

A Valószínűségszámítás I. előadásorozat hatodik előadása.

2001 március 6.

Összefoglaló:

Néhány fontos diszkrét eloszlás.

a.) *Binomiális eloszlás.*

Tekintsünk először egy tipikus példát, amelyikben a binomiális eloszlás megjelenik.

Dobjunk fel egy pénzdarabot n alkalommal egymástól függetlenül, és minden egyes dobásban legyen p , $0 \leq p \leq 1$, a fejdobás valószínűsége. Mi a valószínűsége annak, hogy pontosan k fejdobás következik be?

Ennek valószínűsége $\binom{n}{k} p^k (1-p)^{n-k}$. Valóban, $\binom{n}{k}$ olyan fej-írás sorozat létezik, mely k fej és $n-k$ írás-jelet tartalmaz, és az adott feltételek mellett minden egyes ilyen dobássorozatnak a valószínűsége $p^k (1-p)^{n-k}$. Az ilyen jellegű feladatok leírására vezették be a binomiális eloszlás fogalmát.

Binomiális eloszlás definíciója. Azt mondjuk, hogy a ξ valószínűségi változó binomiális eloszlású n és p paraméterrel, $n = 1, 2, \dots$, $0 \leq p \leq 1$, ha $P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}$ minden $0 \leq k \leq n$ számra, és $P(\xi = k) = 0$ egyébként. Ezt az eloszlást szokás $B(n, p)$ -vel jelölni.

Egy $B(1, p)$ eloszlású ξ valószínűségi változót, azaz egy olyan ξ valószínűségi változót, melyre $P(\xi = 0) = 1 - p$, $P(\xi = 1) = p$ Bernoulli eloszlásúnak is szoktak nevezni.

Jegyezzük meg, hogy valóban eloszlást, definiáltunk, azaz a nem-negatív valószínűségek összege egy, mert a binomiális tétel szerint

$$\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$$

Jegyezzük meg, hogy a $B(n, p)$ és $B(m, p)$ binomiális eloszlások konvolúciója a $B(n + m, p)$ binomiális eloszlás.

Valóban a két eloszlás konvolúciója egy olyan $p(k)$, $0 \leq k \leq n + m$ eloszlás, melyre

$$\begin{aligned} p(k) &= \sum_{j=0}^k \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{k-j} p^{k-j} (1-p)^{(m-(k-j))} \\ &= p^k (1-p)^{n-k} \sum_{j=0}^k \binom{n}{j} \binom{m}{k-j} = \binom{n+m}{k} p^k (1-p)^{n-k}, \end{aligned}$$

mert

$$\sum_{j=0}^k \binom{n}{j} \binom{m}{k-j} = \binom{n+m}{k}.$$

Ez utóbbi állításnak kombinatorikus bizonyítása a következő: Az $\{1, 2, \dots, n + m\}$ számok közül $\binom{n+m}{k}$ módon tudunk k számot kiválasztani. Ez az azonosság jobboldalán álló kifejezés. Ugyanakkor ezen lehetőségeket úgy is összeszámolhatjuk, hogy rögzítve egy j számot, $\binom{n}{j} \binom{m}{k-j}$ módon választhatunk ki az első n számból j -t és az $n+1, \dots, n+m$ számok közül $k-j$ -t, majd összegezve $j = 0, \dots, k$ -ra ugyanazt a számot kapjuk. Ez viszont a baloldalon szereplő kifejezés.

Tanulságos lehet látni, hogyan lehet a fenti azonosságot, pontosabban annak egy lehetséges általánosítását bebizonyítani a hatványsoroknak a múlt órán említett tulajdonságai segítségével. Emlékeztetőül jegyezzük meg, hogy általában nem feltétlenül egész α számokra is definiáljuk az $\binom{\alpha}{k}$ számot mint $\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$, és az $(1+x)^\alpha$ függvény Taylor sorfejtése alapján

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \quad \text{ha } |x| < 1.$$

Továbbá,

$$(1+x)^\beta = \sum_{k=0}^{\infty} \binom{\beta}{k} x^k, \quad \text{ha } |x| < 1,$$

és

$$(1+x)^{\alpha+\beta} = \sum_{k=0}^{\infty} \binom{\alpha+\beta}{k} x^k, \quad \text{ha } |x| < 1.$$

Mivel $(1+x)^\alpha(1+x)^\beta = (1+x)^{\alpha+\beta}$, ezért az együtthatók összehasonlításából kapjuk, hogy

$$\sum_{j=0}^k \binom{\alpha}{j} \binom{\beta}{k-j} = \binom{\alpha+\beta}{k}$$

nem feltétlenül egész α és β számokra.

Feladat:

Adjunk egyszerű, a binomiális eloszlás valószínűségi tartalmát kihasználó bizonyítást arra, hogy a $B(n, p)$ és $B(m, p)$ binomiális eloszlások konvolúciója a $B(n+m, p)$ binomiális eloszlás.

Számítsuk ki egy $B(n, p)$ eloszlású valószínűségi változó várható értékét és szórásnégyzetét. A legegyszerűbb módszer a következő: Legyenek ξ_1, \dots, ξ_n független, $B(1, p)$ eloszlású valószínűségi változók. Ekkor $\xi = \sum_{j=1}^n \xi_j$ $B(n, p)$ eloszlású valószínűségi

változó. Ezért a keresett várható érték és szórásnégyzet $E\xi = \sum_{j=1}^n E\xi_j = np$ és $\text{Var } \xi =$

$$\sum_{j=1}^n \text{Var } \xi_j = np(1-p).$$

Feladat:

Számítsuk ki egy $B(n, p)$ eloszlású valószínűségi változó várható értékét és szórásnégyzetét az e fogalmakat definiáló összegek kiszámításának a segítségével.

A binomiális eloszlás természetes többdimenziós általánosítása a polinomiális eloszlás. Ennek megértéséhez tekintsük a következő feladatot. Adva van r urna. Ezekbe bedobunk összesen n golyót egymástól függetlenül. Az egyes golyók p_1 valószínűséggel esnek az első p_2 valószínűséggel a második, \dots p_r valószínűséggel az r -ik urnába. Mi a valószínűsége annak, hogy a j -ik urnába k_j golyó esik, $1 \leq j \leq r$, $\sum_{j=1}^r k_j = n$?

Ez a valószínűség $\frac{n!}{k_1!k_2!\dots k_r!}p_1^{k_1}\dots p_r^{k_r}$, mert $\frac{n!}{k_1!k_2!\dots k_r!}$ ilyen dobássorozat van, és minden ilyen dobássorozat valószínűsége $p_1^{k_1}\dots p_r^{k_r}$.

Polinomiális eloszlás definíciója $A(\xi_1, \dots, \xi_r)$ r változós véletlen vektor polinomiális eloszlású n , $n = 1, 2, \dots$, és p_j , $1 \leq j \leq r$, $\sum_{j=1}^r p_j = 1$ paraméterekkel, ha

$$P(\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_r = k_r) = \frac{n!}{k_1!k_2!\dots k_r!}p_1^{k_1}p_2^{k_2}\dots p_r^{k_r}$$

az olyan $0 \leq k_j \leq n$, $1 \leq j \leq r$ számokra, melyekre $\sum_{j=1}^r k_j = n$.

Feladat:

Legyen (ξ_1, \dots, ξ_r) polinomiális eloszlású véletlen vektor n és p_j , $1 \leq j \leq r$, $\sum_{j=1}^r p_j = 1$ paraméterekkel. Ekkor $E\xi_j = np_j$, $\text{Var } \xi_j = np_j(1 - p_j)$, $1 \leq j \leq r$ és $\text{Cov}(\xi_j, \xi_k) = -np_jp_k$, $1 \leq j, k \leq r$, $j \neq k$.

b.) *Negatív binomiális eloszlás és ennek speciális esete a geometriai eloszlás.*

Most is tekintsünk először egy tipikus példát, amelyikben ezek az eloszlások megjelennek.

Rögzítsünk egy r pozitív egész számot, és dobjunk fel egy pénzdarabot, amelyik p valószínűséggel esik a fej $1 - p$ valószínűséggel az írás oldalra egymás után többször egymástól függetlenül, egészen addig amikor az r -ik fejdobás megjelenik. Mi a valószínűsége annak, hogy ez a $k + r$ -ik dobás?

Ez a valószínűség $\binom{k+r-1}{r-1}(1-p)^k p^r$. Ez ugyanis azt jelenti, hogy az első $k+r-1$ dobásban pontosan $r-1$ fej és k írás dobás történt, aminek valószínűsége $\binom{k+r-1}{r-1}(1-p)^k p^{r-1}$, valamint a $k+r$ -ik dobás fej, aminek a valószínűsége p , és független az előző eseménytől.

Negatív binomiális és geometriai eloszlás definíciója. Azt mondjuk, hogy a ξ valószínűségi változó negatív binomiális eloszlású r , $r = 1, 2, \dots$, és p , $0 < p \leq 1$, paraméterrel, ha $\xi, r+1, \dots$, értékeket vesz fel és

$$P(\xi = k+r) = \binom{k+r-1}{r-1} (1-p)^k p^r, \quad k = 0, 1, 2, \dots$$

Az $r = 1$ és p paraméterhez tartozó negatív binomiális eloszlást p paraméterű geometriai eloszlásnak is nevezik.

A korábbi gyakorlatok bizonyos eredményeiből következik, hogy a negatív binomiális eloszlás valóban valószínűségeloszlás, azaz $\sum_{k=0}^{\infty} \binom{k+r-1}{r-1} (1-p)^k p^r = 1$, mert egy végtelen pénzfeldobás sorozat esetében egy valószínűséggel előbb-utóbb megjelenik r darab fejdobás. Valójában ez az indoklás hallgatólagosan felhasznál egy igaz, de az előadáson nem bizonyított fontos eredményt. Azt használtuk fel, hogy a végtelen független fej-írás dobássorozatoknak van valószínűségi modellje, ezért alkalmazhatjuk rá a valószínűségszámítás eredményeit. (A nehézséget az okozza, hogy csak bizonyos eddig nem tanult mértékelméleti eredmények segítségével lehet bebizonyítani azt, hogy az ebben a modellben természetes módon definiált valószínűség valóban σ -additív.) De ez az állítás következni fog néhány más később bizonyítandó eredményből is. Tárgyaljuk meg a negatív binomiális eloszlás néhány fontos tulajdonságát.

Feladat:

Bizonyítsuk be valószínűségi megfontolások segítségével, hogy egy r_1 és p paraméterű valamint egy r_2 és p paraméterű negatív binomiális eloszlás konvolúciója egy $r_1 + r_2$ és p paraméterű negatív binomiális eloszlású valószínűségeloszlás. Következésképpen, egy r és p paraméterű negatív binomiális eloszlás előáll mint r darab p paraméterű geometriai eloszlás konvolúciója.

Lássuk be az alábbi feladatot közvetlen számolással is. Elég belátni azt, hogy egy r és p paraméterű negatív binomiális eloszlású valamint egy p paraméterű geometriai eloszlás konvolúciója egy $r+1$ és p paraméterű negatív binomiális eloszlás.

Azt kell belátni, hogy

$$\sum_{j=0}^k \binom{r+j-1}{r-1} (1-p)^j p^r (1-p)^{k-j} p = \binom{k+r}{r} (1-p)^k p^{r+1}.$$

Ezt az azonosságot könnyen látjuk, ha tudjuk, hogy érvényes a következő azonosság:

$$\sum_{j=0}^k \binom{j+r-1}{r-1} = \binom{k+r}{r}.$$

Az utóbbi azonosság viszont következik például a következő kombinatorikai érvelésből. A természetes számok halmazán $\binom{k+r}{r}$ módon jelölhetünk ki $r+1$ számot úgy, hogy a

(nagyság szerint) $r + 1$ -ik szám a $k + r + 1$ szám legyen, és az azonosság jobboldala ezzel egyenlő. Ezt viszont úgyis kiszámolhatjuk, hogy azt tekintjük, hány olyan elrendezés van, melyben az r -ik kijelölt pont a $j + r$, az $r + 1$ -ik pedig a $k + r + 1$ szám, majd összegezzük $0 \leq j \leq k$ -ra. Mivel rögzített j -re az ilyen elrendezések száma $\binom{j+r-1}{r-1}$, innen következik a felírt azonosság.

Megmutatjuk, hogyan lehet a fenti összfüggést analitikus módon bebizonyítani. Tekintsük az r és p paraméterekhez tartozó $g_{r,p}(x)$ generátorfüggvényt, azaz az

$$g_{r,p}(x) = \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} (1-p)^k p^r x^{k+r}$$

összeget. Ha belátjuk, hogy ez az összeg konvergens például az $-1 < x < 1$ intervallumon, valamint $g_{r_1,p}(x)g_{r_2,p}(x) = g_{r_1+r_2,p}(x)$, akkor innen az előző előadás végén tett észrevételekből következik az állítás. A $g_{r,p}(x)$ függvény viszont egyszerűen zárt alakra hozható. Ennek érdekében vegyük észre, hogy

$$\begin{aligned} \binom{k+r-1}{r-1} &= \binom{k+r-1}{k} = \frac{(k+r-1)(k+r-2)\cdots(r+1)r}{k!} \\ &= \frac{r(r+1)\cdots(k+r-2)(k+r-1)}{k!} \\ &= (-1)^k \frac{(-r)(-r-1)\cdots(-r-k+2)(-r-k+1)}{k!} = (-1)^k \binom{-r}{k}, \end{aligned}$$

ahonnan

$$\begin{aligned} g_{r,p}(x) &= \sum_{k=0}^{\infty} (-1)^k \binom{-r}{k} (1-p)^k p^r x^{k+r} = (px)^r \sum_{k=0}^{\infty} \binom{-r}{k} (-(1-p)x)^k \\ &= (px)^r (1 - (1-p)x)^{-r} = \left(\frac{px}{1 - (1-p)x} \right)^r, \end{aligned}$$

ami az $(1+x)^{-r}$ függvény hatványsorának az alakjából látható. Innen viszont könnyen látható a generátorfüggvényekre felírt azonosság.

Számítsuk ki egy negatív binomiális eloszlású valószínűségi változó várható értékét és szórásnégyzetét. Ezt könnyen kiszámíthatnánk a generátorfüggvény deriválásának segítségével, de ehelyett egy geometriai eloszlású valószínűségi változó várható értékét és szórásnégyzetét kiszámoljuk direkt módon, majd az általános esetet visszavezetjük erre.

Legyen ξ geometriai eloszlású valószínűségi változó p paraméterrel, azaz legyen $P(\xi = k+1) = p^k(1-p)$, $k = 0, 1, \dots$. Ekkor

$$E\xi = \sum_{k=0}^{\infty} (k+1)(1-p)^k p = p \sum_{k=0}^{\infty} k(1-p)^{k-1},$$

$E\xi^2 = p \sum_{k=0}^{\infty} k^2(1-p)^{k-1}$, és $\text{Var } \xi = E\xi^2 - (E\xi)^2$. Deriváljuk kétszer a $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, $|x| < 1$ azonosságot. Azt kapjuk, hogy

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1},$$

$$\frac{2}{(1-x)^3} = \sum_{k=2}^{\infty} k(k-1)x^{k-2}.$$

Innen $x = 1-p$ helyettesítéssel $\frac{1}{p^2} = \sum_{k=1}^{\infty} k(1-p)^{k-1}$, $E\xi = p \sum_{k=1}^{\infty} k(1-p)^{k-1} = \frac{p}{p^2} = \frac{1}{p}$,
 $\sum_{k=0}^{\infty} k^2(1-p)^{k-1} = (1-p) \sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} + \sum_{k=1}^{\infty} k(1-p)^{k-1} = \frac{2(1-p)}{p^3} + \frac{1}{p^2}$,
 $E\xi^2 = \frac{2(1-p)}{p^2} + \frac{1}{p}$, $\text{Var } \xi = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1}{p^2} - \frac{1}{p} = \frac{1-p}{p^2}$.

Feladat:

Számítsuk ki egy n és p paraméterű negatív binomiális eloszlású valószínűségi változó várható értékét és szórásnégyzetét annak az ismeretnek a segítségével, hogy egy p paraméterű geometriai eloszlású valószínűségi változó várható értéke $\frac{1}{p}$ és szórásnégyzete $\frac{1-p}{p^2}$.

Tárgyaljuk meg, hogyan lehet egy negatív binomiális eloszlású valószínűségi változó várható értékét és szórásnégyzetét kiszámolni a generátorfüggvénye ismeretében. Sőt, lássunk be egy olyan formulát, mely általánosabb esetben is alkalmazható.

Legyen $\mathcal{P} = \{p_k : k = 1, 2, \dots\}$, $p_k \geq 0$, $k = 0, 1, 2, \dots$, $\sum_{k=0}^{\infty} p_k = 1$, valószínűség eloszlás a nem negatív egész számokon. Vezessük be az ezen eloszláshoz tartozó

$$g(x) = g_{\mathcal{P}}(x) = \sum_{k=0}^{\infty} p_k x^k.$$

generátorfüggvényt. Vegyük észre, hogy a $g(x)$ függvényt definiáló hatványsor konvergens a $-1 < x < 1$ intervallumban. Továbbá kétszeri deriválás és az $x = 1$ (formális) helyettesítés azt sugallja, hogy

$$g'(x) = \sum_{k=1}^{\infty} k p_k x^{k-1}, \quad g''(x) = \sum_{k=2}^{\infty} k(k-1) p_k x^{k-2},$$

valamint

$$g'(1) = \sum_{k=1}^{\infty} k p_k, \quad g''(1) = \sum_{k=2}^{\infty} k(k-1) p_k,$$

és ezen azonosságok segítségével kiszámolhatjuk a keresett várható értéket és szórásnégyzetét.

Felmerül a kérdés, szabad-e az alábbi számolásokat végrehajtani. A problémát az okozza, hogy hatványsorokat szabad tagonként deriválni a konvergenciatartomány belsejében, de mi az $x = 1$ helyettesítést hajtottuk végre, és lehet, hogy a $g(x)$ függvény hatványsora nem terjeszthető ki egy a $(-1, 1)$ intervallumnál nagyobb intervallumra. A negatív binomiális eloszlás generátorfüggvénye esetén egy ilyen kiterjesztés lehetséges, de érdemes az általános esetet is tekinteni, és ekkor nem szabad kizárni ezt a lehetőséget.

Az analízis bizonyos eredményei biztosítják a fenti számolás jogosságát. Egyrészt igaz az, hogy ha egy $h(x) = \sum_{k=0}^{\infty} a_k x^k$ hatványsor konvergens egy $(-A, A)$ intervallumban, és a $h(A) = \sum_{k=0}^{\infty} a_k A^k$ összeg konvergens, akkor $\lim_{x \rightarrow A} h(x) = h(A)$. Másrészt, ha a $h(x)$ Taylor-sor a_k együtthatói nem negatívak, akkor $\lim_{x \rightarrow A} h(x) = h(A)$. Speciálisan, ha ebben az esetben $\lim_{x \rightarrow A} h(x) < \infty$ akkor $h(A) < \infty$. Ezeknek az eredményeknek az alkalmazásával $A = 1$ és $h(x) = g'(x)$ illetve $h(x) = g''(x)$ választással meg lehet mutatni a fenti számolások jogosságát. Érdemes megjegyezni, hogy az általános esetben a $g'(1) = \lim_{x \rightarrow 1} g'(x)$ $g''(1) = \lim_{x \rightarrow 1} g''(x)$ képletet kell alkalmaznunk annak érdekében, hogy a $g'(1)$ és $g''(1)$ számokat értelmezni tudjuk.

Megjegyzés: Az, hogy egy hatványsor hogyan viselkedik a konvergenciakörének szélén, a komplex függvénytan egyik nehéz és fontos kérdése. Annak érdekében, hogy lássunk egy egyszerű példát, mely rávilágíthat arra, hogy vigyázni kell a formális számolások során tekintsük az $h(x) = \frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k$ hatványsort. Ekkor $h(x)$ konvergens a $-1 < x < 1$ intervallumban, $h(1) = \frac{1}{2}$, és $h(x)$ hatványsora az $x = 1$ pontban, a $\sum_{k=0}^{\infty} (-1)^k$ sor, divergens. Ez a példa azért nem mond ellen a fent elmondottaknak, mert ebben a hatványsorban vannak negatív együtthatók is.

Feladat:

Lássuk be, hogy amennyiben egy ξ valószínűségi változó generátorfüggvénye valamely $g(x)$ függvény, akkor $E\xi = g'(1)$, $\text{Var } \xi = g''(1) + g'(1) - g'(1)^2$, ahol az általános esetben a $g'(1) = \lim_{x \rightarrow 1} g'(x)$ és $g''(1) = \lim_{x \rightarrow 1} g''(x)$ képletek definiálják ezeket a mennyiségeket. Számoljuk ki egy r és p paraméterű negatív binomiális eloszlású valószínűségi változó várható értékét és szórásnégyzetét, felhasználva, hogy e valószínűségi változó generátorfüggvénye a $\left(\frac{px}{1 - (1-p)x}\right)^r$ függvény.

Feladat:

Mutassunk példát olyan $\mathcal{P} = \{p_k : k = 1, 2, \dots\}$, $p_k \geq 0$, $k = 0, 1, 2, \dots$, $\sum_{k=0}^{\infty} p_k = 1$, valószínűség eloszlásra a nem negatív egész számokon, melynek $g(x) = \sum_{k=0}^{\infty} p_k x^k$

generátorfüggvénye semmilyen $\varepsilon > 0$ szám esetén nem konvergens a $-1 - \varepsilon < x < 1 + \varepsilon$ intervallumon.

Feladat:

Legyen ξ geometriai eloszlású valószínűségi változó, azaz legyen $P(\xi = k) = p(1 - p)^{k-1}$, $k = 1, 2, \dots$. Lássuk be, hogy ξ teljesíti a következő (diszkrét) örökifjú tulajdonságot.

$$P(\xi = k + l | \xi > l) = p(1 - p)^{k-1} = P(\xi = k).$$

Adjuk meg ennek az azonosságnak a valószínűségszámítási magyarázatát is.

b.) *Hipergeometrikus eloszlás.*

Tekintsünk először egy tipikus példát, amelyikben ilyen eloszlások megjelennek.

Adva van egy urnában M piros és $N - M$ fehér golyó. Visszatevés nélkül kihúzzunk n , $n \leq N$ golyót. Mi annak a valószínűsége, hogy k piros golyót húztunk ki?

Ez a valószínűség $\frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$. Valóban, ha megkülönböztetjük az egyes

golyókat, akkor $\binom{N}{n}$ különböző húzáseredmény alakulhat ki. (Nem különböztetünk meg két húzáseredményt, ha ugyanazokat a golyókat húztuk ki csak más sorrendben.). Olyan húzássorozat, melyben az M piros golyóból k -t, az $N - M$ fehér golyóból pedig $n - k$ -t húzzunk $\binom{M}{k} \binom{N-M}{n-k}$ van. Mivel ezek azok az n hosszú húzássorozatok, melyekben pontosan k piros golyót húzzunk, és az egyes húzássorozatok valószínűsége megegyezik, innen következik az állítás.

A hipergeometrikus eloszlás definíciója. Azt mondjuk, hogy a ξ valószínűségi változó hipergeometrikus eloszlású N , M és n paraméterekkel, ahol N , M és n nem negatív egész számok, $M < N$, $n < N$, ha

$$P(\xi = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}, \quad 0 \leq k \leq n.$$

Ilyen módon valóban eloszlást definiáltunk. A $\sum_{k=0}^n P(\xi = k) = 1$ azonossággal ekvivalens $\sum_{k=0}^n \binom{M}{k} \binom{N-M}{n-k} = \binom{N}{n}$ azonosságot (más jelöléssel) beláttuk a binomiális eloszlás vizsgálatában.

Feladat:

Számoljuk ki annak valószínűségét, hogy a lottóban pontosan három találatunk lesz.

Számoljuk ki egy hipergeometrikus eloszlású valószínűségi változó várható értékét és szórásnégyzetét.

Feladat:

Számoljuk ki egy N , M és n paraméterű hipergeometrikus eloszlású ξ valószínűségi változó várható értékét és szórásnégyzetét, azaz egy olyan valószínűségi változó várható értékét és szórásnégyzetét, melyre

$$P(\xi = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}, \quad 0 \leq k \leq n.$$

Mutassuk meg, hogy $E\xi = n \frac{M}{N}$, $\text{Var } \xi = \frac{nM}{N} \left(1 - \frac{M}{N}\right) \frac{n-n}{N-1}$.

Segítség: Tekintsünk egy urnamodellt, melyben ξ jelöli azt, hogy N piros és $N-M$ fehér golyóból n visszatevés nélküli húzásban hány piros golyót húzunk. Legyen $\xi_j = 1$, ha a j -ik húzás eredménye piros, $\xi_j = 0$, ha a j -ik húzás eredménye fehér, $1 \leq j \leq n$. Ekkor a $\xi = \sum_{j=1}^n \xi_j$ valószínűségi változó várható értékét és szórásnégyzetét kell kiszámolnunk. Hogyan kell ezt csinálni? Emlékezzünk arra, hogy mint korábban megmutattuk, ebben a modellben annak a valószínűsége, hogy valamely húzás(ok)ban piros golyót húzunk nem függ attól, hogy hanyadik húzást tekintettük.

Ha egy urnahúzásban N és $N-M$ nagy, azaz sok fehér és piros golyó van az urnában, és fix számú golyót kihúzunk, akkor a húzáseredmény szempontjából alig van jelentősége annak, hogy a kihúzott golyókat visszadobjuk-e vagy sem. Ilyen jellegű állítást fogalmaz meg a következő (egyszerű) feladat.

Feladat:

Ha $N \rightarrow \infty$, $\lim_{N \rightarrow \infty} \frac{M}{N} = p$, $0 < p < 1$, n rögzített egész szám, akkor

$$\lim_{N \rightarrow \infty} \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{minden } 0 \leq k \leq n \text{ számra.}$$

A hipergeometrikus eloszlás természetes általánosítása a polihipergeometrikus eloszlás. Ennek megértése érdekében tekintsük a következő feladatot: Egy urnában r különböző színű golyó van, N_1 1-es, N_2 2-es, \dots N_r r -es színű golyó. Legyen $N = \sum_{j=1}^r N_j$.

Ezekből a golyókból kihúzzunk n -et visszatevés nélkül. Mi annak a valószínűsége, hogy k_1 1-es, k_2 2-es, \dots k_r r -es színű golyót húzzunk ki?

A válasz:

$$\frac{\binom{N_1}{k_1} \binom{N_2}{k_2} \cdots \binom{N_r}{k_r}}{\binom{N}{n}},$$

ahol $n = \sum_{j=1}^r k_j$. Ennek az állításnak a bizonyítása hasonlóan történhet mint a hipergeometrikus eloszlás bevezetése előtt tekintett feladaté. Ennek a feladatnak az alapján vezették be a következő fogalmat.

A polihipergeometrikus eloszlás definíciója. Azt mondjuk, hogy a (ξ_1, \dots, ξ_r) véletlen vektor polihipergeometrikus eloszlású N_1, \dots, N_r és n paraméterekkel, ahol N_j , $1 \leq j \leq r$, és n nem negatív egész számok, és $N = \sum_{j=1}^r N_j$ jelöléssel $n < N$, ha

$$P(\xi_1 = k_1, \dots, \xi_r = k_r) = \frac{\binom{N_1}{k_1} \binom{N_2}{k_2} \cdots \binom{N_r}{k_r}}{\binom{N}{n}},$$

ahol $n = \sum_{j=1}^r k_j$.

Feladat:

Legyenek ξ_1, \dots, ξ_L és η_1, \dots, η_M valószínűségi változók ugyanazon a valószínűségi mezőn. Lássuk be, hogy

$$\text{Cov} \left(\sum_{j=1}^L \xi_j, \sum_{k=1}^M \eta_k \right) = \sum_{j=1}^L \sum_{k=1}^M \text{Cov}(\xi_j, \eta_k).$$

Feladat:

Legyen (ξ_1, \dots, ξ_r) véletlen vektor polihipergeometrikus eloszlású N_1, \dots, N_r és n paraméterekkel. Számítsuk ki az $E\xi_j$ várható értékeket, a $\text{Var} \xi_j$ szórásnégyzetet minden $1 \leq j \leq r$ számra és a $\text{Cov}(\xi_j, \xi_k)$ kovarianciafüggvényt, ha $1 \leq j, k \leq r$, $j \neq k$.

Segítség: A hipergeometrikus eloszlásról szóló hasonló feladat módszere természetes módon adaptálható ebben az esetben is.

d.) *Poisson eloszlás.*

A Poisson eloszlást közvetlenül fogjuk definiálni. Azok a tulajdonságai, amelyek miatt fontos szerepet játszik a valószínűségi számításban később fognak kiderülni.

Poisson eloszlás definíciója. *Egy ξ valószínűségi változó λ paraméterű Poisson eloszlású, ha ξ nem negatív egész értékeket vesz fel, és*

$$P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Ez valóban valószínűség eloszlás, mert

$$\sum_{k=0}^{\infty} P(\xi = k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

Számítsuk ki egy λ paraméterű Poisson eloszlású valószínűségi változó generátorfüggvényét. Ez

$$\sum_{k=0}^{\infty} P(\xi = k) x^k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} x^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda x)^k}{k!} = e^{-\lambda} e^{\lambda x} = e^{\lambda(x-1)}.$$

A következő állítást fontossága miatt fogalmazzuk meg Tétel formájában.

Tétel *Ha ξ és η két független, λ illetve μ paraméterű Poisson eloszlású valószínűségi változó, akkor $\xi + \eta$ $\lambda + \mu$ paraméterű Poisson eloszlású valószínűségi változó.*

Bizonyítás:

$$\begin{aligned} P(\xi + \eta = k) &= \sum_{j=0}^k P(\xi = j, \eta = k - j) = \sum_{j=0}^k P(\xi = j) P(\eta = k - j) \\ &= \sum_{j=0}^k \frac{\lambda^j}{j!} e^{-\lambda} \frac{\mu^{k-j}}{(k-j)!} e^{-\mu} = \frac{e^{-(\lambda+\mu)}}{k!} \sum_{j=0}^k \frac{k!}{j!(k-j)!} \lambda^j \mu^{(k-j)} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} (\lambda + \mu)^k. \end{aligned}$$

Feladat:

Mutassuk meg, hogy a Poisson eloszlás generátorfüggvényének alakjából látható, hogy ha ξ és η két független, λ illetve μ paraméterű Poisson eloszlású valószínűségi változó, akkor $\xi + \eta$ $\lambda + \mu$ paraméterű Poisson eloszlású valószínűségi változó.

A fenti eredmény következménye, hogy amennyiben véges sok független, Poisson eloszlású valószínűségi változónak vesszük az összegét, az ismét Poisson eloszlású lesz, melynek paramétere az egyes valószínűségi változók paramétereinek

az összege. A következő feladat állítása, melynek érdekes következményei vannak, tekinthető úgy mint ennek az állításnak a megfordítása. Abban ugyanis egy alkalmas konstrukció segítségével egy Poisson eloszlású valószínűségi változót bontunk fel független kisebb paraméterű Poisson eloszlású valószínűségi változók összegére.

Feladat:

Legyen adva k darab urna, és ezekbe dobjunk be véletlen ξ számú golyót, ahol ξ Poisson eloszlású valószínűségi változó $\lambda > 0$ paraméterrel. Legyenek az egyes dobások eredményei egymástól és a ξ valószínűségi változótól függetlenek. Tegyük fel továbbá, hogy minden egyes dobásnál a golyó az j -ik urnába $p_j \geq 0$ valószínűséggel esik, $j = 1, \dots, k$, $\sum_{j=1}^k p_j = 1$. Jelölje η_j a j -ik urnába eső golyók számát. Ekkor az η_j , $j = 1, \dots, k$ valószínűségi változók függetlenek, és η_j Poisson eloszlású λp_j paraméterrel, $j = 1, \dots, k$.

Lássuk be az előző feladat segítségével a következő állítást:

Feladat:

Legyen adva ξ Poisson eloszlású valószínűségi változó λ paraméterrel. Dobjunk le egymástól és a ξ valószínűségi változótól függetlenül véletlenül egyenletesen ξ darab pontot az egységintervallumra, azaz tegyük fel, hogy minden pont $b - a$ valószínűséggel esik valamely $[a, b] \subset [0, 1]$ intervallumba. Ekkor a $[0, 1]$ intervallum tetszőleges felbontására $[s_0, s_1], [s_1, s_2], \dots, [s_{k-1}, s_k]$ intervallumokra, $0 = s_0 < s_1 < \dots < s_k = 1$, az egyes intervallumokba eső pontok száma egymástól független Poisson eloszlású valószínűségi változó $s_j - s_{j-1}$, $1 \leq j \leq k$, paraméterrel.

Számítsuk ki egy ξ λ paraméterű Poisson eloszlású valószínűségi változó várható értékét és szórásnégyzetét.

$$E\xi = \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda + \lambda} = \lambda,$$

$$\begin{aligned} E\xi^2 &= \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=2}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} + \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} \\ &= e^{-\lambda} \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda = \lambda^2 + \lambda, \end{aligned}$$

és $\text{Var } \xi = E\xi^2 - (E\xi)^2 = (\lambda + \lambda^2) - \lambda^2 = \lambda$.

Jegyezzük meg, hogy a Poisson eloszlás várható értékére és szórásnégyzetre kapott eredmények összhangban vannak azzal a ténnyel, hogy független Poisson eloszlású valószínűségi változók összege olyan Poisson eloszlású valószínűségi változó, melynek paramétere az összeadandók paramétereinek az összege. Ugyanis mind a várható érték mind a szórásnégyzet additív független valószínűségi változók összegzése esetén.

Feladat:

Legyenek ξ_j , $j = 1, \dots, r$, független Poisson eloszlású valószínűségi változók λ_j , $1 \leq j \leq r$, paraméterekkel. Mutassuk meg, hogy

$$P\left(\xi_1 = k_1, \dots, \xi_r = k_r \mid \sum_{j=1}^r \xi_j = n\right) = \frac{n!}{k_1! \cdots k_r!} \prod_{j=1}^r \frac{\lambda_j^{k_j}}{\left(\sum_{s=1}^r \lambda_s\right)^{k_j}},$$

ha $\sum_{j=1}^r k_j = n$. Azaz a ξ_1, \dots, ξ_r vektor feltételes eloszlása feltéve, hogy $\sum_{j=1}^r k_j = n$

a polinomiális eloszlás n és $p_j = \frac{\lambda_j}{\sum_{s=1}^r \lambda_s}$, $1 \leq j \leq r$, paraméterekkel.

Jegyezzük meg, hogy ha jól megértjük a valószínűségi mezők és változók szerkezetét a fenti állítás kiolvasható abból a korábban megfogalmazott feladat eredményéből is, mely megad egy lehetőséget arra, hogy egy Poisson eloszlású valószínűségi változót felbontsunk kisebb paraméterű független Poisson eloszlású valószínűségi változók összegére. Ugyanis, ha abban a ledobott ξ golyó száma $\xi = \sum_{j=1}^r \xi_j$, akkor $\xi = n$ feltétel választással, azaz pontosan n pontot dobva le, a fenti eredményt kapjuk.

Lemma. Minden $n = 1, 2, \dots$ számra tekintsünk egy S_n binomiális eloszlású valószínűségi változót $B(n, p_n)$ eloszlással, azaz n és p_n paraméterekkel. Tegyük fel továbbá, hogy a $\lambda_n = np_n$, $n = 1, 2, \dots$, számok teljesítik a $\lim_{n \rightarrow \infty} \lambda_n = \lambda$ feltételt valamilyen $\lambda > 0$ számmal. Ekkor

$$\lim_{n \rightarrow \infty} P(S_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

minden $k = 0, 1, 2, \dots$ számra.

Bizonyítás:

$$\begin{aligned} P(S_n = k) &= \binom{n}{k} \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \left(\frac{\lambda_n}{n}\right)^k = \\ &= \frac{\lambda_n^k}{k!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \frac{n(n-1) \cdots (n-k+1)}{n^k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}, \end{aligned}$$

ha $n \rightarrow \infty$ minden $k = 0, 1, 2, \dots$ számra, mert $\frac{\lambda_n^k}{k!} \rightarrow \frac{\lambda^k}{k!}$, $\left(1 - \frac{\lambda_n}{n}\right)^n \rightarrow e^{-\lambda}$,

$\left(1 - \frac{\lambda_n}{n}\right)^{-k} \rightarrow 1$, és $\frac{n(n-1) \cdots (n-k+1)}{n^k} \rightarrow 1$, ha $n \rightarrow \infty$.