás: *Legyen x_n valós számoknak olyan sorozata, melyre $0 \le x_n < 1$ 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: Heuisztikusan 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 $-\log(1 - x) = x + \dfrac{x^2}{2} + \dfrac{x^3}{3} + \cdots$, 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, melyek 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 $\dfrac{1}{2}\sum\limits_{k=1}^{n} x_k < \sum\limits_{k=1}^{n} -\log(1-x_k) < 2\sum\limits_{k=1}^{n} x_k$ 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.