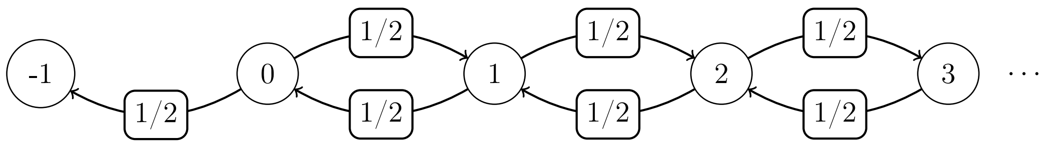
Valószínűségszámítás 1 gyakorlat, 3. feladatsor, megoldások

1. Alapfeladatok

1. A számegyenes 0 pontjából egy bolha indul, minden másodpercben véletlenszerűen ugrik jobbra vagy balra egyet. Mekkora a valószínűsége, hogy valaha ráugrik a (-1)-re?

   Első megoldás. Rajzoljunk Markov láncot:

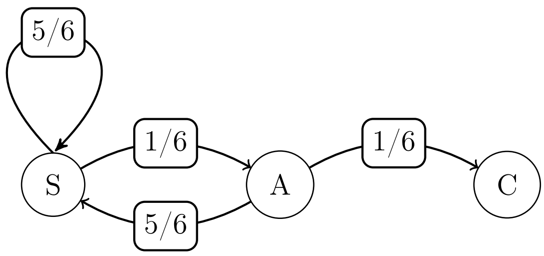
   Jelölje p_i annak a valószínűségét, hogy az i-ből indulvavalaha rálépünk a -1-re. Ekkor persze $p_{-1} = 1$. Mit tudunk mondani a p_i-ről, ha $i \geq 0$? Az i csúcsból indulva $1/2$ eséllyel lépünk balra, ahonnan p_{i-1} annak a valószínűsége, hogy valaha rálépünk a (-1)-re. Illetve $1/2$ eséllyel lépünk jobbra, ahonnan ugyanez p_{i+1} valószínűséggel történik. Vagyis $p_i = \frac{1}{2} \cdot p_{i-1} + \frac{1}{2} \cdot p_{i+1}$.

   Vagyis a p_i egy végtelen hosszú számtani sorozat, ami végig 0 és 1 között marad. Ez csak úgy lehetséges, ha konstans. Mivel $p_{-1} = 1$, ezért $p_i = 1$ minden i-re. Vagyis bárhonnan indulva 1 valószínűséggel rálépünk a (-1)-re.

   Második megoldás. Egyszerűbb egyenlethez jutunk, ha észrevesszük, hogy annak a valószínűsége, hogy 1-ből eljutunk a -1-be, pont annak a két valószínűségnek a szorzata, hogy 1-bl eljutunk a 0-ba, és hogy 0-ból eljutunk a -1-be. Ez utóbbi két valószínűsége mindkettő p_0: az első definíció szerint, míg a második az egyenes (és a véletlen bolyongás) eltolási szimmetriája miatt. Vagyis $p_1 = p_0^2$, így a fenti egyenletek közül elég az elsőt felírni: $p_0 = 1/2 + 1/2 \cdot p_0^2$. Ez ugyan másodfokú, de szerencsénkre csak egyetlen gyöke van, a $p_0 = 1$.

2. Egy dobókockával addig dobálunk, amíg kétszer egymás után hatost nem dobunk. Mennyi a dobásaink számának várható értéke?

   Megoldás. Készítsünk ábrát.

   Mint korábban, jelöljük x_i-vel az i állapotból való szükséges lépések várható számát, az alábbi egyenleteket kapjuk:

$$\begin{aligned} x_S &= 1 + \frac{5}{6} \cdot x_S + \frac{1}{6} \cdot x_A, \\ x_A &= 1 + \frac{5}{6} \cdot x_S + \frac{1}{6} \cdot 0. \end{aligned}$$

   Ezeket megoldva kapjuk, hogy $x_S = 42$.

3. Egy vizsgaidőszak előtt az évfolyam 50 egyformán felkészült hallgatóból áll. 5 tantárgyból vizsgázik mindenki, melyeken a (galád) vizsgáztatók 5-5 embert (teljesen véletlenszerűen) megbuktatnak, egy hallgatót esetleg többen is. Mennyi a sikeresen vizsgázó (vagyis egy tantárgyból sem bukó) hallgatók számának várható értéke?

   Intuitív megoldás: Képzeljük azt, hogy azok a diákok, akik egy vizsgán megbuknak kiesnek, és vizsgáljuk a következő vizsgán csak azokat, akik még versenyben vannak. Az első vizsga után a bent maradtak (várható) értéke $45 = (9/10) \cdot 50$. A második vizsgán 5 diák fog megbukni, de ezek véletlenszerűen kerülnek kisorsolásra, és a már kiesettek is bukhatnak még egyszer. Vagyis ez a kör is várhatóan megtizedeli a versenyzőket, két kör utána a bent maradtak várható értéke $45 \cdot (9/10) = 50 \cdot (9/10)^2$. És így tovább, minden körben várhatóan megtizedelődnek, így öt vizsga utána várható bent maradtak várható értéke $50 \cdot (9/10)^5$.

Pontos megoldás: Mekkora valószínűséggel fog egy adott diák mind az öt tantárgyból sikeresen vizsgázni? Mindből $1/10$ valószínűséggel bukik meg, tehát $(9/10)^5$ valószínűséggel fog sikeresen vizsgázni. Mivel 50 diák van, ezért várhatóan $50 \cdot (9/10)^5$ fog közülük sikeresen vizsgázni minden tantárgyból.

Ez még úgy is igaz, hogy az egyes diákok bukásai nem függetlenek egymástól, mert tudjuk, hogy minden tantárgyból összesen 5-en buknak, nem pedig mindenki egymástól függetlenül $1/10$ valószínűséggel. A várható érték nem csak akkor adódik össze, ha független valószínűségi változókat adunk össze. Kicsit precízebben fogalmazva legyen X a sikeresen vizsgázó diákok száma. Legyen továbbá I_i annak az indikátora hogy az i-edik diák mindenből átment. Vagyis ha az i-edik diák mindenből átment, akkor $I_i = 1$, ha valamiből bukott, akkor $I_i = 0$.

Persze $X = I_1 + I_2 + \ldots + I_{50}$, ami azt fejezi ki, hogy bárhogy történjék a véletlenség a feladatban, a kifejezés két oldalán lévő véletlen számok egyenlők. Ekkor persze a várható értékük is ugyanaz, vagyis

$$\mathbb{E}(X) = \mathbb{E}(I_1 + \ldots + I_{50}).$$

A várható érték additivitása miatt

$$\mathbb{E}(X) = \mathbb{E}(I_1) + \ldots + \mathbb{E}(I_{50}) = 50 \cdot \mathbb{E}(I_1) = 50 \cdot \left(\frac{9}{10}\right)^5.$$

4. Egy egység sugarú körlapban véletlenszerűen választunk egy pontot. Mekkora a pont középponttól való távolságának várható értéke?

Intuitív megoldás. Ennek a valószínűségi változónak (véletlen pontnak a középponttól mért távolsága) az értékkészlete $[0, 1]$, de a változó nem egyenletes eloszlású, a *sűrűségfüggvénye* az $s(x) = 2x$ függvény. Ez kifejezi, hogy az egyes értékek egymáshoz viszonyított súlya lineárisan nő, ahogy növeljük az értéket (és ezzel a körvonal sugarát). Azért $2x$, hogy az integrál 1 legyen, ez felel meg annak, hogy az összes eset együttes valószínűsége 1. Vagyis a várható értéket kiintegrálva azt kapjuk, hogy

$$\int_{[0,1]} x \cdot 2x \; dx = [(2/3)x^3]_0^1 = 2/3.$$

Tehát a keresett várható érték $2/3$.

Profi megoldás. Legyen Z a középponttól való távolság. Ekkor $0 \le Z \le 1$, így a továbbiakban csak erre az intervallumra szorítkozunk (tehát az eloszlásfüggvény a $(-\infty, 0)$ intervallumon 0, az $(1, \infty)$ intervallumon 1).

$$F(r) = \mathbb{P}(Z < r) = \frac{r^2\pi}{1^2\pi} = r^2,$$

ebből deriválással adódik, hogy

$$f(r) = F'(r) = 2r\mathbb{I}(0 \le r \le 1),$$

és könnyen ellenőrizhető, hogy ez valóban sűrűségfüggvény:

$$F(r) = \int_{-\infty}^{r} f(x)\, dx$$

minden r-re.

$$\mathbb{E}Z = \int_0^1 r \cdot 2r \; \mathrm{d}r = \left[\frac{2r^3}{3}\right]_0^1 = \frac{2}{3}.$$

5. Aladár megrögzötten gyűjti a Spar ajándék matricáit, minden 2000 forint vásárlás után ingyen kap 5 matricát. Szeretné kigyűjteni a teljes albumot, amihez szüksége lesz mind a 200 féle matricára. Tegyük fel, hogy minden matrica amit kap véletlenszerű. Várhatóan hány matricát kell összegyűjtenie, hogy mind a 200 féléből legyen neki legalább egy? Mennyi pénzt kell ehhez a Sparban elköltenie?

   Megoldás. Bontsuk a várakozást szakaszokra aszerint, hogy épp hányféle matricája van már Aladárnak. Amikor egy sem, akkor a következő matrica biztosan új fajta lesz, így csak 1 véletlen matricára van szükség. Ha már van egyféle matricája, akkor (199/200) valószínűséggel kap újat. Mint már korábban láttuk, ekkor várhatóan (200/199) húzásra van szüksége, hogy újfajta matricát találjon. Ugyanígy, ha már i-féle matricája van, akkor $(200-i)/200$ valószínűséggel húz újat, tehát várhatóan $200/(200-i)$ húzés múlva lesz $(i+1)$-féle matricája.

   Tehát a teljes várható érték:

$$1 + \frac{200}{199} + \frac{200}{198} + \ldots + \frac{200}{1} = 200 \cdot \left(\frac{1}{200} + \frac{1}{199} + \ldots + \frac{1}{2} + 1 \right) \approx 200 \cdot \log(201).$$

   Vagyis Aladárnak várhatóan kb. $200 \cdot \log(201) = 1060.66$ matricát kell szereznie, hogy kigyűljön az album, ami $(1060/5) \cdot 2000 = 424000$ forint környékén mozog.

6. Egy baktériumfajta minden példánya időegységenként p valószínűséggel kettéosztódik, $1-p$ valószínűséggel elpusztul. Kezdetben egy baktériumunk van. Mekkora a valószínűsége, hogy a baktériumfaj kihal?

   Megoldás. Jelölje q annak a valószínűségét, hogy a baktériumfaj kihal. Ha az első baktérium elpusztul, akkor persze kihalt a faj. Ha nem, akkor 2 baktériumunk lesz, amiknek a leszármazottai egymástól függetlenül q valószínűséggel fognak kihalni. Ebből az alábbi egyenletet kapjuk:

$$q = (1-p) \cdot 1 + p \cdot q^2 \quad \Rightarrow \quad 0 = p \cdot q^2 - q + (1-p).$$

$$q = \frac{1 \pm \sqrt{1-4p(1-p)}}{2p} = \frac{1 \pm \sqrt{(1-2p)^2}}{2p} = \frac{1 \pm |1-2p|}{2p} = \frac{1 \pm (1-2p)}{2p}.$$

$$q = \frac{2-2p}{2p} = \frac{1-p}{p}, \text{ vagy } q = \frac{2p}{2p} = 1.$$

   Az abszolút érték jelet azért lehet elhagyni, mert amúgy is $\pm$ jel van előtte. Viszont ennél sokkal fontosabb kérdés, hogy akkor melyik a valódi megoldás a kettő közül?

   Ha $p < 1/2$, akkor $(1-p)/p > 1$, tehát nem lehet ez a megoldás. Vagyis ebben az esetben a megoldás $q = 1$.

   Ha $p = 1/2$, akkor a két megoldás egybeesik, tehát itt sincs gond. De mi a helyzet, ha $p > 1/2$, ekkor mindkét megoldás hihető? Ezt nem fogjuk tudni precízen megválaszolni, de mégis, melyiket érezzük igaznak?

   Két érvelést is találhatunk. Először is, mennyi az n-edik generáció méretének várható értéke? Az nulladik generációé 1, az elsőé $(1-p) \cdot 0 + p \cdot 2 = 2p$. Ez pont azt fogalmazza meg, hogy egy baktériumnak a következő generációban várhatóan $2p$ leszármazottja lesz. Mivel az elsőben várhatóan $2p$ baktérium van, ezért a másodikban várhatóan $(2p)^2$ baktérium lesz. És így tovább, az n-edikben $(2p)^n$. A mi esetünkben $2p > 1$, tehát a baktériumok várható száma exponenciálisan nő. Ami érzésre ellentmond a $q = 1$ megoldással, hiszem az azt jelentené, hogy egy valószínűséggel kihal a baktériumfaj. Ez az érvelés matematikailag helytelen, mert az n-edik generáció várható értéke nőhetne attól is, hogy ugyan nagy valószínűséggel addigra kihal a baktérium, de nagyon kis valószínűséggel nagyon-nagyon-nagyon sok baktérium lesz.

   A másik érvelés folytonosságra épül: mi a helyzet, ha $p = 1$? Ekkor világos, hogy a baktérium 0 valószínűséggel hal ki, tehát ebben az esetben a két megoldás közül az $(1-p)/p$ érvényesül. Ábrázoljuk a $f(p) = 1$ és $g(p) = (1-p)/p$ függvényeket a $[0,1]$ intervallumon. Egy adott p-re a q a két függvényérték közül lehet valamelyik. Azt már tudjuk, hogy eleinte az f érvényesül. A $p = 1/2$ pontban egybeesnek, majd a végén, a $p = 1$-ben már a g érvényesül. Ha elhisszük, hogy a feladat megoldása (q) a paraméter (p) változtatására nézve folytonos, akkor minden $p > 1/2$ esetben a g függvényt kell választanunk!

   Mindkét érvelés arra utal, hogy $p > 1/2$ esetben az $(1-p)/p$ a helyes megoldás. Ez így is van, de nem fogjuk bizonyítani.

2. Ismerd fel a nevezetes eloszlást!

1. Tegyük fel, hogy az új internetelőfizetők mindegyike a többiektől függetlenül 20% valószínűséggel speciális kedvezményt kap. Mennyi a valószínűsége, hogy 10 ismerősünk közül, akik most fizettek elő, legalább négyen részesülnek a kedvezményben?

   Megoldás. Bernoulli, $n = 10$, $p = 0,2$.

   $$\mathbb{P}(X \geq 4) = 1 - \binom{10}{0}(0,2)^0 \cdot (0,8)^{10} - \binom{10}{1}(0,2)^1 \cdot (0,8)^9 - \binom{10}{2}(0,2)^2 \cdot (0,8)^8 - \binom{10}{3}(0,2)^3 \cdot (0,8)^7.$$

2. Négy szabályos dobókockával dobunk sokszor egymás után addig, amíg elő nem fordul, hogy a négy dobásból legalább három hatos. Jelölje Y, hogy hányszor kell dobni ehhez. Adjuk meg Y eloszlását.

   Megoldás. Jelölje p a siker valószínűségét, vagyis (Bernoulli-val számolva):

   $$p = \binom{4}{3}\left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right) + \binom{4}{4}\left(\frac{1}{6}\right)^4.$$

   Maga az Y geometriai eloszlású, p paraméterrel.

3. Pisti minden nap leül felesezni. Először megiszik egy pohárral, majd ezután minden pohár után (egymástól függetlenül) 0.4 valószínűséggel iszik még egyet. Mi az általa hetente megivott felesek számának eloszlása? (A napi mennyiségek is függetlenek egymástól.)

   Megoldás. Minden nap geometriai, 0.4 paraméterrel, így összesen negatív binomiális, $n = 7$, $p = 0.4$.

4. Egy forgalmas útszakaszon azt figyelik, hogy öt perc alatt hány autó lépi át a megengedett sebességhatárt. A tapasztalatok alapján feltételezzük, hogy annak valószínűsége, hogy van ilyen autó, ugyanannyi, mint annak, hogy nincs. Milyen eloszlást használhatunk (további információ hiányában) a gyorshajtók számának modellezésére? Mennyi a valószínűsége, hogy pontosan három autó lépi át a megengedett sebességhatárt öt perc alatt?

   Megoldás. Poisson λ paraméterrel, amit a $\mathbb{P}[X = 0] = e^{-\lambda} = 1/2$ egyenletből kapunk: $\lambda = \log 2$. Ezek után $\mathbb{P}[X = 3] = (1/2)\frac{\log(2)^3}{6}$

5. Válasszunk egy számot 1 és n között, mindegyiket azonos valószínűséggel. Mennyi a kiválasztott szám várható értéke?

   Megoldás. Diszkrét egyenletes, $\mathbb{E}(X) = (n + 1)/2$.

6. A francia labdarúgó-válogatott keretében huszonhárom játékos van, közülük hárman kapusok. A szövetségi kapitány a holnapi mérkőzésre a 11 kezdőjátékost teljesen véletlenszerűen választja ki. Mennyi a valószínűsége, hogy pontosan kettő kapus fog szerepelni?

   Megoldás. Hipergeometrikus eloszlás, $N = 23$, $K = 3$, $n = 11$.

   $$\mathbb{P}(X = 2) = \frac{\binom{3}{2}\binom{20}{9}}{\binom{23}{11}}.$$

3. Gyakorló feladatok

1. Egy tízemeletes ház földszintjén 15 ember száll be a liftbe. Egymástól függetlenül mindenki választ a 10 emelet közül egyet (mindegyiket azonos valószínűséggel), ahol kiszáll. Várhatóan hány emeleten áll meg a lift? Csak a kiszállásokat vesszük figyelembe, vagyis a lift ott áll meg, ahol legalább egy ember kiszáll.

   Megoldás. Legyen X_j annak indikátora, hogy a j. emeleten megáll a lift ($j = 1, 2, \ldots, 10$). Ekkor a kérdés valójában az X_j-k összegének várható értéke, ami a várható érték additív tulajdonsága alapján, illetve az alapján, hogy egy emeleten akkor nem áll meg a lift, ha mindenki a többi 9 emelet valamelyikét választja, és ezek a választások egymástól függetlenek:

$$\mathbb{E}\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} \mathbb{E}(X_j) = 10 \cdot 1 \cdot \mathbb{P}(X_1 = 1) = 10 \cdot \left(1 - \left(\frac{9}{10}\right)^{15}\right) = 7,94.$$

Az X_j indikátorok nem függetlenek, de a várható érték linearitásánál nem is kell semmilyen feltétel a várható érték létezésén kívül.

2. Öt házaspár leül véletlenszerű sorrendben egy kerek asztal köré (minden elhelyezkedés egyformán valószínű). Mennyi az olyan házaspárok számának várható értéke, akik egymás mellett ülnek?

   Megoldás. Legyen X_j annak indikátora, hogy a j. házaspár egymás mellé ül, X pedig az összes, egymás mellé kerülő házaspár száma. Ekkor a várható érték additív tulajdonsága alapján

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{j=1}^{5} X_j\right) = \sum_{j=1}^{5} \mathbb{E}(X_j) = 5 \cdot \mathbb{P}(X_1 = 1) = 5 \cdot \frac{2}{9},$$

   hiszen mind az öt házaspár $2/9$ valószínűséggel kerül egymás mellé: akárhová ül a pár egyik tagja, a pár másik tagja a helyek szimmetriája miatt a fennmaradó 9 hely mindegyikére azonos valószínűséggel kerül, és ezek közül 2 jó.

3. Egy kalapban 9 cédula van, ezeken az $1, 2, \ldots, 9$ számok szerepelnek. Addig húzunk visszatevéssel, amíg 5-nél nagyobb számot nem kapunk (minden húzásnál minden cédulát azonos valószínűséggel választunk). Mennyi a kihúzott számok összegének a várható értéke?

   Megoldás. Legyen x a keresett várható érték. A szokásos Markov láncos meggondolásunkból az első dobás értéke szerint szétválasztva a teljes várható érték tételből azt kapjuk, hogy

$$x = \frac{1}{9}\cdot(x+1)+\frac{1}{9}\cdot(x+2)+\frac{1}{9}\cdot(x+3)+\frac{1}{9}\cdot(x+4)+\frac{1}{9}\cdot(x+5)+\frac{1}{9}\cdot 6+\frac{1}{9}\cdot 7+\frac{1}{9}\cdot 8+\frac{1}{9}\cdot 9 = \frac{5}{9}x+5.$$

   Az egyenletet megoldva $x = 45/4$. Ugyanezt kapjuk, ha összeszorozzuk a húzások várható számát (geometriai, $\mathbb{E}(X) = 1/p = 9/4$) az egy húzásban szereplő szám várható értékével (diszkrét egyeletes, $\mathbb{E}(Y) = (n+1)/2 = 5$).