A Valószínűségszámítás I. előadássorozat negyedik témája.

Diszkrét eloszlású valószínűségi változók, azok függetlensége és várható értéke. Felcserélhető valószínűségi változók.

Először bevezetem 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 $\xi = \xi(\omega)$ egy $\Omega \to R^1$ leképezés. Az a megkötés, hogy a $\xi(\omega)$ függvény legyen mérhető azt jelenti, hogy minden x valós számra az $\{\omega\colon\ \xi(\omega) < x\} \in \mathcal{A}$ feltétel teljesül.*

Bizonyos feladatokban nemcsak valós értékű valószínűségi változókkal dolgozunk. Ezért célszerű olyan valószínűségi változókat is definiálni, amelyek értékeiket egy általános térben veszik fel. (Például bizonyos esetekben érdemes olyan valószínűségi változókat tekinteni, amelyek értéke komplex szám, vagy nem egyetlen szám, hanem egy szám-n-es. Az utóbbi valószínűségi változókat vektor értékűeknek szokták hívni.) Annak érdekében, hogy ezt megtehessük tekintsünk valamilyen X halmazt, és e halmaz kitüntetett részhalmazainak egy $\mathcal{X}$ osztályát, amely halmazok σ-algebrát alkotnak. Egy ilyen $(X, \mathcal{X})$ rendszert mérhető térnek és az $\mathcal{X}$ σ-algebra elemeit mérhető halmazoknak szokás nevezni az irodalomban.

Értékeit egy általános térben felvevő valószínűségi változó fogalma. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező és egy $(X, \mathcal{X})$ mérhető tér (azaz egy X halmaz és annak bizonyos részhalmazaiból álló σ-algebra.) Egy az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált és $\xi = \xi(\omega)$ értékeit az X térben felvevő (mérhető) függvényt X-tér értékű valószínűségi változónak nevezünk. Az, hogy ez a $\xi(\omega)$ függvény mérhető azt jelenti, hogy minden $Y \in \mathcal{X}$ halmazra teljesül az $\{\omega\colon\ \xi(\omega) \in Y\} \in \mathcal{A}$ feltétel.*

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 definiált és értékeit nem feltétlenül a számegyenesen felvevő 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 az első látásra talán mesterkéltnek tűnő feltételt, hogy a valószínűségi változó legyen *mérhető* függvény. Ennek az az oka, hogy természetes az a kívánság, hogy beszélhessünk az $\{\omega\colon\ \xi(\omega) < x\}$ esemény bekövetkezésének a valószínűségéről minden x számra. Viszont csak $\mathcal{A}$-beli halmazoknak definiáltuk a valószínűségét. 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. Látni fogjuk bizonyos itt nem bizonyított mértékelméleti eredmények segítségével, hogy ez a kívánság teljesül.

Az a probléma is felmerül, hogy meg kell értenünk az először bevezetett valószínűségi változó és az utána definiált értékeit egy általános térben felvevő valószínűségi változó definiciója közötti kapcsolatot is. E kapcsolat megértésének az érdekében a következő a mértékelméletben bebizonyított eredményre hívatkozom. Ha $\xi(\omega)$ mérhető valós értékű függvény egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn e fogalom eredeti definiciója értelmében, azaz $\{\omega\colon\ \xi(\omega) < x\} \in \mathcal{A}$ minden x valós számra, akkor az mérhető a következő általánosabb értelemben is. Jelölje $\mathcal{B}$ a számegyenes Borel-mérhető részhalmazainak σ-algebráját, azaz azt a legszűkebb σ-algebrát a számegyenes részhalmazain, amely az összes $\{u\colon\ u < x\}$ alakú halmazt tartalmazza. A $\xi(\omega)$ függvény eredeti értelemben vett mérhetőségéből következik az $\{\omega\colon\ \xi(\omega) \in B\} \in \mathcal{A}$ tulajdonság minden $B \in \mathcal{B}$ (azaz Borel mérhető) halmazra. Ez azt jelenti, hogy egy a valószínűségi változó eredeti definicióját teljesítő $\xi(\omega)$ valós értékű függvény tekinthető speciálisan, mint egy értékeit egy általános térben felvevő valószínűségi változó, ha az $(X, \mathcal{X})$ mérhető teret, mint az $(R, \mathcal{B})$ teret tekintjük, ahol R a számegyenes, és $\mathcal{B}$ a számegyenes Borel halmazainak σ-algebrája.

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.

Diszkrét eloszlású valószínűségi változók és azok eloszlásának a definiciója. *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. (Ebben az esetben a valószínűségi változók definiciójában szereplő $(X, \mathcal{X})$ mérhető teret úgy definiáljuk, mint az X halmazt és az annak összes részhalmazaiból álló $\mathcal{X}$ σ-algebrát.) Egy diszkrét eloszlású valószínűségi változó eloszlását meghatározzák azok az $x_1, x_2, \dots,$ értékek amelyeket az felvesz és az összes $p_n = P(\xi = x_n)$ valószínűség. (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 függetlenségének definiciója. *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, $\dots$, 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 halmazban veszik fel az értékeiket. Lehetséges, hogy ily módon sok nulla valószínűségi eseményt is bevezettünk, de ez nem okoz problémát.

2. megjegyzés. Annak, hogy bevezettük valószínűségi változók eloszlásának a definicióját mélyebb oka van. A valószínűségszámításban olyan problémákkal foglalkozunk, amelyek csak attól függnek, hogy különböző eseményeknek mennyi a valószínűségük, azaz, hogy különböző valószínűségi változók milyen értéket milyen valószínűséggel vesznek fel. Arra vagyunk kiváncsiak, hogy ezen ismereteknek milyen következményei vannak. Viszont a valószínűségszámításban 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

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

Másrészt

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

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 $A_j=\{x_1,x_2,\dots\}$, $j\in\{1,\dots,k\}\setminus\{j_1,\dots,j_s\}$ 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 valószínűségi változók rendszere az előző lemma eredménye szerint. Ezért minden $\{j_1, \dots, j_s\} \subset \{1, \dots, k\}$ indexhalmazra

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

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 (csak nulla vagy 1 értéket fölvevő) $\chi_{A_1}(\omega), \dots, \chi_{A_k}(\omega)$ indikátorfüggvényeknek teljesíteniük kell, ha azok függetlenek.

Ugyancsak hasznos bizonyos esetekben a következő lemma állítása.

Lemma diszkrét eloszlású független valószínűségi változók függvényeinek egy tulajdonságáról. *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

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

és ezt kellett igazolni.

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 $P(\xi_j \ne \xi'_j$ 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

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

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

A (diszkrét eloszlású) valószínűségi változók fogalmának a segítségével meg tudom fogalmazni az egyik az első téma jegyzetében tárgyalt eredmény egy hasznos általánosítását. Tekintettük azt a feladatot, hogy ha egy urnában különböző színű golyók vannak, és azokat visszatevés nélkül egymás után kihúzzuk, akkor mi annak a valószínűsége, hogy egy adott húzásban előírt színű golyót húzunk ki. Most egy olyan lemmát fogalmazok meg, amely hasznos általánosítása az e feladat megoldásában szereplő legfontosabb gondolatnak. Ennek érdekében 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.

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. Az alábbiakban bebizonyítom felcserélhető valószínűségi változók egy (egyszerű) tulajdonságát, majd megmutatom, hogy ez segít általánosítani és jobban megérteni az első téma jegyzetében tárgyalt urnabeli húzásokkal kapcsolatos feladatok megoldását.

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 $\pi(1) = j_1$, $\pi(2) = j_2$, $\dots$, $\pi(k) = j_k$. Felírhatjuk a

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

azonosságot. Ezért a valószínűségi változók felcserélhetősége miatt

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

Az utolsó azonosság bal és jobboldalának összehasonlításából következik a lemma állítása.

Az urnából való visszatevéses húzásról szóló feladat megoldásában ezen lemma eredményét használtuk fel. Tartalmazzon egy urna n fehér és m piros golyót, es húzzuk ki ezen golyókat egymás után visszatevés nélkül. Vezessük be a következő ξ_j, $1 \le j \le n+m$, valószínűségi változókat: $\xi_j = F$, ha a j-ik húzás fehér, és $\xi_j = P$, ha a j-ik húzás piros. Ekkor $\xi_1, \dots, \xi_{n+m}$ felcserélhető valószínűségi változók, és alkalmazható rá a fenti lemma.

Ezért $P(\xi_{j_1} = P, \xi_{j_2} = P) = P(\xi_1 = P, \xi_2 = P)$ minden $1 \le j_1 < j_2 \le n+m$ indexre, azaz annak a valószínűsége, hogy a j_1-ik és j_2-ik húzásban piros golyót húzunk megegyezik annak a valószínűségével, hogy az első és második húzásban piros golyót húzunk. A vizsgált feladat megoldásában ilyen összefüggéseket használtunk fel.

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.

Lássunk egy másik tipusú feladatot, amelynek vizsgálatában hasznos a felcserélhető valószínűségi változók használata. A következő feladat hasonlít a harmadik előadáson tárgyalt 5. es 5a. feladathoz.

Feladat:

Egy szabályos érmét feldobunk egymás után egészen addig, amíg ötödször megjelenik egy fejdobás. Azután a dobássorozatot abbahagyjuk. Lássuk be, hogy annak a valószínűsége, hogy az utolsó dobást kettővel megelőző dobás eredménye írás ugyanannyi, mint annak a valószínűsége, hogy az első dobás eredménye írás. Mennyivel egyenlőek ezek a valószínűségek?

Megoldás. Tekintsük az első $l-1$ dobás értékeit azon feltétel mellett, hogy a k-ik fejdobás az l-ik dobásban jelenik meg. Vegyük észre, hogy ezek valószínűségi változók (eszerint a feltételes eloszlás szerint) felcserélhetőek. Ezért annak a feltételes valószínűsége, hogy az $l-2$-ik dobás eredménye írás, feltéve, hogy a k-ik fejdobás az l-ik dobásban következik be, ugyanannyi, mint annak a valószínűsége, hogy az első dobás eredménye írás ugyanazon feltétel mellett. Mivel ezen két esemény feltételes valószínűsége, ama feltétel mellett, hogy a k-ik fejdobás az l-ik dobásban jelenik meg minden l-re megegyezik, ezért ezen események (feltétel nélküli) valószínűsége is egyenlő. Ez a valószínűség $\frac{1}{2}$.

Házi feladat.

Egy szabályos érmét feldobunk egymás után végtelen sokszor. Rögzítsünk egy k egész számot, és definiáljunk minden $l \ge k$ egész számra egy $1 \le j_1(l) < j_2(l) < \cdots < j_k(l) \le l$ sorozatot. Tekintsük ezt a sorozatot egészen addig, amíg az k-edik fejdobás megjelenik. Ha ez az l-ik dobásban történik akkor tekintsük a $j_1(l)$-ik, $j_2(l)$-ik és $j_k(l)$-ik dobásokat. Lássuk be, hogy annak a valószínűsége, hogy a $j_1(l)$-ik, $j_2(l)$-ik és $j_k(l)$-ik dobások mindegyike fej, megegyezik annak a valószínűségével, hogy az első k dobás fej, ami 2^{-k}-val egyenlő.

Következő témánk a valószínűségszámítás egyik legfontosabb fogalmához, a várható értékhez kapcsolódik. Egyelőre, a jobb megértés érdekében, csak a diszkrét eloszlású valószínűségi változó várható értékét vezetem be. Tárgyalom annak legfontosabb tulajdonságait. Nagyon fontos azon feladatok megoldásának jó megértése, amelyek arról szólnak, hogy hogyan tudjuk valószínűségi változók várható értékét effektíve kiszámítani.

Diszkrét eloszlású valószínűségi változók várható értékének a definiciója. *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. A ξ valószínűséget definiáló összeg úgy is értelmezhető, mint az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált ξ (elemi) függvény Lebesgue integrálja a P mérték szerint. Ez az észrevétel érthetőbbé teszi néhány a várható érték tulajdonságairól szóló eredmény hátterét, illetve azt, hogy hogyan érdemes definiálni a várható értéket az általános esetben.

2. 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.

3. 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* kifejezés 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.

Felidézek néhány eredményt, amelyek megmagyarázzák, miért követeltük meg a várható érték definiciójában azt, 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 $\sum_{n=1}^{\infty} y_{k(n)}$ 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 igaz. Sőt, ebben az esetben a következő negatív jellegű állítás érvényes.

Legyen $\sum_{n=1}^{\infty} x_n$ 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 $\lim_{N\to\infty} \sum_{n=1}^{N} x_{k(n)} = \alpha$.

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 $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$. Tekintsük e sor következő átrendezését: $S = 1 + \sum_{k=1}^{\infty} \left(\frac{1}{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}-1} - \frac{1}{k} \right)$. 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 $E\xi = \sum_{k=0}^{100} k \binom{100}{k} 2^{-100}$ 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

$$E\xi = \sum_{k=0}^{100} k \binom{100}{k} 2^{-100} = \sum_{k=1}^{100} k \binom{100}{k} 2^{-100} = 100 \sum_{k=0}^{99} \binom{99}{k} 2^{-100}$$

$$= 50 \sum_{k=0}^{99} \binom{99}{k} \left(\frac{1}{2}\right)^{99} = 50 \cdot \left(\frac{1}{2} + \frac{1}{2}\right)^{99} = 50.$$

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, amelynek bizonyítását később ismertetem, megmagyarázza, hogy a fenti heurisztika miért adott helyes eredményt, és ennek segítségével hogyan lehet kiszámítani sok esetben egyszerűen valószínűségi változók várható értékét.

Tétel a várható érték additivításáról. *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 $c_1\xi_1 + c_2\xi_2 + \dots + c_k\xi_k$ 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 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 $ES = 100 \cdot 3.5 = 350$.

A várható érték tulajdonságainak bizonyításában nagyon hasznos az alábbi segédtétel.

Segédtétel a várható érték tulajdonságairól. *Legyen ξ egy diszkrét eloszlású valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amely valamilyen $x_1, x_2, \dots,$ értékeket vesz fel, és legyen $B_k = \{\omega\colon\ \xi(\omega) = x_k\}$, $k = 1, 2, \dots,$ a ξ valószínűségi változó értékei által meghatározott teljes eseményrendszer. Legyen C_j, $j = 1, 2, \dots,$ egy ennél finomabb teljes eseményrendszer, azaz tegyük fel, hogy a $C_j \in \mathcal{A}$ halmazok diszjunktak, $\bigcup\limits_{j=1}^{\infty} C_j = \Omega$, és minden C_j halmazhoz tartozik olyan B_k halmaz, amelyre $C_j \subset B_k$. Ekkor $B_k = \bigcup\limits_{j\colon C_j \subset B_k} C_j$ minden k indexre. Ha $C_j \subset B_k$, akkor rendeljük hozzá a C_j halmazhoz az $y_j = x_k$ számot. A $\sum\limits_{k=1}^{\infty} x_k P(B_k)$ és $\sum\limits_{j=1}^{\infty} y_j P(C_j)$ összegek egyszerre abszolut konvergensek, és ha ezek az összegek abszolut konvergensek, akkor*

$$E\xi = \sum_{k=1}^{\infty} x_k P(B_k) = \sum_{j=1}^{\infty} y_j P(C_j).$$

A fenti Segédtétel tartalát a következő módon is interpretálhatjuk. Ha az $E\xi$ várható értéket definálό $E\xi = \sum x_k P(B_k)$, $B_k = \{\omega\colon\ \xi(\omega) = x_k\}$ összegben szereplő B_k halmazokat kisebb C_j darabokra bontjuk, akkor a várható értéket $E\xi = \sum y_j P(C_j)$ alakban is felírhatjuk, ahol y_j a $\xi(\omega)$ függvény értéke a C_j (illetve az őt tartalmazó halmazon. A várható érték ilyen felírása emlékeztet az integrálok definiciójában szereplő integrálközelítő összegekre.

Valójában itt sokkal többről van szó, mint puszta formális analógiáról. A mértékelméletben bevezették az úgynevezett Lebesgue integrál fogalmát, ami általánosabb, mint a Riemann integrál. A várható érték, amiről beszélünk, a $\xi(\omega)$ valószínűségi változónak, azaz az $(\Omega, \mathcal{A}, P)$ téren definiált (mérhető) függvénynek a Lebesgue integrálja a P mérték szerint. Erre a tényre a későbbiekben nem lesz szükségünk, de sok meggondolás hátterében ennek ismerete rejtőzik. Így például a várható érték additívitása az integrál additívitásának az átfogalmazása. Bár itt nem foglalkozunk a Lebesgue integrál fogalmának kidolgozásával, mert erre nincs szükségünk, valójában az ehhez szükséges munka jelentős részét megtesszük. Diszkrét eloszlású valószínűségi változók várható érték definiciójának bevezetésével az egyszerű (elemi) függvények Lebesgue integrálját definiáljuk, és a várható érték tulajdonságainak a bizonyításával a Lebesgue integrál legfontosabb tulajdonságait bizonyítjuk.

A Segédtétel bizonyítása: Először azt mutatjuk meg, hogy a $\sum\limits_{k=1}^{\infty} x_k P(B_k)$ és $\sum\limits_{j=1}^{\infty} y_j P(C_j)$ összegek egyszerre abszolut konvergensek.

Ha a $\sum\limits_{k=1}^{\infty} x_k P(B_k)$ összeg abszolut konvergens, akkor létezik olyan $L < \infty$ szám, amelyre $\sum\limits_{k=1}^{\infty} |x_k| P(B_k) = L$, és ezért minden N indexre

$$\sum_{j=1}^{N} |y_j| P(C_j) \le \sum_{k=1}^{\infty} |x_k| P(B_k) = L.$$

Innen következik, hogy $\sum\limits_{j=1}^{\infty} |y_j| P(C_j) \le L$, azaz a $\sum\limits_{j=1}^{\infty} y_j P(C_j)$ összeg is abszolut konvergens.

Ha a $\sum\limits_{j=1}^{\infty} y_j P(C_j)$ összeg abszolut konvergens, akkor $\sum\limits_{j=1}^{\infty} |y_j| P(C_j) \le L < \infty$ alkalmas L számmal, ahonnan tetszőleges N indexre

$$\sum_{k=1}^{N} |x_k| P(B_k) = \sum_{k=1}^{N} \sum_{j\colon\, C_j \subset B_k}^{\infty} |y_j| P(C_j) \le \sum_{j=1}^{\infty} |y_j| P(C_j) \le L < \infty,$$

és ezért a $\sum\limits_{k=1}^{\infty} x_k P(B_k)$ összeg is abszolut konvergens. Ezzel beláttuk, hogy a két tekintett összeg egyszerre abszolut konvergens.

Ha a fenti sorok abszolut konvergensek akkor felírhatjuk minden k indexre az

$$x_k P(B_k) = \sum_{j\colon\, C_j \subset B_k} y_j P(C_j)$$

azonosságot. Ezeket az azonosságokat összegezve minden k-ra, és felhasználva azt, hogy egy abszolut konvergens sor tagjai tetszőleges átrendezés után ugyanoda konvergál kapjuk, hogy

$$E\xi = \sum_{k=1}^{N} x_k P(B_k) = \sum_{k=1}^{N} \sum_{j\colon\, C_j \subset B_k}^{\infty} y_j P(C_j) = \sum_{j=1}^{\infty} y_j P(C_j),$$

és ezt kellett bizonyítanunk.

A várható érték additívitását kifejező $E(\xi_1+\xi_2)=E\xi_1+E\xi_2$ azonosság bizonyítása. Az ξ_1 és ξ_2 valószínűségi változók vegyék fel az $x_1, x_2, \dots$ értékeket. Vezessük be a $C(j,k)=\{\omega\colon\ \xi_1(\omega)=x_j, \xi_2(\omega)=x_k\}$, $j,k=1,2,\dots$, eseményeket. A Segédtételt alkalmazva a $B(k)=\{\omega\colon\ \xi_1(\omega)=x_k)$ és az előbb bevezetett $C(j,k)$ halmazokkal kapjuk, hogy

$$E\xi_1=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}x_jP(C(j,k)).$$

Hasonlóan a Segédtétel alkalmazása a $\tilde{B}(k)=\{\omega\colon\ \xi_2(\omega)=x_k)$ és a $C(j,k)$ halmazokkal azt adja, hogy

$$E\xi_2=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}x_kP(C(j,k)),$$

E két azonosságot összeadva, kapjuk, hogy

$$E\xi_1+E\xi_2=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(x_j+x_k)P(C(j,k)).$$

Másrészt, ismét a Segédtétel alapján (a $B(j,k)=\{\omega\colon\ \xi_1(\omega)+\xi_2(\omega)=x_j+x_k)\}$ és $C(j,k)$ halmazokkal) érvényes az

$$E(\xi_1+\xi_2)=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(x_j+x_k)P(C(j,k))$$

reláció is. Ezekből az összefüggésekből következik az $E\xi_1+E\xi_2=E(\xi_1+\xi_2)$ azonosság.

Egy ξ valószínűségi változó valamilyen $g(\xi)$ függvényének $Eg(\xi)$ várható értékét kiszámolhatnánk a várható érték eredeti definiciójának a felhasználásával úgy, hogy először kiszámoljuk a $g(\xi)$ valószínűségi változó eloszlását. De sokszor egyszerűbb a várható érték kiszámolását az alábbi egyszerű tétel segítségével végezni.

Tétel diszkrét eloszlású valószínűségi változó függvényének a várható értékéről. *Legyen ξ diszkrét eloszlású valószínűségi változó, amely bizonyos $x_1, x_2, \dots$ értéket vesz fel, és legyen $g(x)$ az $X=\{x_1,x_2,\dots\}$ halmazon definiált valós értékű függvény. Ekkor*

$$Eg(\xi)=\sum_{j=1}^{\infty}g(x_j)P(\xi=x_j),$$

feltéve, hogy a $\sum\limits_{j=1}^{\infty}g(x_j)P(\xi=x_j)$ összeg abszolut konvergens.

Bizonyítás: Vezessük be a $C_j=\{\omega\colon\ \xi(\omega)=x_j\}$ halmazokat és $y_j=g(x_j)$ számokat, $j=1,2,\dots$. Ekkor alkalmazva a Segédtételt a $B_k=\{\omega\colon\ g(\xi(\omega))=x_k\}$ és a fenti C_j halmazokkal az $\eta=g(\xi)$ valószínűségi változóra megkapjuk a Tétel állítását.

A következő eredmény független valószínűségi változók szorzatának a várható értékéről szól.

Tétel független valószínűségi változók szorzatáról. *Legyenek* $\xi_1, \xi_2, \dots, \xi_n$ *független, diszkrét eloszlású valószínűségi változók, amelyek mindegyikére létezik az* $E\xi_j$, $1 \le j \le n$ *várható érték. Ekkor az* $E\xi_1\xi_2\cdots\xi_n$ *várható érték is létezik, és*

$$E\xi_1\xi_2\cdots\xi_n = E\xi_1 E\xi_2\cdots E\xi_n.$$

Megjegyés: A fent megfogalmazott eredmény rendkívül fontos tulajdonsága a *független* valószínűségi változóknak.

Bizonyítás: Tegyük fel, hogy a $\xi_1, \dots, \xi_n$ valószínűségi változók $x_1, x_2, \dots,$ értékeket vesznek fel. Ekkor

$$\begin{aligned} E\xi_1\cdots E\xi_n &= \sum_{j_1=1}^{\infty} x_{j_1} P(\xi_1 = x_{j_1})\cdots \sum_{j_n=1}^{\infty} x_{j_n} P(\xi_n = x_{j_n}) \\ &= \sum_{j_1=1}^{\infty}\cdots\sum_{j_n=1}^{\infty} x_{j_1}\cdots x_{j_n} P(\xi_1 = x_{j_1})\cdots P(\xi_n = x_{j_n}) \\ &= \sum_{j_1=1}^{\infty}\cdots\sum_{j_n=1}^{\infty} x_{j_1}\cdots x_{j_n} P(\xi_1 = x_{j_1}, \cdots, \xi_n = x_{j_n}), \end{aligned}$$

a ξ_j, $1 \le j \le n$ valószínűségi változók függetlensége miatt, valamint azért mert a fenti sorok mindegyike abszolut konvergens. (Ugyanis abszolut konvergens sorok szorzata is abszolut konvergens. Ez azt jelenti, hogy ha egy $a(i)$ és $b(j)$, $1 \le i, j < \infty$, sorozat abszolut konvergens, akkor a $c(i,j) = a(i)b(j)$ sorozat is abszolut konvergens.) Viszont a Segédtétel alapján a fenti azonosság jobboldalán szereplő kifejezés $E\xi_1\xi_2\cdots\xi_n$, ezért

$$E\xi_1\xi_2\cdots\xi_n = E\xi_1 E\xi_2\cdots E\xi_n.$$

A fenti eredmények alkalmazásaként megoldjuk a következő feladatot.

Feladat:

1.) Egy szabályos dobókockát feldobunk tízszer. Számoljuk ki a dobásösszeg harmadik hatványának a várható értékét.

Megoldás: Vezessük be a következő valószínűségi változókat: $\xi_j(\omega) = k$, ha a j-ik dobás eredménye k, $1 \le j \le 10$, $1 \le k \le 6$. Ekkor az $E\left(\sum\limits_{j=1}^{10}\xi_j(\omega)\right)^3$ várható értéket kell kiszámítanunk. Ennek érdekében tekintsük a $\left(\sum\limits_{j=1}^{10}\xi_j(\omega)\right)^3$

kifejezést és értsük meg milyen tagokat kapunk, ha elvégezzük a beszorzásokat. Egyrészt megjelenik 10 darab ξ_j^3 alakú kifejezés, és $E\xi_j^3 = E\xi_1^3$ minden ilyen tagra. Ezenkívül megjelenik $3 \cdot 10 \cdot 9$ darab $\xi_j^2\xi_k$, $j \neq k$, alakú kifejezés, mert a lehetséges (j,k) párokat $10 \cdot 9$ módon választhatjuk ki, és a k (a négyzetre nem emelt tényező) három helyen szerepelhet a szorzatban. (Tehát például a $\xi_1\xi_2^2$ alakú tagnak 3 lesz az együtthatója a szorzatban.) Továbbá minden ilyen tagra $E\xi_j^2\xi_k = E\xi_j^2 E\xi_k = E\xi_1^2 E\xi_1$. Továbbá, hasonló meggondolások alapján láthatjuk, hogy $10 \cdot 9 \cdot 8$ módon jelenhet meg $\xi_j\xi_k\xi_l$ alakú tag, ahol a j, k és l indexek mind különbözőek, és ezekre $E\xi_j\xi_k\xi_l = (E\xi_1)^3$. Valóban, a $\xi_j\xi_k\xi_l$ alakú alakú tagok összeszámlálásánál vegyük észre, hogy $1 \le j < k < l \le 10$ alakú számhármasokat $\binom{10}{3}$ féleképp választhatunk, a ξ_j tényező a szorzatban 3-féleképp jelenhet meg, a szorzat első, második vagy harmadik tagjában, a ξ_k tényező ezután 2-féleképp, a ξ_l tényező pedig egyféleképp választható. Másfajta tag nem jelenikmeg a szorzatban. Innen a várható érték additívitását kihasználva azt kapjuk, hogy $E\left(\sum\limits_{j=1}^{10}\xi_j(\omega)\right)^3 =$ $10E\xi_1^3 + 270E\xi_1 E(\xi_1^2) + 720\left(E\xi_1\right)^3 = 735 + 14332.5 + 30870 = 45937.5$, mert $E\xi_1 = 3.5$, $E\xi_1^2 = \frac{91}{6}$ és $E\xi_1^3 = 73.5$.