Határeloszlás-tételek. A binomiális és a Poisson eloszlások kapcsolata. A központi határeloszlás tétel. Valószínűségi változókra vonatkozó konvergencia fogalmak.

Binomiális eloszlás

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

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

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 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.

Poisson-eloszlás

A valószínűség-számításban és a statisztikában a **Poisson-eloszlás** egy diszkrét valószínűségi eloszlás, a binomiális eloszlás határeloszlása. Kifejezi az adott idő alatt ismert valószínűséggel megtörténő események bekövetkezésének számát (például: egy telefonközpontba adott időszakban és időtartamban beérkezett telefonhívások száma, vagy egy radioaktív anyag adott idő alatt elbomló atomjainak száma).

Az X valószínűségi változó λ paraméterű Poisson-eloszlást követ – vagy rövidebben: Poisson-eloszlású – pontosan akkor, ha

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

ahol $0 < \lambda$ konstans.

Ha binomiális eloszlások olyan sorozatát vesszük, melyben az eloszlások *n* paramétere úgy tart a végtelenbe, hogy közben az *np* szorzat konstans marad (*p* így nyilván a 0-hoz tart), akkor határeloszlásként Poisson-eloszlást kapunk.

A központi határeloszlás-tétel az általános esetben

A matematikai statisztika módszereinek jelentős része arra a feltevésre épül, hogy a megfigyelt mennyiség normális eloszlású. Azt, hogy a megfigyelt mennyiségek igen gyakran (közelítőleg) normális eloszlást követnek, egyrészt a tapasztalat mutatja, másrészt elméletileg a központi határeloszlás-tételek támasztják alá.

2.14. Tétel. *Legyenek* $\xi_1, \xi_2, \ldots$ *független, azonos eloszlású valószínűségi változók,* $S_n=\xi_1+\cdots+\xi_n$. *Tegyük fel, hogy* $\sigma^2=\mathbb{D}^2\xi_1$ *véges és pozitív, és legyen* $m=\mathbb{E}\xi_1$. *Ekkor*

$$\boxed{\lim_{n\to\infty} P\left(\frac{S_n-nm}{\sqrt{n}\sigma}<x\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\,dt,} \tag{2.30}$$

$\forall x \in \mathbb{R}$. □

A (2.30) képlet jelentése: S_n standardizáltjának eloszlásfüggvénye a standard normális eloszlásfüggvényhez tart $n\to\infty$ esetén. Megjegyezzük, hogy a központi határeloszlás tétel általánosabb alakjai a fentinél jóval gyengébb feltételek esetén állítják, hogy valószínűségi változók összegei normális eloszláshoz tartanak. A 2.14. Tétel bizonyítását a karakterisztikus függvények módszerével fogjuk megadni.

A 2.14. TÉTEL BIZONYÍTÁSA.* A folytonossági tétel miatt elegendő belátni, hogy az S_n standardizáltjának karakterisztikus függvénye konvergál a standard normális eloszlás karakterisztikus függvényéhez. E célból vezessük be a következő jelöléseket:

$$\varphi_n(t)=\mathbb{E}\exp\left\{\frac{it(S_n-nm)}{\sqrt{n}\sigma}\right\}, \qquad \varphi(t)=\mathbb{E}\exp\left\{it(\xi_1-m)\right\}.$$

A 6.7. Tétel alapján

$$\varphi_n(t) = \left[\varphi\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n.$$

Viszont a 6.11. Tétel miatt

$$\varphi(t) = 1 - \frac{\sigma^2 t^2}{2} + g(t),$$

ahol $g(t)/t^2 \to 0$, ha $t \to 0$. Ezek alapján

$$\varphi_n(t) = \left[1 - \frac{\sigma^2 t^2}{2n\sigma^2} + g\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n.$$

Ez a kifejezés viszont $\varphi_n(t) = \left[1 - \frac{t^2/2 - \varepsilon_n(t)}{n}\right]^n$ alakra hozható, ahol $\varepsilon_n(t) \to 0$, ha $n \to \infty$. Így

$$\lim_{n\to\infty} \varphi_n(t) = e^{-t^2/2},$$

azaz $\varphi_n(t)$ a standard normális eloszlás karakterisztikus függvényéhez konvergál. ☐

Valószínűségi változó

A **valószínűségi változó** a valószínűség-számítás egyik legfontosabb fogalma. Lényegében olyan jelenségek matematikai megfogalmazására, modellezésére alkalmas, melyek véletlentől függő értéket vesznek fel. Ilyen lehet például egy kockadobás eredménye, egy folyó vízállása, vagy az utcán szembe jövő emberek testmagassága.

Bár a valószínűségi változó szemléletes jelentése viszonylag könnyen megragadható, a precíz matematikai meghatározás a huszadik századig váratott magára, és egészen komoly függvénytani illetve mértékelméleti eszközöket használ fel.

A valószínűségi változók két leggyakrabban emlegetett fajtája a **diszkrét** és a **folytonos** valószínűségi változó. Szemléletesen a diszkrét valószínűségi változó olyan, ami elkülönült értékeket tud csak felvenni, a folytonos pedig olyan, ami – legalább egy intervallumon – bármilyen értéket felvehet. Diszkrét valószínűségi változó például az, ami egy kockadobás eredményét írja le, vagy azt, hogy egy üzletbe következőnek betoppanó 8 vendég közül hány férfi. Ezzel szemben folytonosnak tekinthető az a valószínűségi változó, ami azt írja le, hogy az ugyanebbe az üzletbe betoppanó következő vevő milyen magas, vagy hogy egy fáról leszüretelt őszibarack mekkora súlyú, hisz ezek a változók – legalább is egy intervallomon – akármilyen értéket felvehetnek.