Binomiális eloszlás

A Wikipédiából, a szabad enciklopédiából

Az X valószínűségi változó n és p paraméterű **binomiális eloszlást követ** – vagy rövidebben **binomiális eloszlású** – pontosan akkor, ha

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, ..., n \quad ,$$

ahol $0 < p < 1$. Azt, hogy az X valószínűségi változó n és p paraméterű binomiális eloszlást követ, a következő módon szoktuk jelölni: $X \sim B(n,p)$. Speciálisan, ha $X \sim B(1,p)$, akkor X-et Bernoulli-eloszlásúnak nevezzük.

A képlet szemléletes jelentése

A binomiális eloszlású valószínűségi változóval a *visszatevéses mintavétel* ragadható meg, vagyis olyan helyzeteket lehet vele modellezni, ahol egy véletlen kísérletet tetszőlegesen sokszor lehet megismételni ugyanolyan körülmények között, miközben azt figyeljük meg, hogy az n ismétlés során hányszor következett be egy adott esemény. A képletben az esemény bekövetkezésének gyakoriságát jelöli k – ez magyarázza a $k = 0, 1, 2, \ldots, n$ feltételt –, a figyelemmel követett esemény valószínűségét pedig p – ami pedig a $0 < p < 1$-et.

Húzzunk például egy pakli magyar kártyából egy lapot, nézzük meg, hogy makk-e a húzott lap színe, majd tegyük vissza, keverjük meg a paklit, és húzzunk újra. Ha a húzást mondjuk ötször ismételjük ezzel az eljárással, akkor a binomiális eloszlás adja meg annak a valószínűségét, hogy $k = 0, 1, 2, 3, 4$ vagy 5 esetben húztunk makkot.

Gondoljuk át, hogy mi is a valószínűsége annak, hogy az öt húzott lap közül mondjuk háromnak lesz a színe makk. Tetszőleges húzás esetén a makk színű lap húzásának valószínűsége 1/4, hiszen a lapok 1/4-e makk színű, így az egyéb színek húzásának valószínűsége 3/4. Emiatt annak a valószínűsége, hogy például az első három húzás lesz makk, és a második kettő nem, az pontosan $\frac{1}{4}\cdot\frac{1}{4}\cdot\frac{1}{4}\cdot\frac{3}{4}\cdot\frac{3}{4} = (1/4)^3\cdot(3/4)^2$ lesz. De természetesen nem csak így állhat elő az a helyzet, hogy az öt húzott kártyából pontosan háromnak makk a színe. Lehet, hogy az utolsó három lap makk színű, lehet, hogy az első, a harmadik, az ötödik stb. Ezeknek a lehetőségeknek egyenként mindig $(1/4)^3\cdot(3/4)^2$ a valószínűsége, így csak az a dolgunk, hogy összeszámoljuk hányféle sorrendben húzható három makk és két nem makk színű lap. Ez megegyezik azzal, ahányféleképpen egy ötelemű halmazból három elem kiválasztható, vagyis kombinatorikából tudjuk, hogy $\binom{5}{3}$-mal. Így az eredeti kérdésre a válaszunk: annak a valószínűsége, hogy a magyar kártyából visszatevéssel öt lapot húzva pontosan háromszor húzunk makk színű lapot,

$$\binom{5}{3}(1/4)^3(3/4)^2.$$

Ugyanezzel a gondolatmenettel kapjuk annak a valószínűségét, hogy pontosan 2, 4 vagy általában *k* darab makk színű lapot húzunk az öt húzás alatt. Ez *k*-ra ($k = 0, 1, \ldots, 5$)

$$\binom{5}{k}(1/4)^k(3/4)^{5-k}.$$

Látható, hogy az 5 és 1/4 paraméterű binomiális eloszlás képletét kaptuk meg.

Itt most a véletlen kísérletünk a lap húzása és annak megfigyelése volt, hogy a húzott lap makk színű-e vagy nem. Azt hogy ezt a kísérletet többször *ugyanolyan körülmények között* lehetett megismételni, éppen a húzott lap visszatevése tette lehetővé. Gondoljuk meg, hogy amennyiben nem tesszük vissza a lapot, és elsőre makkot húzunk, akkor a második makk húzásának valószínűsége már nem 1/4, hiszen már nem a lapok 1/4-e makk színű, hanem csak 7/31-e. (Ezt a fajta mintavételt, a visszatevés nélküli mintavételt a hipergeometrikus eloszlással tudjuk leírni.) Az adott esemény, amit az 5 ismétlés során megfigyeltünk, az volt, hogy hányszor következik be makk színű lap húzása.

Hasonlóan, ha egy erdőben ücsörögve egy erdei iramszarvas populáció tagjait figyeljük meg egy lesről, és tudjuk, hogy ebben a populációban az erdei iramszarvasok 10%-a zöldhátú, akkor annak a valószínűségét, hogy 11 megfigyelt erdei iramszarvasból pontosan *k* lesz zöldhátú, a 11 és 1/10 paraméterű binomiális eloszlás írja le.

A binomiális eloszlást jellemző számok

Várható értéke $\mathbf{E}(X) = np$

Szórása $\mathbf{D}(X) = \sqrt{np(1-p)}$