A MATEMATIKA TANÍTÁSA

MÓDSZERTANI FOLYÓIRAT

IV. ÉVFOLYAM 1996 4

MOZAIK OKTATÁSI STÚDIÓ

Fazekas István – Tómács Tibor

A valószínűségszámítás szemléletes oktatásáról

Ezen cikket mindazon gyakorló matematikatanároknak és tanárjelölteknek ajánljuk, akik az általános, illetve középiskolai óráikon a törzsanyagon kívül a valószínűségszámítás törvényszerűségeiről is szeretnének szólni.

Az általános tapasztalat szerint a valószínűségszámítás iskolai oktatása megmarad azon feladatok és tételek körében, melyek kombinatorikai eszközökkel kezelhetők. A mélyebb törvények bemutatására viszonylag kevés figyelem jut. Ha belegondolunk, akkor ez természetes is, hiszen ezen tételeknek már a pontos kimondása is mély matematikai fogalmakat igényel. Nem beszélve ezek bizonyításáról, melyek gyakran még főiskolákon, sőt egyetemeken sem kerülnek bemutatásra, éppen a megfelelő matematikai apparátus bonyolultsága miatt.

Jelen cikkünkben arra vállalkozunk, hogy néhány példán keresztül bemutassuk, mindezek ellenére van mód a valószínűségszámítás legfontosabb törvényeinek megértetésére szinte bármilyen előképzettségű diák számára, természetesen bizonyítás nélkül, pusztán a szemléletre hagyatkozva.

Véletlen kísérleteket fogunk elvégezni, s ezek kimeneteleit vizsgáljuk. A valószínűségszámítás törvényei nagyszámú ismétlés esetére vonatkoznak, így a szükséges kísérleteknek csak egy töredékét tudjuk ténylegesen végrehajtani, hosszú kísérletsorozatokat számítógépen szimulálunk.

Az alábbiakban a nagy számok törvényének, a bolyongásnak és a centrális határeloszlás tételnek a számítógépes demonstrációjáról lesz szó. Nem célunk a precíz matematikai háttér bemutatása, valamint annak elemzése, hogy a generált véletlen számok mennyiben teljesítik a függetlenség és az azonos eloszlás feltételét.

1. A nagy számok törvénye

A nagy számok Bernoulli-féle törvénye – miszerint a relatív gyakoriság sztochasztikusan konvergál a valószínűséghez – egyben arra is alkalmas, hogy a diákok képet alkossanak a véletlen jelenségek természetéről, illetve hogy összevessék a véletlen bennük élő intuitív fogalmát a valósággal.

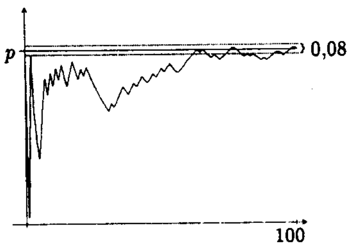
Igen fontos, hogy legelőször arra kérjük meg a diákokat, hogy végezzenek el egy kísérletsorozatot gondolatban. Az elvégzendő kísérlet lényege, hogy egy pénzérmét dobunk fel többször, és azt figyeljük, hogy egy-egy dobás után a fej vagy az írás oldal van-e felül. A diákok írják le egy körülbelül 100 dobásból álló kísérletsorozat eredményeit úgy, ahogyan ők elképzelik a kimeneteleket.

Ezek után végezzük el ténylegesen is (nem számítógéppel) az érme feldobásait, és az eredményt hasonlítsuk össze a gondolati kísérlet eredményével. Tapasztalni fogjuk, hogy a tényleges kísérletben több olyan viszonylag hosszú szakaszt is találunk, melyek csupa fejből, vagy csupa írásból állnak. Ezzel ellentétben ilyen szakaszok a gondolati kísérletekben csak elvétve szoktak előfordulni. Vagyis a véletlen kísérletben nincs törekvés a kiegyenlítődésre, ellentétben azzal, ahogyan azt az átlagember gondolná [4].

A relatív gyakoriság alakulása tekintetében azonban a józan ész már ugyanazt sugallja, amit a tapasztalat mutat. Nevezetesen azt, hogy szabályos érmével dobva a fej dobások és az összes dobások aránya nagy ismétlésszám esetén az 1/2 körül ingadozik. Javasoljuk, hogy az előbb ténylegesen elvégzett kb. 100 dobás esetén számítsuk ki és ábrázoljuk a relatív gyakoriságokat.

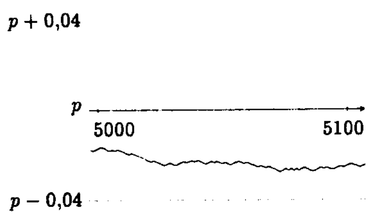
Ha már kialakult a valószínűség valamilyen „természetes" fogalma a diákokban, akkor térhetünk rá a számítógéppel szimulált „mesterséges" kísérletek bemutatására. Annyi feltehetően elfogadtatható, hogy a számítógép „képes utánozni" egy (szabályos vagy szabálytalan) érme egymás utáni feldobásait.

Egy adott p valószínűségű esemény előfordulási száma a kísérlet n ismétléséből legyen $k(n)$. A $k(n)/n$ hányadost nevezzük az esemény relatív gyakoriságának. Koordinátarendszerben ábrázoljuk az $(n, k(n)/n)$ pontokat, és kössük össze az egymás utáni pontokat szakaszokkal. Egy ilyen töröttvonalat (ún. trajektóriát) mutat az 1. ábra. Erről leolvasható, hogy már $n = 100$ ismétlés esetén a relatív gyakoriság a megadott p valószínűség körül ingadozik.



1. ábra

Ez a tendencia n növelésével egyre inkább jellemző. A 2. ábrán a trajektória $n = 5000$ és $n = 5100$ közötti szakasza látható. Ez a szakasz teljes egészében a p körüli 0,08 szélességű sávban fut. Érdemes hosszú trajektóriákat többször generálni, hogy megismerjük azok „tipikus viselkedését".



2. ábra

Az oktatónak természetesen tudnia kell, hogy a valószínűségszámítás matematikai modelljének kialakításához a relatív gyakoriság tapasztalati tulajdonságaiból kiindulva fektetjük le a valószínűség axiómáit, majd az ezekből kifejtett matematikai elméletben már tételként jelenik meg a relatív gyakoriság és a valószínűség kapcsolatát leíró nagy számok törvénye. A tapasztalati tények és az elméleti tételek viszonyának elemzésébe azonban ezen a szinten nem célszerű belemenni.

2. A bolyongás

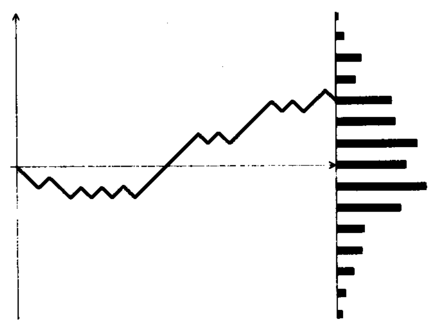
A bolyongást legegyszerűbben egy szerencsejátékkal illusztrálhatjuk. Dobjunk fel egy pénzérmét, s ha a fej lesz felül, akkor nyerünk egy forintot, ellenkező esetben veszítünk egy forintot. Játsszuk el ezt n-szer! A kérdés, hogy mennyi lesz a játék végén a nyereményünk. Jelöljünk η_n-nel a nyereményt. (Ha veszteségről van szó, akkor természetesen η_n értéke negatív.) Ha – a nem szabályos érme lehetőségére való tekintettel – egy dobásnál p valószínűséggel nyerünk és $q = 1 - p$ valószínűséggel veszítünk, akkor annak valószínűsége, hogy az n fordulós játékban $\rho_n = k$-szor nyerünk és ebből következően $(n - k)$-szor veszítünk

$$(1) \quad P(\rho_n = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \ldots, n.$$

Ez egyben annak a valószínűsége is, hogy $k - (n - k) = 2k - n$ a nyereményünk értéke. Nyilván

$$\eta_n = 2\rho_n - n$$

teljesül. (1) a jól ismert binomiális eloszlás.

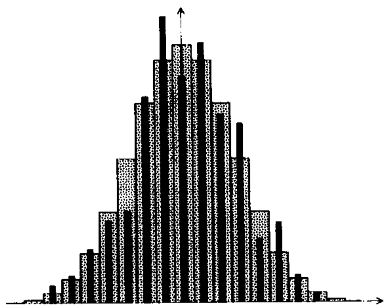


3. ábra

A bolyongást a fenti játékból úgy kaphatjuk, hogy kiindulunk az origóból, és +1-et lépünk minden fej és −1-et minden írás dobás esetén. A bolyongás trajektóriája (út–idő-diagramja) a (k, η_k), $k = 0, 1, \ldots, n$, pontokat összekötő töröttvonal. A bolyongás trajektóriájával együtt a bolyongás és a binomiális eloszlás kapcsolata is szemléltethető a számítógépen. A 3. ábra egy szimmetrikus (azaz $p = 1/2$) bolyongást szemléltet. (Az n-et körülbelül 5 és 30 között érdemes választani.)

Az n dobásból álló játékot ismételtük meg 200-szor, s közben számoltuk, hogy az $x = n$-ben húzott függőleges egyenes szóba jöhető pontjain hányszor „csapódott" be a trajektória. A becsapódások alakították ki a hisztogramot. A játék minden egyes indításakor a trajektóriát letöröltük, az ábrán így a 200. trajektória látható.

Ha a hisztogram téglalapjainak magasságát osztjuk a kísérletsorozatok számával (jelen esetben 200-zal), akkor a különböző nyeremények relatív gyakoriságait kapjuk, melyek a nagy számok törvénye alapján a valószínűségek körüli értékeket adják. Ennek a módosított hisztogramnak a képe látható a 4. ábrán, a háttérben a szürke oszlopok pedig az (1) által meghatározott elméleti eloszlás hisztogramja.



4. ábra

3. A centrális határeloszlás tétel

A centrális határeloszlás tételek a normális eloszlás szerepét mutatják. A centrális határeloszlás tétel alapján az η_n, bolyongás standardizáltja eloszlásban konvergál a standard normális eloszláshoz.
Képletben

$$\lim_{n\to\infty} P\left(\frac{\eta_n - n(p-q)}{\sqrt{4npq}} < x\right) = \Phi(x),$$

ahol

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, \quad -\infty < x < \infty,$$

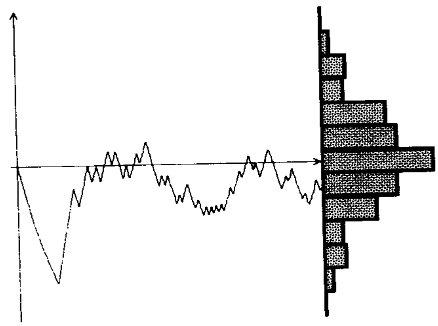
a standard normális eloszlásfüggvény. Így ha x-nek egy Δx sugarú környezetét vizsgáljuk, akkor

$$\lim_{n\to\infty} P\left(x - \Delta x \le \frac{\eta_n - n(p-q)}{\sqrt{4npq}} < x + \Delta x\right) = \int_{x-\Delta x}^{x+\Delta x} \varphi(t)\, dt, \tag{2}$$

ahol

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}, \quad -\infty < t < \infty,$$

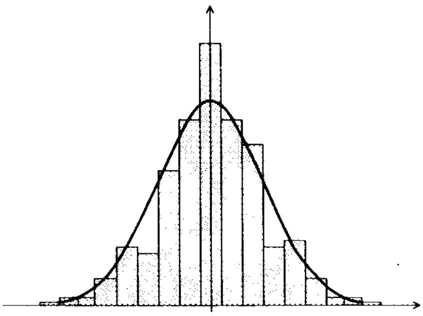
a standard normális sűrűségfüggvény, aminek a gráfját szokás haranggörbének nevezni.



5. ábra

Számítógéppel a (2) képletet szemléltetjük. (A programban n-t körülbelül 80 és 100 között érdemes megválasztani.) A standardizált bolyongás (azaz $\eta_n - n(p-q)/\sqrt{4npq}$) trajektóriáit ábrázoljuk. Hasonlóan a bolyongáshoz, itt is vizsgáljuk, hogy egy trajektória hol „csapódik be". A „becsapódási falat" osszuk fel $2\Delta x$ hosszúságú intervallumokra! Számoljuk meg, hogy a trajektóriák becsapódási pontjai hányszor esnek az egyes részintervallumokba, s az eredményt ábrázoljuk egy hisztogramon! Az 5. ábra a szimmetrikus bolyongás esetén 200 trajektóriából kialakult hisztogramot mutatja.

Ha a hisztogramot úgy normáljuk, hogy az oszlopok összterülete éppen 1 legyen, akkor a sűrűséghisztogramot kapjuk. Ekkor az egyes oszlopok területe éppen az illető intervallumba esés relatív gyakorisága. A (2) képlet alapján (nagy n és nagy ismétlésszám esetén) az egyes oszlopok területe közelítőleg a φ haranggörbe alatti területtel egyenlő. Ezt úgy tudjuk szemléltetni, ha a sűrűséghisztogramot összehasonlítjuk az elméleti haranggörbével. (Lásd 6. ábra!)



6. ábra

Megjegyezzük, hogy mind a három tárgyalt esetben a bolyongás trajektóriáit vizsgáltuk. Amikor ezt normálás nélkül tettük, akkor n növelésével a trajektóriák „szétterültek" (3. és 4. ábra). Amikor a nagy számok törvényének megfelelően normáltunk (azaz n-nel osztottunk), akkor a trajektória egy ponthoz (a várható értékhez) tartott (1. és 2. ábra). Amikor pedig a centrális határeloszlás tételnek megfelelően normáltunk (azaz $\sqrt{n}$ -nel osztottunk), akkor a trajektóriák n növelésével a haranggörbének megfelelő sűrűséggel haladtak. Igen fontos arra is rámutatni, hogy amíg a nagy számok törvénye egyetlen trajektóriával is szemléltethető (hiszen az erős törvény majdnem biztos konvergenciát állít), addig a centrális határeloszlás tételbeli törvényszerűség nagy számú „hosszú" trajektória által alakul csak ki (hisz ott eloszlásban való konvergenciáról van szó).

3. További témakörök

Az érdeklődő kollégáknak javasoljuk az alábbi tételek számítógépes szemléltetésének kialakítását is.

Az iterált logaritmus tétel [5]:

$$P\left(\overline{\lim_{n\to\infty}} \frac{\eta_n - n(p-q)}{\sqrt{4pq}\sqrt{2n\log\log n}} = 1\right) = 1$$

és

$$P\left(\underline{\lim_{n\to\infty}} \frac{\eta_n - n(p-q)}{\sqrt{4pq}\sqrt{2n\log\log n}} = -1\right) = 1 .$$

Ez a tétel durván azt jelenti, hogy a fenti módon normált trajektória a $(-1, 1)$ sávot „járja be".

Az arkusz szinusz törvények a szimmetrikus bolyongásra (l. [3], [4]). Az utolsó visszatérésre vonatkozó arkusz szinusz törvény szerint: egy szimmetrikus bolyongást végző részecskét $2n$ időpontig vizsgálva, annak valószínűsége, hogy a részecske utoljára a $2k$-adik időpillanatban tér vissza az origóba $\frac{\binom{2k}{k}\binom{2n-2k}{n-k}}{2^{2n}}$ $(k = 0, 1, \ldots, n)$. A hosszú vezetésre vonatkozó arkusz szinusz törvény szerint: egy szimmetrikus bolyongást végző részecskét $2n$ időpontig vizsgálva, annak valószínűsége, hogy a részecske $2k$ egységnyit a pozitív, $2n-2k$ egységnyit a negatív oldalon tölt $\frac{\binom{2k}{k}\binom{2n-2k}{n-k}}{2^{2n}}$, ahol $k = 0, 1, \ldots, n$. Ezen tételek alapján azt is mondhatjuk, hogy két, egyformán erős játékos esetén az, hogy az utolsó egyenlítés a játék közepe körül következzék be, illetve az, hogy a játék folyamán mindkét játékos nagyjából egyforma ideig vezessen, viszonylag kicsi.

A homogén szériák hossza. Erdős és Rényi egy tétele szerint egy szabályos érmével n-szer dobva, a leghosszabb „tiszta fej"-blokk hossza aszimptotikusan $\log_2 n$ ([1], [6]).

A szerzők szívesen fogadják a kollégák észrevételeit, ill. személyes tapasztalatait.

Irodalom

[1] Erdős P. – Révész P.: On the length of the longest head-run. Csiszár I. – Elias P. (editors): Topics in Information Theory. (Coll. Math. Soc. J. Bolyai, 16), 219-228, Norht-Holland, Amsterdam,1977.

[2] Fazekas I.: Bevezetés a valószínűségszámításba. (egyetemi jegyzet), Debrecen, 1992., Eger, 1993.

[3] Feller W.: Bevezetés a valószínűségszámításba és alkalmazásaiba. Műszaki Könyvkiadó, Budapest, 1978.

[4] Rényi A.: Napló az információelméletről. Gondolat, Budapest, 1976.

[5] Rényi A.: Valószínűségszámítás. Tankönyvkiadó, Budapest, 1973.

[6] Székely J. G.: Paradoxonok a véletlen matematikájában. Műszaki Könyvkiadó, Budapest, 1982.