Székelyhidi László

Valószínűségszámítás és matematikai statisztika

****

Budapest, 1998

Előszó

Ez a jegyzet a valószínűségszámításnak és a matematikai statisztikának azokat a fejezeteit tárgyalja, amelyek a matematikatanár szakos hallgatók képzésében szerepet játszanak, figyelembe véve a tanárképző intézmények tanterveit és a felhasználható alapismereteket. Mivel a valószínűségelmélet tárgyalása ezen a szinten nem lehetséges a Lebesgue-féle mértékelmélet alapján, így a jegyzetben az úgynevezett naiv felépítést követjük. Mindazonáltal igyekszünk azokat a legfontosabb fogalmakat és tételeket bemutatni (utóbbiakat esetenként bizonyítás nélkül), amelyek kellő mértékben reprezentálják a modern valószínűségelmélet és a matematikai statisztika gyakorlati alkalmazhatóságát, módszereit.

A valószínűségszámítással kapcsolatos fejezetek megírása során elsősorban az [5] és [10] munkákra támaszkodtunk, azok terminológiáját, jelölésrendszerét használtuk. A statisztikáról szóló fejezetek forrásmunkája főleg [2] és [9] volt. A feladatok összeállítása során elsősorban a [7], [9], [10] munkák feladatanyagát használtuk fel.

Az egyes fejezetekhez feladatok csatlakoznak, melyek a gyakorlás mellett az anyag mélyebb elsajátítását hivatottak elősegíteni.

A jegyzet írásakor erősen támaszkodtunk az irodalomjegyzékben felsorolt művekre, ezekből számos feladatot átvettünk, esetenként az adatok, illetve a fogalmazás kisebb módosításával.

TARTALOM

1. Bevezetés

1.1. Véletlen tömegjelenségek

A valószínűségszámítás tárgya a véletlen tömegjelenségek vizsgálata. A tapasztalat azt mutatja, hogy a gyakorlatban léteznek olyan jelenségek, folyamatok, melyek végbemenetelét nem határozzák meg egyértelműen azok a körülmények, amelyeket a jelenség, illetve folyamat tanulmányozásakor tekintetbe tudunk, illetve tekintetbe kívánunk venni. A gyakorlati életből számos ilyen példát sorolhatunk fel. Például, egy adott folyó vízállása egy adott pillanatban számos olyan tényezőtől függ, melyek mindegyikének figyelembevétele szinte lehetetlen és gyakorlati szempontból nem is lényeges. Az ilyen jelenségeket, eseményeket a továbbiakban véletlen jelenségeknek, véletlen eseményeknek, röviden *eseményeknek* nevezzük. A *véletlen tömegjelenség* olyan véletlen jelenség, mely több alkalommal lényegében azonos figyelembe vett körülmények között megismétlődik, illetve megismételhető.

1.2. Kísérletek, események

A véletlen tömegjelenségek egyes lefolyásainak megfigyeléseit véletlen kísérleteknek, röviden *kísérleteknek* nevezzük. Egy kísérletnek tehát általában több, esetleg végtelen sok lehetséges kimenetele van. Ha például a kísérlet abból áll, hogy egy szabályos dobókocka feldobásakor a felső lapon levő számot figyeljük, akkor ennek a kísérletnek 6 lehetséges kimenetele van. Két ilyen kocka feldobásakor a lehetséges kimenetelek száma 36. Egy folyó vízállásának megfigyelésekor a lehetséges kimenetelek száma annyi, ahány különböző vízállás elvileg elképzelhető, s mivel a vízállás elvileg bármilyen nem negatív szám lehet, a lehetséges kimenetelek száma végtelen. Természetesen a lehetséges kimenetelek száma akkor is végtelen, ha feltételezzük, hogy a vízállás valamilyen ésszerű felső korlát alatt marad.

Egy kísérlet lehetséges kimeneteleit *elemi eseményeknek* nevezzük. Egy kísérlettel kapcsolatban tehát először is azt kell eldöntenünk, hogy mit tekintünk elemi eseményeknek. Egy pénzérme feldobásakor elvileg az is elképzelhető, hogy az érme az élére esik, mégis, ennél a kísérletnél csak azt a két elemi eseményt szokás tekintetbe venni, hogy a dobás eredménye fej vagy írás. A kísérlettel kapcsolatos elemi események összességét *eseménytérnek* nevezzük és ebben a jegyzetben általában Ω-val jelöljük. Így például a kockadobási kísérlethez tartozó eseménytér a következő lehet: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Választhatnánk természetesen más eseményteret is, például lehetne $\Omega = \{a, b, c, d, e, f\}$, ahol az a, b, c, d, e, f betűk rendre az $1, 2, 3, 4, 5, 6$ számok dobását jelentik, ez azonban csak jelölésben tér el az előzőtől, s matematikailag célszerű minél egyszerűbben leírni az eseményteret. Két szabályos pénzérme egyidejű feldobásával (vagy ami ugyanaz, egy érme kétszer egymás utáni feldobásával) kapcsolatos kísérletnél az eseménytér a következő lehet: $\Omega = \{FF, FI, IF, II\}$, ahol például FF azt jelenti, hogy az egyes számú kockával és a kettes számú kockával is fejet dobtunk, FI azt, hogy az egyes számúval fejet, a kettes számúval pedig írást, stb.

Egy kísérlettel kapcsolatban az elemi eseményeken kívül általában más eseményeket is le lehet írni. A kockadobási kísérletnél például egy esemény az, hogy páros számot dobunk, a pénzfeldobás esetén pedig például az, hogy mindkét érmével ugyanazt dobjuk. Általában az elemi események bármely együttesét véletlen eseménynek, röviden *eseménynek* nevezzük. Akkor mondjuk, hogy a kísérlet elvégzésekor valamely esemény *bekövetkezik*, ha a kísérlet tényleges kimenetele egy olyan elemi esemény, amely az illető eseményhez tartozik. A kockadobási kísérletnél például az az esemény, hogy páros számot dobunk, akkor következik be, ha a kísérlet kimenetele 2, 4 vagy 6.

A fentiekben vázoltuk azokat az alapfogalmakat, amelyek a valószínűségszámítás kialakulásakor fogalmazódtak meg, ezek azonban nem tekinthetők precíz matematikai definícióknak. A matematikai absztrakció során elvonatkoztatunk a fenti fogalmak konkrét, gyakorlati jelentésétől, s kizárólag ezek jellemző tulajdonságaira és a köztük fennálló kapcsolatokra figyelünk. A valószínűségszámítás modern matematikai elmélete ugyanolyan axiomatikus alapokon nyugszik, mint a matematika egyéb elméletei. Az axiomatikus felépítés lényege az, hogy néhány (lehetőleg minél kisebb számú) úgynevezett alapfogalmat választunk, melyeket nem definiálunk, s a matematikai logika segítségével leírjuk ezek legfontosabb tulajdonságait, kapcsolatait, melyeket axiómáknak nevezünk. Az axiómák tekinthetők bebizonyítatlan tételeknek. Egy matematikai diszciplína axiomatikus felépítésekor természetesen alapvető fontosságú, hogy az axiómák logikailag ellentmondásmentesek legyenek, ennek ellenőrzése azonban általában nem könnyű feladat. Másrészt célszerű minél kevesebb axiómát választani az elmélet könnyebb áttekinthetősége kedvéért, ami úgy érhető el, hogy az axiómák logikai függetlenségére törekszünk: lehetőleg egyetlen axiómát se lehessen logikai úton levezetni a többiből. További követelményeket is szokás az axiómarendszerekkel szemben támasztani, ennek részleteivel azonban itt nem foglalkozunk. Visszatérve a valószínűségszámításra, ennek modern elmélete a halmazelmélet axiomatikus elméletén és az úgynevezett mértékelméleten alapul, felépítése A. N. Kolmogorov[1] nevéhez fűződik. Mivel itt nincs mód arra, hogy ezeket az elméleteket akár ismertnek tételezzük fel, akár részletesen tárgyaljuk, a valószínűségszámítás egy egyszerűsített, úgynevezett naiv felépítését fogjuk követni, mely matematikailag eléggé precíz, a gyakorlati alkalmazások számára eléggé általános, de a tárgyaláshoz szükséges előismeretek nem haladják meg az elemi halmazelmélet, algebra és kombinatorika szintj ét. Ebben a megközelítésben a fentiekben vázolt valószínűségelméleti fogalmaknak egy – egy halmazelméleti fogalom, illetve reláció felel meg. E gondolatok részletes kifejtésére a 3. fejezetben kerül sor.

Amint említettük, a számunkra szükséges technikai apparátus nem lép túl az elemi halmazelmélet, algebra és kombinatorika szintjén. A szükséges halmazelméleti, illetve algebrai alapismeretekkel a hallgatók más tárgyak keretein belül részletesen megismerkednek, itt csupán azokat a kombinatorikai fogalmakat foglaljuk össze, melyek a jegyzetben sűrűn előfordulnak, s esetükben célszerűnek látszik a jelölések, terminológia stb. rögzítése. Mint látni fogjuk, a gyakorlati alkalmazásokban fellépő klasszikus valószínűségi mezőkön a különféle problémák megoldása általában kombinatorikai módszerekkel történik.

[1] Andrej Nyikolajevics Kolmogorov szovjet matematikus, 1903–1987

2. A kombinatorika elemei

A kombinatorika a véges halmazok elméletével foglalkozik, s napjainkban is dinamikusan fejlődő tudományága a matematikának, számos alkalmazással. A valószínűségszámítás szempontjából elsősorban az úgynevezett csoportalkotási, illetve elrendezési problémák érdekesek, így a következőkben ezek legfontosabb típusait tekintjük át. A csoportalkotási problémák olyan kérdésekkel kapcsolatosak, hogy egy véges számú elemet tartalmazó halmazból például hogyan és hányféleképpen lehet kiválasztani bizonyos számú elemet bizonyos előírásoknak megfelelően, míg az elrendezési problémáknál egy véges halmaz elemei különböző szempontok szerinti elrendezéseinek száma iránt érdeklődünk.

2.1. Permutációk

Legyen adott egy n elemű halmaz. Általában feltételezhetjük, hogy a halmaz elemei természetes számok, ez nyilván nem jelenti az általánosság korlátozását. Ha a halmaz elemeit valamilyen módon sorba rendezzük, akkor azt mondjuk, hogy a halmaz elemeit *permutáljuk*. Egy ilyen elrendezést a halmaz elemei egy *permutációjának* nevezünk. Attól függően, hogy a halmaz elemeit mind megkülönböztetjük egymástól, vagy közülük bizonyosakat azonosnak tekintünk, beszélünk *ismétlés nélküli*, illetve *ismétléses* permutációkról. Az, hogy a halmaz bizonyos elemei között nem teszünk különbséget, annyit jelent, hogy az ilyen elemek egymás közötti cseréjével nyert elrendezéseket nem tekintjük különböző permutációknak. Legyen például a halmaz $\{1, 2, \ldots, n\}$. Ez egyben az úgynevezett természetes elrendezés, melyben az elemek nagyság szerint vannak sorba rendezve. Ez egy permutáció, melyet egyszerűen $(1, 2, \ldots, n)$ módon jelölünk. Egy másik lehetséges permutációt kapunk, ha az első két elemet felcseréljük: $(2, 1, 3, 4, \ldots, n)$. A kérdés: hogyan kapható meg az n elem összes (ismétlés nélküli) permutációinak P_n száma? Ha az első helyen a halmaz bármelyik elemét lerögzítjük, akkor az ezzel az elemmel kezdődő permutációk száma nyilván annyi, ahányféleképpen a fennmaradó $n-1$ elemet permutálhatjuk, azaz P_{n-1}. Mivel az első helyre n különböző elemet rögzíthetünk, ezért nyilván

$$P_n = nP_{n-1}$$

érvényes, hacsak $n \geq 2$. Mivel nyilván $P_1 = 1$, a fenti egyenletből teljes indukcióval kapjuk n elem ismétlés nélküli permutációinak számát:

$$P_n = n(n-1)(n-2)\dots 2 \cdot 1 = n!$$

Egy egyszerű példa: 8 ember között hányféleképpen lehet 8 különböző gyümölcsöt szétosztani úgy, hogy mindenki pontosan egyet kapjon? Ha a 8 embert egymás mellé sorba állítjuk, akkor nyilván annyi különböző szétosztásra van lehetőség, ahányféle különböző sorrendje van a 8 gyümölcsnek, azaz $P_8 = 40320$.

A fenti probléma úgy módosítható, hogy van n darab golyó, melyek közül k darab egyforma, a többi $n-k$ pedig ezektől és egymástól is különböző. Ilyen feltétel mellett hány különböző elrendezés lehetséges? Ha ezen elrendezések számát $P_n^{(k)}$ jelöli, akkor vizsgáljuk meg, mi a kapcsolat $P_n^{(k)}$ és P_n között. Nyilvánvaló, hogy a $P_n^{(k)}$ számú permutáció mindegyikéből $P_k = k!$ számú ismétlés nélküli permutációt kaphatunk, ha a k egyforma golyót mégis különbözőnek tekintjük, így $k!P_n^{(k)} = P_n$, amiből adódik, hogy

$$P_n^{(k)} = \frac{n!}{k!}.$$

Hasonló gondolatmenetet alkalmazhatunk, ha az n elem között több egyenlő elemből álló csoport is előfordul. Legyen ezen csoportok elemszáma $k_1, k_2, \dots, k_m$, ahol tehát $k_1 + k_2 + \dots + k_m = n$ és $k_i \geq 1$, $i = 1, 2, \dots, m$, akkor az ismétléses permutációk száma

$$P_n^{(k_1,k_2,\dots,k_m)} = \frac{n!}{k_1!k_2!\dots k_m!}.$$

A fentieket illusztrálja a következő egyszerű példa: hány különböző tipposzlopot tudunk 13 + 1 mérkőzést tartalmazó totószelvényen kitölteni úgy, hogy 8 mérkőzést 1-esre, 4-et x-re és 2-t 2-esre tippelünk? A választ a fenti képletből kapjuk: $P_{14}^{(8,4,2)} = \frac{14!}{8!4!2!} = 45045$.

A következő problémák csoportok különböző feltételek melletti képzésével kapcsolatosak.

2.2. Kombinációk

Legyen adott n különböző elem, melyekből k elemű csoportokat képezünk. Nyilván feltételezzük, hogy $1 \leq k \leq n$. Egy ilyen csoportot az n elem egy k-ad osztályú kombinációjának nevezünk. Kérdés: hány különböző ilyen kombináció képezhető? Ha ezen kombinációkban minden elem legfeljebb egyszer szerepelhet, akkor *ismétlés nélküli kombinációkról* beszélünk, ha többször is, akkor *ismétléses kombinációkról.* Természetesen egy – egy kombináción belül az elemek sorrendje nem számít.

Először határozzuk meg az *n elem k-ad osztályú ismétlés nélküli kombinációinak* C_n^k számát. Ha n elem közül k darabot kiválasztunk, akkor $n-k$ darabot nem választunk ki, tehát a keresett ismétlés nélküli kombinációk száma megegyezik az n elem olyan ismétléses permutációinak számával, melyek között k, illetve $n-k$ darab egyforma, azaz

$$C_n^k = P_n^{(k,n-k)} = \frac{n!}{k!(n-k)!}.$$

Erre a képletre gyakori előfordulása miatt külön jelölést szokás alkalmazni: $\binom{n}{k}$, mely így olvasandó: "n alatt a k". Formálisan az $\binom{n}{k}$ úgynevezett *binomiális együttható* csak $1 \le k < n$ értékekre van értelmezve, a $0! = 1$ megállapodással azonban minden $0 \le k \le n$ értékre kiterjeszthetjük. Ekkor pl. $\binom{n}{0} = 1$. Az $\binom{n}{k}$ binomiális együttható valójában minden $0 \le k \le n$ értékre megadja egy n elemű halmaz összes k elemű részhalmazainak számát.

Az ismétlés nélküli kombinációk alkalmazására egy lehetséges példa a lottójátékkal kapcsolatos: hányféleképpen lehet szabályosan kitölteni egy ötöslottószelvényt? A kérdés nyilván az, hogy 90 elemből hányféleképpen lehet 5-öt kiválasztani, s ennek megfelelően a válasz: 90 elem 5-ödosztályú ismétlés nélküli kombinációinak száma, $\binom{90}{5} = 43949268$.

A következőkben az ismétléses kombinációkkal foglalkozunk. Legyen adott n különböző elem, s ezekből készítsünk k elemű csoportokat, de úgy, hogy egy – egy elemet akárhányszor felhasználhatunk. Ekkor tehát k lehet n-nél nagyobb is. Az ilyen csoportokat az *n elem k-ad osztályú ismétléses kombinációinak* nevezzük. Jelölje ezek számát $C_n^{k,i}$. Megmutatjuk, hogy ez a szám egyenlő $n+k-1$ elem ismétlés nélküli kombinációinak számával. Legyen az n elem $1, 2, \dots, n$, és legyen egy k-ad osztályú ismétléses kombináció a következő: $n_1, n_2, \dots, n_k$. Természetesen feltehetjük, hogy

$$1 \le n_1 \le n_2 \le \dots \le n_k \le n,$$

hiszen kombinációk esetén a különböző sorrendek között nem teszünk különbséget. Ehhez hozzárendelhetjük az n_1, n_2+1, $n_3+2, \dots, n_k+k-1$ sorozatot, amely nyilván az $1, 2, \dots, n+k-1$ elemek ismétlés nélküli kombinációja. Másrészt az is világos, hogy ezen a módon különböző ismétléses kombinációkból különböző ismétlés nélküli kombinációkat kapunk. Hasonló gondolatmenet alapján, ha $n_1, n_2, \dots, n_k$ az $1, 2, \dots, n+k-1$ elemeknek k-ad osztályú ismétlés nélküli kombinációja, melyről feltehetjük, hogy eleget tesz az

$$1 \le n_1 < n_2 < \dots < n_k \le n$$

feltételnek, akkor ebből elkészíthetjük az $1, 2, \ldots, n$ elemek egy k-ad osztályú ismétléses kombinációját, s különböző ismétlés nélküli kombinációkból itt is különböző ismétléses kombinációkat nyerünk. Ilyen módon az n elemből készíthető k-ad osztályú ismétléses, és az $n+k-1$ elemből készíthető k-ad osztályú ismétlés nélküli kombinációk száma megegyezik:

$$C_n^{k,i} = \binom{n+k-1}{k}.$$

A formula alkalmazására tekintsük a következő problémát: három egyforma, szabályos dobókockát feldobva hányféle eredményt kaphatunk, ha a kockákat nem különböztetjük meg egymástól? Mivel egy kockával hatféle számot dobhatunk, s a hat szám mindhárom kockán ismétlődve is szerepelhet, így hat elem harmad-osztályú ismétléses kombinációinak számát kell meghatározni, ami a fenti képlet alapján: $C_6^{3,i} = \binom{6+3-1}{3} = \binom{8}{3} = 56$.

2.3. Variációk

A csoportképzés egy másik lehetséges formája az, amikor n különböző elemből úgy választunk ki k darabot, hogy az elemek sorrendjét is figyeljük. Egy ilyen csoport neve *k-ad osztályú variáció*. Ha a csoportok képzésekor minden elemet csak egyszer használhatunk fel, akkor ismétlés nélküli, ha többször is, akkor ismétléses variációkról beszélünk. Az ismétlés nélküli variációk egyik leggyakoribb alkalmazása az úgynevezett *visszatevés nélküli mintavétel*, amin azt értjük, hogy egy n elemű halmazból egyenként kiveszünk k elemet, és ezeket a húzás sorrendjében egymás mellé rakjuk. Hasonlóan, ismétléses variációkat kapunk akkor, ha a kihúzás után az elem valamilyen azonosítóját (pl. számát) felírjuk, majd az elemet visszatesszük. Ez a *visszatevéses mintavétel*.

Most meghatározzuk n elem k-ad osztályú ismétlés nélküli variációinak V_n^k számát. Világos, hogy minden egyes k-ad osztályú ismétlés nélküli kombinációból $k!$ darab ismétlés nélküli variációt kaphatunk, ha a kiválasztott k elemet minden lehetséges módon sorba rendezzük. Így

$$V_n^k = k! C_n^k = k! \binom{n}{k} = \frac{n!}{(n-k)!} = n(n-1)\ldots(n-k+1).$$

Egy egyszerű példa a fentiek illusztrálására: négy különböző színű selyemből hányféle háromszínű zászlót lehet varrni, ha a zászlón a színek elrendezése a magyar zászlóéhoz hasonló? A kérdés nyilván 4 elem 3-ad osztályú variációinak száma, amely a fentiek szerint: $V_4^3 = 4 \cdot 3 \cdot 2 = 24$.

Végül, az n elemből képezhető k-ad osztályú ismétléses variációk $V_n^{k,i}$ számának meghatározása a következőképpen történhet: az $1, 2, \ldots, n$ elemek mindegyikét párosítjuk az összes $(k-1)$-ed osztályú ismétléses variációval. Tehát $V_n^{k,i} = nV_n^{k-1,i}$. Mivel nyilván $V_n^{1,i} = n$, ebből indukcióval adódik, hogy

$$V_n^{k,i} = n^k.$$

Az ismétléses variációk alkalmazására egy példa: legalább hány tipposzlopot kellene kitöltenünk a totón ahhoz, hogy biztosan legyen $13+1$ találatunk? Mivel a kérdés a 3 elemből $(1, 2, x$-ből) képezhető $13 + 1 = 14$ tagú sorozatok számára vonatkozik, a válasz: $V_3^{14,i} = 3^{14} = 4782969$.

2.4. Feladatok

2.1. Bizonyítsuk be a következő azonosságokat:

i) $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$;

ii) $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n$;

iii) $\sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j} = \binom{m+n}{k}$;

iv) $\binom{n}{j}\binom{n-j}{l} = \binom{n}{j+l}\binom{j+l}{l}$.

2.2. Hozzuk egyszerűbb alakra a következő kifejezéseket, és számítsuk ki numerikus értéküket:

$$\frac{P_8}{P_6}; \qquad \frac{P_{10}+P_8}{P_7}; \qquad \frac{8P_7-P_8}{P_{10}};$$
$$\frac{9!+10!}{8!-6!}; \qquad 1+\frac{P_9}{P_8}; \qquad \frac{P_6-4!}{3!}.$$

2.3. Oldjuk meg a következő egyenletet:

$$\frac{n!}{3!(n-3)!} = 4\frac{(n-2)!}{2!(n-4)!}, \qquad (n>3).$$

2.4. Oldjuk meg a következő egyenletet:

$$\frac{(n+1)!}{n!} + 7 = 4(n-1).$$

2.5. Határozzuk meg a következő kifejezések numerikus értékét:

$$\frac{V_8^3+V_7^4}{V_6^3}; \qquad \frac{V_7^4-V_6^3}{V_7^3-V_6^2}; \qquad \frac{V_7^4-P_5}{V_5^2}.$$

2.6. Hozzuk egyszerűbb alakra a következő kifejezéseket:

$$\frac{V_n^5+V_n^4}{V_n^2}; \qquad \frac{V_n^{k+2}+V_n^{k+1}}{V_n^k}.$$

2.7. Oldjuk meg a következő egyenletet:

$$\frac{x!}{V_4^2(x-2)!} = 4\frac{(x-1)!}{V_4^3(x-3)!} \qquad (x \neq 2, x \neq 3).$$

2.8. Oldjuk meg a következő egyenletet:

$$\frac{V_x^2}{2} + 4 = \frac{V_{x+1}^2}{2} - 2.$$

2.9. Oldjuk meg a következő egyenletet:

$$\binom{n}{n-3} \cdot 3! = \binom{n-2}{2} V_5^2.$$

2.10. Oldjuk meg a következő egyenleteket:

$$\binom{x}{2} + 4 = \binom{x+1}{2} - 2;$$

$$\binom{x}{2} + \binom{x-1}{2} = \frac{(x+1)!}{x!} + 4;$$

$$\binom{x-1}{2} + \frac{(x-1)!}{(x-2)!} = x - 1 + \frac{V_x^2}{2} - 2.$$

2.11. Hány olyan hatjegyű telefonszámot alkothatunk az $1, 3, 4, 6, 8, 9$ számjegyekből, amelyben a harmadik jegy 4-es, és minden számjegy csak egyszer fordul elő?

2.12. Hány hétjegyű szám állítható elő a $0, 1, 2, 3, 4, 5, 6$ számjegyekből, ha egy számjegy csak egyszer szerepelhet a számban?

2.13. Az A, B, C, D, E, F, G betűkből hány olyan permutáció képezhető, amelyben a magánhangzók egymás mellett szerepelnek?

2.14. Hány szó alkotható a KOMBINATORIKA szóból?

2.15. Hány ötjegyű szám írható fel a $2, 2, 2, 4, 4$ számjegyekből?

2.16. Egy pénzérmét nyolcszor egymás után feldobunk. Hányféle olyan dobássorozat van, amelyben 4 fej és 4 írás fordul elő?

2.17. Egy sakkversenyen 16 sakkozó vesz részt. Körmérkőzést játszanak, mégpedig úgy, hogy minden pár kétszer mérkőzik, világos és sötét színekkel felváltva. Hány mérkőzésre kerül sor a versenyen?

2.18. Hányféleképpen lehet egy dobókockát négyszer egymás után úgy feldobni, hogy legalább az egyik dobás 6-os legyen?

2.19. Hány különböző lottószelvényt tölt ki az, aki három számot állandónak vesz minden szelvényen?

2.20. A raktárban 15 üveg bor van: 10 üveg fehér és 5 üveg vörös bor. Hányféleképpen választhatunk ki 6 üveget úgy, hogy köztük kettőben legyen vörös bor?

2.21. Három egyforma kockát feldobunk. Hányféle eredményt kaphatunk, ha a kockákat nem különböztetjük meg?

2.22. 20 termék közül, melyek között 8 bizonyos "a", a többi "nem a" tulajdonságú, kiválasztunk 5-öt. Hány olyan választás létezik, melyben 2 db "a" és 3 db "nem a" tulajdonságú?

2.23. A $3, 5, 6, 8, 9$ számjegyekből legalább hány számjegyből álló számokat kell felírnunk, hogy legalább 625 különböző számot kapjunk?

2.24. Hány négyjegyű, 2000-nél nem nagyobb számot képezhetünk az $1, 2, 3, 4, 5$ számjegyekből, ha

i) a számjegyek ismétlődését nem engedjük meg;

ii) a számjegyek ismétlődését megengedjük?

2.25. Az egyik áruházba egy árucikkből naponta 6 egyforma nagyságú láda érkezik. Egy-egy láda minősülhet I., II. vagy III. osztályúnak. Hányféle lehetőség adódik az áru minőség szerinti napi eloszlására?

2.26. Egy üzletben 20 eladó közül 11 szakképzett, 9 szakképzetlen. Hányféleképpen lehet közülük 8 tagú csoportot kiválasztani úgy, hogy a csoportban a szakképzetlenek száma legfeljebb 2 legyen?

2.27. Egy kockával öt egymás utáni dobásból álló sorozatot dobunk. Hány olyan dobássorozat lehet, amelyben egy darab 1-es és egy darab 2-es dobás fordul elő (a dobások sorrendjét is figyelembe véve)?

3. Eseményalgebra

Amint azt az első fejezetben jeleztük, a következőkben a valószínűségelméletet igyekszünk úgynevezett axiomatikus módon tárgyalni olyan mélységben, amely csupán azt tételezi fel, hogy az olvasó rendelkezik bizonyos elemi halmazelméleti, illetve algebrai és analízisbeli ismeretekkel. A tárgyalás is csak részben lesz axiomatikus: célunk az, hogy az olvasót matematikai alapokon ismertessük meg ennek az elméletnek az alapjaival és legfontosabb alkalmazásaival.

3.1. Eseménytér

Az elemi valószínűségelmélet alapvető matematikai struktúrája az *eseményalgebra*.

Legyen Ω rögzített, nem üres halmaz. Az Ω halmazt *eseménytérnek*, elemeit *elemi eseményeknek* nevezzük. Az Ω részhalmazait *eseményeknek* nevezzük. Így az elemi események azonosíthatók az Ω egyelemű részhalmazaival, tehát maguk is események. Azokat az eseményeket, amelyek nem elemi események, *összetett eseménynek* nevezzük. Az üres halmazt *lehetetlen eseménynek*, az Ω-t *biztos eseménynek* nevezzük, s néha az előbbit O-val, az utóbbit I-vel jelöljük. Általában vegyesen fogjuk használni a halmazelméleti és a valószínűségelméleti terminológiát, illetve jelölésrendszert. Ha az ω elemi esemény adott, A pedig tetszőleges esemény, akkor azt mondjuk, hogy az A esemény *bekövetkezik*, ha $\omega \in A$. Ha A, B tetszőleges események, akkor azt mondjuk, hogy az A esemény *maga után vonja* a B eseményt, ha $A \subset B$. Vagyis az A esemény pontosan akkor vonja maga után a B eseményt, ha valahányszor A bekövetkezik, bekövetkezik B is. Így például a lehetetlen esemény minden eseményt maga után von, és minden esemény maga után vonja a biztos eseményt. Világos, hogy két esemény pontosan akkor egyenlő, ha kölcsönösen maguk után vonják egymást, azaz egyszerre következnek be. A következőkben a halmazelméleti műveletek mintájára értelmezünk néhány, az események között végezhető műveletet. A továbbiakban feltételezzük, hogy a szóban forgó események egy rögzített Ω eseménytér eseményei.

3.2. Műveletek eseményekkel

Ha A tetszőleges esemény, akkor azt az eseményt, amely pontosan akkor következik be, ha az A nem következik be, az A esemény *ellentettjének* nevezzük, és $\overline{A}$-sal jelöljük. Halmazelméleti nyelven $\overline{A}$ az A komplementere. Nyilvánvalóan $\overline{\overline{A}} = A$. Világos továbbá, hogy a lehetetlen esemény és a biztos esemény egymás ellentettjei.

Ha A és B két tetszőleges esemény, akkor ezek $A + B$ *összegén* azt az eseményt értjük, amely pontosan akkor következik be, ha az A és B valamelyike (esetleg mindkettő) bekövetkezik. Halmazelméleti nyelven $A + B$ az A és B halmazok egyesítése. Az összeadás művelete értelemszerűen terjeszthető ki tetszőleges számú (esetleg végtelen sok) eseményre.

Ha A és B két tetszőleges esemény, akkor ezek AB *szorzatán* azt az eseményt értjük, amely pontosan akkor következik be, ha az A és B mindegyike bekövetkezik. Halmazelméleti nyelven AB az A és B halmazok metszete. A szorzás művelete ugyancsak kézenfekvő módon terjeszthető ki tetszőleges számú (esetleg végtelen sok) eseményre.

Ha A és B két tetszőleges esemény, akkor ezeket *egymást kizáró* eseményeknek nevezzük, ha szorzatuk a lehetetlen esemény. Halmazelméleti nyelven ez nyilván azt jelenti, hogy az A és B halmazok diszjunktak. A lehetetlen esemény például nyilván bármely más eseménnyel együtt egymást kizáró eseménypárt alkot. Az elemi események is egymást kizáró események. Megjegyezzük, hogy kettőnél több esemény esetében azt szokás mondani, hogy ezek *egymást páronként kizárják*, ha közülük bármely kettő egymást kizárja. Ha $A_1, A_2, \ldots, A_n$ egymást páronként kizáró események, és összegük a biztos esemény, akkor azt mondjuk, hogy az $A_1, A_2, \ldots, A_n$ események *teljes eseményrendszert* alkotnak. Például bármely esemény az ellentettjével együtt teljes eseményrendszert alkot. Egy általánosabb értelmezésre visszatérünk a 4. fejezetben.

Ha A és B két tetszőleges esemény, akkor az A és B események $A-B$ *különbségén* azt az eseményt értjük, amely pontosan akkor következik be, ha az A esemény bekövetkezik, a B pedig nem. Halmazelméleti nyelven $A - B$ az A és B halmazok (halmazelméleti) különbsége, vagyis az A halmaz és a B komplementerének metszete. Ilyen módon $A - B = A\overline{B}$. Világos, hogy bármely A esemény esetén $\overline{A} = I - A$, azaz, az A esemény ellentettje éppen a biztos esemény és az A különbsége.

Ha A és B két tetszőleges esemény, akkor az A és B események $A \circ B$ *szimmetrikus differenciáján* azt az eseményt értjük, amely pontosan akkor következik be, ha az

A és B közül pontosan az egyik következik be. Az $A+B$ eseménnyel ellentétben tehát az $A \circ B$ nem következik be akkor, ha az A is és a B is bekövetkezik. A halmazelméleti szóhasználat itt tehát azonos a valószínűségszámítás nyelvezetével. Ha az összeadás műveletét a "megengedő vagy" logikai művelettel interpretáljuk, akkor a szimmetrikus differencia a "kizáró vagy" logikai műveletének felel meg. Például egymást kizáró események szimmetrikus differenciája nyilván egyenlő összegükkel.

A fentiekben értelmezett műveletek természetesen különböző algebrai tulajdonságokkal rendelkeznek. Ezen tulajdonságok némelyike azonos a megfelelő algebrai művelet hasonló tulajdonságával, de vannak lényeges eltérések. Az eseményekkel való formális "számolás" egyszerűsítése érdekében célszerű ezeket a műveleti szabályokat áttekinteni. Itt felsorolunk néhány egyszerű tulajdonságot, melyek sora természetesen folytatható. Ezek, és a hasonló műveleti szabályok bizonyítása a megfelelő halmazelméleti azonosságok bizonyításához hasonlóan történhet. Az alábbiakban feltételezzük, hogy a szereplő események mind egy rögzített eseménytérhez tartoznak.

Az összeadás és szorzás főbb azonosságai:

1) $A + A = A$ (idempotens),
2) $A + B = B + A$ (kommutatív),
3) $A + (B + C) = (A + B) + C$ (asszociatív),
4) $A + (B \cdot C) = (A + B) \cdot (A + C)$ (disztributív);

1') $A \cdot A = A$ (idempotens),
2') $A \cdot B = B \cdot A$ (kommutatív),
3') $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ (asszociatív),
4') $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$ (disztributív).

Az ellentett eseménnyel kapcsolatos főbb azonosságok:

5) $A + \overline{A} = I$,
6) $A + O = A$,
7) $A + I = I$;

5') $A \cdot \overline{A} = O$,
6') $A \cdot I = A$,
7') $A \cdot O = O$.

Látható, hogy a felsorolt azonosságokban az i)-vel és í)-vel jelzettek analógiát mutatnak, mégpedig ha az i)-ben minden eseményt az ellentettjére, minden +-t ·-ra és minden ·-t +-ra cserélünk, akkor megkapjuk í)-t. Ez az úgynevezett *dualitási elv*, mely a következő, úgynevezett *de Morgan*[2]-azonosságok következménye:

[2] Augustus de Morgan skót matematikus, 1806-1871

8) $\overline{A+B} = \overline{A} \cdot \overline{B}$,
8') $\overline{A \cdot B} = \overline{A} + \overline{B}$.

Illusztrációként bebizonyítjuk például 8)-at. Jelölje M, illetve N a 8) egyenlőség bal, illetve jobboldalán álló eseményt. Ha az M bekövetkezik, akkor a definíció szerint $A+B$ nem következik be, azaz sem A, sem B nem következik be. Másszóval, $\overline{A}$ és $\overline{B}$ mindegyike bekövetkezik, tehát bekövetkezik N. Megfordítva, ha N bekövetkezik, akkor definíció szerint bekövetkezik $\overline{A}$ és $\overline{B}$ mindegyike, tehát sem A, sem B nem következik be. Így nem következik be $A+B$ sem, ami azt jelenti, hogy $\overline{A+B}$, azaz M bekövetkezik. Ezzel megmutattuk, hogy M és N egyszerre következnek be, tehát egyenlők. Megjegyezzük, hogy a de Morgan-képletek alapján a dualitási elv minden eseményekkel kapcsolatos azonosságra alkalmazható.

A fentiekben felsorolt 1)-7) és 1')-7') azonosságok lehetővé teszik a valószínűségelmélet egy másfajta axiomatikus megalapozását, amely a halmazelméletre nem hivatkozik. Nevezetesen az olyan algebrai struktúrát, amelyben értelmezve van egy $(A,B) \mapsto A+B$ "összeadás", egy $(A,B) \mapsto A \cdot B$ "szorzás" és egy $A \mapsto \overline{A}$ "ellentett képzés", valamint van két speciális elem, az O és az I úgy, hogy teljesülnek az 1)-7) és 1')-7') azonosságok, *Boole*[3]*-algebrának* szokás nevezni. Az események összessége Boole-algebra, melyet *eseményalgebrának* nevezünk. A valószínűségelmélet általunk tárgyalandó változatának alapja egy ilyen eseményalgebra. Megjegyezzük, hogy ebben a megközelítésben az elemi események jelentősége elhalványul, hiszen kiindulópontunk nem egy eseménytér, hanem az összes események halmaza, az eseményalgebra.

Ezen a ponton egy fontos megjegyzést kell tennünk. A modern valószínűségelméletben az eseményalgebra szerepét egy általánosabb struktúra, az úgynevezett *σ-algebra* játssza. Ennek definíciója a következő. Legyen Ω egy tetszőleges, nem üres halmaz, $\mathcal{A}$ pedig az Ω részhalmazainak egy olyan összessége, mely rendelkezik az alábbi tulajdonságokkal:

i) az Ω beletartozik $\mathcal{A}$-ba: $\Omega \in \mathcal{A}$;
ii) bármely $\mathcal{A}$-beli halmaz komplementere is beletartozik $\mathcal{A}$-ba: ha $A \in \mathcal{A}$, akkor $\overline{A} \in \mathcal{A}$;
iii) az $\mathcal{A}$-ba tartozó halmazok bármely sorozatának egyesítése is beletartozik $\mathcal{A}$-ba: ha $A_1, A_2, \dots \in \mathcal{A}$, akkor $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Ekkor $\mathcal{A}$-t σ-algebrának nevezzük. Látható, hogy az eseményalgebra és a σ-algebra közti legfontosabb különbség az, hogy míg az eseményalgebra csupán véges sok esemény összeadására nézve zárt, a σ-algebra bármely *megszámlálhatóan végtelen* sok elemének egyesítését is tartalmazza. Az i) és ii) alapján $\emptyset$ is beletartozik $\mathcal{A}$-ba, s az ii) és iii) (valamint a de-Morgan-azonosságok) alapján az $\mathcal{A}$-ba

[3] George Boole angol matematikus, 1815-1864

tartozó halmazok bármely sorozatának metszete is beletartozik $\mathcal{A}$-ba. Ebben a megközelítésben Ω reprezentálja az eseményteret, s az Ω halmaznak csupán azok a részhalmazai tekintendők eseménynek, amelyek $\mathcal{A}$-ba tartoznak. Az általunk vizsgálandó esetekben leggyakrabban Ω egy véges halmaz, s ilyenkor a Boole-algebra és a σ-algebra fogalma egybeesik.

3.3. Feladatok

3.1. Határozzuk meg egy n elemi eseményt tartalmazó eseménytér összes eseményeinek számát, illetve összes összetett (nem elemi) eseményeinek számát!

3.2. Bizonyítsuk be, hogy bármely eseményalgebrában érvényesek az alábbi azonosságok:

i) $AA = A$;
ii) $A + A = A$;
iii) $AB = BA$;
iv) $A + B = B + A$;
v) $A(BC) = (AB)C$;
vi) $A + (B + C) = (A + B) + C$;
vii) $A(B + C) = AB + AC$;
viii) $A + BC = (A + B)(A + C)$.

3.3. Bizonyítsuk be, hogy bármely eseményalgebrában érvényesek az alábbi azonosságok:

i) $A\overline{A} = O$;
ii) $A + \overline{A} = I$;
iii) $AI = A$;
iv) $A + O = A$;
v) $AO = O$;
vi) $A + I = I$.

3.4. Bizonyítsuk be, hogy bármely eseményalgebrában érvényesek az alábbi azonosságok:

i) $A \subset A$;
ii) ha $A \subset B$ és $B \subset C$, akkor $A \subset C$;
iii) $O \subset A \subset I$;
iv) $AB \subset A \subset A + B$;
v) $\overline{A + B} = \overline{A}\,\overline{B}$;
vi) $\overline{AB} = \overline{A} + \overline{B}$;
vii) $A + B = A + (\overline{A}B)$;
viii) $A = AB + A\overline{B}$.

3.5. Bizonyítsuk be, hogy bármely eseményalgebrában érvényesek az alábbi azonosságok!

i) $A = A \cap (A \cup B)$;

ii) $A \cup B = A \cup (\overline{A} \cap B)$;
iii) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;
iv) $(A - B) - C = (A - C) - (B - C)$;
v) $(A \cap B) - C = (A - C) \cap (B - C)$;
vi) $A - (B \cup C) = (A - B) - C$.

3.6. Vizsgáljuk meg, hogy milyen, az A és B eseményekre tett feltevés mellett érvényesek az alábbi egyenlőségek!

i) $A = A \cup B$;
ii) $B = A \cup B$;
iii) $B = A \cap B$;
iv) $A = A \cap \overline{B}$;
v) $\overline{A} = A \cap B$;
vi) $\overline{A} = A \cap B$;
vii) $A \cup B = A \cap B$.

3.7. 10 darab terméket megvizsgálunk abból a szempontból, hogy hány selejtes darabot találunk köztük. Mik lesznek a kísérletet leíró eseménytér pontjai?

3.8. Négy pénzérme ismételt feldobásakor figyeljük a *fej* dobások számát. Mi lesz az eseménytér?

3.9. Egy kocka és egy pénzérme együttes feldobásakor hogyan reprezentálható az eseménytér?

3.10. Az előző feladat feltételei mellett adjuk meg a következő eseményeket elemi események halmazaként: fej és páros szám; fej, valamint írás és prímszám; írás és páratlan szám.

3.11. Feldobunk egy ötforintost, egy tízforintost és egy kockát. Írjuk le az eseményteret az elemi események felsorolásával!

3.12. Egy adott időponttól kezdve megfigyeljük egy telefonközpontba bizonyos idő alatt beérkező hívások számát. Adjuk meg a kísérletet leíró eseményteret!

3.13. Feldobunk egy érmét. Ha az eredmény fej, akkor még egyszer dobunk, ha írás, akkor még kétszer. Mik lesznek a kísérletet leíró eseménytér pontjai?

3.14. Egy pénzdarabot kétszer egymás után feldobunk. A kísérletnél összesen
i) hány elemi esemény értelmezhető;
ii) hány esemény értelmezhető?

3.15. Egy cukrászdában három hűtőpult működik. Jelentse A_i azt az eseményt, hogy az i-edik hűtőpult egy éven belül elromlik ($i = 1, 2, 3$). Fejezzük ki az A_i eseményekkel a következőket:
i) csak az első romlik el;
ii) mindhárom elromlik;
iii) egyik sem romlik el;
iv) az első és második nem romlik el;
v) az első és második elromlik, a harmadik nem;
vi) csak egy hűtőpult romlik el;
vii) legfeljebb egy hűtőpult romlik el;
viii) legfeljebb két hűtőpult romlik el;
ix) legalább egy hűtőpult elromlik.

3.16. Jelentse A azt az eseményt, hogy egy gyár 3 százaléknál alacsonyabb selejttel dolgozik; B azt, hogy 2 százaléknál alacsonyabb a selejt. Mit jelent az $A - B$, $B - A$, AB, $A + B$ esemény?

3.17. A 32 lapos magyar kártyából kihúzunk négy lapot. Jelöljük rendre A,B,C,D-vel azt az eseményt, hogy a kihúzott lapok között van piros, zöld, tök, makk. Fejezzük ki A, B, C, D, illetve $\overline{A}, \overline{B}, \overline{C}, \overline{D}$ segítségével a következő eseményeket: a kihúzott lapok között
i) csak piros van;
ii) csak piros és zöld van;
iii) vagy csak piros, vagy csak zöld van.

3.18. Egy üzletbe porszívókat szállítanak. Az átvétel során három készüléket ellenőrzünk. Az A esemény azt jelenti, hogy legalább egy készülék hibás, a B esemény jelentése pedig, hogy mindhárom készülék kifogástalan. Vizsgáljuk meg, mit jelentenek a következő események: $A + B$, AB, $A - B$, $A \circ B$!

3.19. Jelentse A azt az eseményt, hogy 10 kockadobásból egyszer sem kaptunk 6-ost. Mit jelent az $\overline{A}$?

3.20. Feldobunk két játékkockát.
i) A kísérletnek hány lehetséges kimenetele van?
ii) A kísérletnél összesen hány esemény értelmezhető?

4. A valószínűség matematikai fogalma

4.1. Gyakoriság, relatív gyakoriság

Egy véletlen jelenséggel kapcsolatos kísérlet egyszeri elvégzésekor általában semmiféle előzetes bizonyosságunk nem lehet a kísérlet kimenetelét illetően, hiszen mint mondottuk, a véletlen jelenségeknek az a sajátossága, hogy a számba vehető körülmények a jelenség lefolyását nem határozzák meg egyértelműen. Látszólag ugyanez a helyzet akkor is, ha a kísérletet nemcsak egyszer, hanem többször egymás után, egymástól függetlenül elvégezzük: a különböző kimeneteleknek elvileg bármilyen sorozatát megkaphatjuk. A tapasztalat azonban azt mutatja, hogy nagy számú kísérlet elvégzésekor mégis felfedezhető valamilyen törvényszerűség. Azt tapasztaljuk ugyanis, hogy ha az egyes kimenetelek *gyakoriságait* figyeljük, akkor a különböző események bekövetkezési gyakoriságainak hányadosa stabilitási tendenciát mutat. Egy adott kísérletsorozatban az A esemény bekövetkezéseinek gyakorisága legyen k_A: ez az a szám, ahányszor az A esemény a kísérletsorozatban bekövetkezett. Ha B egy másik esemény, akkor azt tapasztaljuk, hogy a kísérletek számának növelésével a $\frac{k_A}{k_B}$ hányados aránylag kis ingadozásokat mutat. Ennek alapján – a tapasztalattal megegyezően – célszerű feltételezni, hogy a kísérletek számának minden határon túli növelésekor a $\frac{k_A}{k_B}$ hányados egy meghatározott, csupán az A és B eseménytől függő, bizonyos – később pontosan definiálandó – értelemben vett "határértékhez" tart. Ezt a meghatározott számértéket a $B = I$ esetben az A esemény *valószínűségének* nevezzük. Így az A esemény valószínűsége a $\frac{k_A}{k_I}$ hányadosnak, vagyis az A esemény *relatív gyakoriságának* a később meghatározandó értelemben vett "határértéke". Meg kell jegyeznünk, hogy itt semmi esetre sem szabad a közönséges analízisbeli határérték fogalmára, illetve konvergenciára gondolnunk. A "bizonyos értelemben vett határérték" fogalmának pontos definíciójához szükséges a későbbiekben (ld. 15.2. pont) bevezetésre kerülő sztochasztikus konvergencia fogalma. Ha a kísérletet összesen n-szer végeztük el, akkor nyilván $k_I = n$, s ha ebből az A esemény k-szor következett be, akkor $k_A = k$, így az A esemény relatív gyakorisága $\frac{k}{n}$. Itt n a kísérlet összes elvégzéseinek száma, k pedig az A esemény bekövetkezéseinek száma. Vegyük észre, hogy míg egy esemény gyakorisága és relatív gyakorisága függ a kísérletek számától, sőt, általában magától a kísérletsorozattól is, a valószínűség csupán az eseménytől függ. Másrészt, míg a gyakoriság és a relatív gyakoriság egy adott kísérletsorozatban egzakt módon meghatározott, kiszámítható numerikus adat, ezzel szemben a valószínűség egy "ideális" érték, amelynek létezését csupán feltételezhetjük, amelyet ugyan a tapasztalat általában nem cáfol, de egy esemény valószínűségének a fentiekben adott értelmezése semmiképpen sem tekinthető matematikai értelemben vett definíciónak.

Vizsgáljuk meg a relatív gyakoriság néhány egyszerű tulajdonságát. Először is nyilvánvaló, hogy bármely esemény relatív gyakorisága 0 és 1 közé esik. Az is világos, hogy a biztos esemény relatív gyakorisága 1. Tegyük most fel, hogy A és B két egymást kizáró esemény az adott kísérletsorozatban. Ekkor az $A + B$ esemény gyakorisága nyilván megegyezik az A és B gyakoriságainak összegével, s ugyanez igaz a megfelelő relatív gyakoriságokra is. Mivel a valószínűség a relatív gyakoriság bizonyos értelemben vett határértéke, ezért az említett három tulajdonságot célszerű a valószínűséggel kapcsolatban is feltételezni, tehát bármely esemény valószínűsége 0 és 1 közé esik, a biztos esemény valószínűsége 1, végül egymást kizáró események összegének valószínűsége egyenlő valószínűségeik összegével.

4.2. Valószínűségi mező

A fenti heurisztikus bevezető után célszerű lenne a valószínűség pontos matematikai definícióját megadni. Ezt ismét axiomatikus módon tesszük meg, figyelembe véve a fenti gondolatmenetet.

Legyen adott egy eseményalgebra, és tételezzük fel, hogy ennek minden A eseményéhez hozzá van rendelve egy $P(A)$-val jelölt valós szám a következő tulajdonságokkal:

1) $0 \leq P(A)$,
2) $P(I) = 1$,
3) ha $A \cdot B = O$, akkor $P(A + B) = P(A) + P(B)$.

Ezt a $P(A)$ számot az A esemény valószínűségének nevezzük, az eseményalgebrát a valószínűséggel együtt pedig *valószínűségi mezőnek.* Ily módon tehát a valószínűség egy olyan valós értékű függvénynek tekinthető az eseményalgebrán, amely rendelkezik a fenti három tulajdonsággal, melyek a *valószínűség axiómái.* Ezt a függvényt szokás *valószínűségeloszlásnak* is nevezni. A harmadik tulajdonságot a valószínűség *additivitásának* nevezzük, s teljes indukcióval azonnal következik, hogy két esemény helyett tetszőleges véges számú eseményre is fennáll, amennyiben azok egymást páronként kizárják.

Amint azt az előző fejezet végén megjegyeztük, az általunk vizsgálandó esetekben leggyakrabban Ω egy véges halmaz, s ilyenkor a Boole-algebra és a σ-algebra fogalma egybeesik. Bizonyos esetekben azonban lényeges, hogy egyrészt az eseménytér lehet végtelen is, másrészt az eseménytér nem minden részhalmaza esemény, így nem is lehet a valószínűségéről beszélni. Amennyiben az események összessége σ-algebra, a valószínűség 3) axiómájának helyét a következő veszi át:

3') ha A_1, A_2, ... ,$A_n, \ldots$ egymást páronként kizáró események véges vagy végtelen sorozata, akkor

$$P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Ezt a tulajdonságot a valószínűség *teljes additivitásának* nevezzük, és a továbbiakban 1) és 2) mellett a 3') fennállását tételezzük fel, amennyiben nem feltétlenül véges eseménytérrel kapcsolatos valószínűségi mezőről beszélünk. A valószínűségi mező matematikai definíciója tehát a következő: legyen Ω egy nem üres halmaz, $\mathcal{A}$ az Ω részhalmazainak egy σ-algebrája, P pedig egy valószínűségi függvény az $\mathcal{A}$-n. Ekkor az $(\Omega, \mathcal{A}, P)$ hármast *valószínűségi mezőnek* nevezzük. Az Ω halmazt *eseménytérnek*, az $\mathcal{A}$-beli halmazokat *eseményeknek*, a $P(A)$ számot pedig az A esemény *valószínűségének* nevezzük. Mint azt korábban említettük, az általunk vizsgált esetekben leggyakrabban az Ω egy véges halmaz lesz, $\mathcal{A}$ pedig az Ω összes részhalmazaiból álló σ-algebra, de látni fogunk példákat végtelen eseményterekre is, melyeknek nem minden részhalmaza esemény.

Pozitív valószínűségű események (véges vagy végtelen) sorozatát *teljes eseményrendszernek* nevezzük, ha egymást páronként kizárják, és valószínűségeik összege 1. Általánosabban, egymást páronként kizáró események egy tetszőleges halmazát szokás teljes eseményrendszernek nevezni, ha összegük a biztos esemény. Például egy tetszőleges eseménytéren az összes elemi események teljes eseményrendszert alkotnak.

A fentiek illusztrálása céljából tekintsünk egy egyszerű példát: a kockadobással kapcsolatos kísérletet. Mint a korábbiakban láttuk, ebben az esetben az eseménytér azonosítható az $\Omega = \{1, 2, 3, 4, 5, 6\}$ halmazzal, tehát az elemi események az $1, 2, 3, 4, 5, 6$ számok. Az eseményalgebra eseményei az Ω összes részhalmazai. Például az az esemény, hogy prímszámot dobunk, azonosítható a $\{2, 3, 5\}$ részhalmazzal; az az esemény, hogy 1-nél nagyobb és 2-nél kisebb számot dobunk, azaz, a lehetetlen esemény pedig az üres halmazzal. Ahhoz, hogy valószínűségi mezőt kapjunk, értelmeznünk kell az Ω összes részhalmazain egy valós értékű függvényt, amely rendelkezik az 1) - 3) tulajdonságokkal. Könnyű belátni, hogy a következő definíció megfelelő: ha A egy tetszőleges részhalmaza Ω-nak, akkor $P(A)$ jelentse az A elemszámának $\frac{1}{6}$-od részét. Világos, hogy ilyen módon valószínűségi mezőt kapunk, melyben az elemi események egyformán valószínűek: minden elemi esemény valószínűsége $\frac{1}{6}$. Ez összhangban van azzal a tapasztalati ténnyel, hogy egy szabályos dobókocka szabályos feldobásának sokszori ismétlésekor az egyes számok relatív gyakoriságai egyre kevésbé térnek el $\frac{1}{6}$ - tól, pontosabban az ingadozások egyre kisebbek. Fontos dolog azonban megjegyezni, hogy eseményalgebránkon számtalan különböző módon értelmezhetünk valószínűséget. Legyen például $P(A) = 1$, ha az A tartalmazza a 6-ost, egyébként 0. Könnyen ellenőrizhető, hogy ez a függvény is rendelkezik a kívánt tulajdonságokkal, s így egy másik valószínűségi mezőt kaptunk. Hogyan lehetne ezt a valószínűségi mezőt gyakorlatilag "interpretálni"? A válasz egyszerű: ez egy olyan "teljesen szabálytalan" dobókockával kapcsolatos kísérletet ír le, amellyel csak 6-ost lehet dobni! Az persze más kérdés, hogy egy ilyen "csaló " dobókockának a gyakorlatban nem sok jelentősége lehetne, de elvileg elképzelhető ilyen kocka készítése, például úgy, hogy a 6-ossal átellenes lapot valamilyen erősen mágneses anyagból készítjük és a kockadobást vasból készült asztalon hajtjuk végre... Ezzel a példával azt akartuk bemutatni, hogy egy adott eseményalgebrán általában számos különböző valószínűségi mező értelmezhető, melyek elméleti szempontból egyformán tanulmányozhatók. Természetesen a gyakorlati alkalmazásoknál azt a valószínűségi függvényt célszerű választani, amely az illető problémát a lehető legjobban modellezi. Ennél a példánál maradva könnyen leírhatjuk az eseményalgebránkon értelmezhető összes lehetséges valószínűségeloszlást. Vegyük észre, hogy tetszőleges esemény valószínűségét egyértelműen meghatározzák a benne foglalt elemi események valószínűségei, hiszen azok egymást páronként kizárják, összegük az illető esemény, így a valószínűség additivitási tulajdonsága felhasználható. Másrészt az elemi események valószínűségei 0 és 1 között tetszőlegesen választhatók úgy, hogy összegük 1 legyen. Ezért eseményalgebránkon bármely valószínűségeloszlás egyértelműen jellemezhető azzal a rendezett $(p_1, p_2, p_3, p_4, p_5, p_6)$ szám hatossal, melynek elemei rendre az $1, 2, 3, 4, 5, 6$ elemi események valószínűségei. Az összes lehetséges valószínűségeloszlást tehát a következőképpen nyerhetjük: legyen $(p_1, p_2, p_3, p_4, p_5, p_6)$ egy nem negatív komponensű vektor, melyre $p_1+p_2+p_3+p_4+p_5+p_6 = 1$ teljesül, és legyen a P valószínűség definíciója tetszőleges A esemény esetén:

$$P(A) = \sum_{i \in A} p_i,$$

ahol az üres szumma zérusnak tekintendő. Könnyű belátni, hogy így tetszőleges véges eseménytéren leírhatjuk az összes lehetséges valószínűségeloszlásokat.

4.3. Feladatok

4.1. Bizonyítsuk be, hogy bármely valószínűségi mezőn bármely esemény valószínűsége legfeljebb 1!

4.2. Végezzük el 10 alkalommal a következő kísérletsorozatot: 10-szer dobjunk fel egy szabályos pénzérmét. Vizsgáljuk az egyes kísérletsorozatokban a fej, ill. az írás dobásának gyakoriságát, ill. relatív gyakoriságát!

4.3. Legyen adott az $\{1, 2, 3\}$ alaphalmaz, s tekintsük az ennek összes részhalmazaiból álló $\mathcal{A}$ eseményalgebrát. Adjunk meg ezen legalább három különböző valószínűségi függvényt!

4.4. Tekintsük a természetes számok halmazának mint alaphalmaznak az összes részhalmazaiból álló $\mathcal{A}$ σ-algebrát. Írjuk le ezen az összes lehetséges valószínűségi függvényt!

4.5. Legyen $\Omega = \mathbb{R}$, a valós számok halmaza, és jelölje $\mathcal{A}$ az Ω mindazon részhalmazainak összességét, amelyek vagy megszámlálhatók, vagy komplementerük megszámlálható. Bizonyítsuk be, hogy $\mathcal{A}$ σ-algebra!

4.6. Az előző feladatban értelmezett $\mathcal{A}$ σ-algebrán értelmezzük a P függvényt a következő módon: bármely $\mathcal{A}$-beli A esetén legyen $P(A) = 0$, ha A megszámlálható és $P(A) = 1$, ha A nem megszámlálható. Bizonyítsuk be, hogy P valószínűségi függvény, tehát $(\Omega, \mathcal{A}, P)$ valószínűségi mező!

4.7. Legyen adott az előző két feladatban értelmezett valószínűségi mező. Adjunk meg az $\mathbb{R}$-nek olyan részhalmazát, amely nem esemény!

4.8. Olyan súlyeloszlású kockánk van, amelynél 1-től 6-ig bármely pontszám dobásának valószínűsége arányos az illető pontszámmal. Írjuk le a megfelelő valószínűségi mezőt!

5. A klasszikus valószínűségi mező

A valószínűségszámítás kialakulása a szerencsejátékokkal kapcsolatos. A klasszikus valószínűség definícióját Moivre[4] adta meg 1711-ben: *"Egy esemény valószínűsége egyenlő az esemény bekövetkezésére nézve kedvező esetek és az összes esetek számának hányadosával, feltéve, hogy ezek az esetek egyenlően valószínűek."* Visszagondolva a relatív gyakorisággal kapcsolatban leírtakra, ez valójában annak a definíciója, ha a "kedvező esetek" helyett "bekövetkezések számá"-t, "összes esetek" helyett pedig "összes kísérletek számá"-t mondunk. A Moivre-féle "definíció " valójában tétel – egy speciális valószínűségi mező esetében.

Legyen $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ véges eseménytér, jelölje $\mathcal{A}$ az Ω összes részhalmazainak családját, és a P valószínűségi függvényt értelmezzük a következő módon: legyen minden elemi esemény valószínűsége ugyanaz a p szám. A valószínűség additivitása alapján a biztos esemény valószínűsége tehát szükségképpen np, így $np = 1$, azaz $p = \frac{1}{n}$. Ha A egy olyan esemény, amely k számú elemi eseményt tartalmaz, akkor ismét az additivitás alapján kapjuk, hogy

$$P(A) = \frac{k}{n},$$

ami éppen a Moivre-féle formula. Az olyan valószínűségi mezőt, mely véges számú elemi eseményt tartalmaz, amelyek egyenlően valószínűek, *klasszikus valószínűségi mezőnek* nevezzük. Klasszikus valószínűségi mezőben tehát egy esemény valószínűsége megegyezik a benne foglalt elemi események számának és az összes elemi esemény számának hányadosával. Ahhoz, hogy egy esemény valószínűségét meghatározhassuk, csupán a benne foglalt elemi események számát és az összes elemi esemény számát kell ismerni. A probléma arra redukálódik, hogy különböző véges halmazok elemszámát kell meghatározni, mely az esetek többségében kombinatorikai megfontolásokkal végezhető el. Ezek a megfontolások esetenként igen komplikáltak is lehetnek, egy – egy gyakorlati probléma megfogalmazásakor ugyanis nem mindig világos azonnal, hogy a matematikai modell megalkotásakor miket kell elemi eseményeknek tekinteni, s ezekből miként épülnek fel a kérdéses események. Az ilyen problémák megoldására jól felhasználhatók a 2. fejezetben összefoglalt kombinatorikai alapismeretek. Az alábbiakban néhány egyszerű példán illusztráljuk a módszer alkalmazását.

[4] Abraham de Moivre angol matematikus, 1667-1754

5.1. Véletlen húzás

Egy urnában hat golyó van $1, 2, 3, 4, 5, 6$ számjegyekkel ellátva. Az urnából véletlenszerűen kihúzunk egy golyót. Mennyi a valószínűsége annak, hogy a kihúzott golyón

i) négyes számot találunk;
ii) páratlan számot találunk;
iii) ötnél kisebb számot találunk?

Nyilvánvaló, hogy ennél a problémánál az eseménytér elemi eseményei az egyes kihúzott golyókkal, illetve a golyókon talált számokkal, másszóval az $1, 2, 3, 4, 5, 6$ számokkal azonosíthatók. Tehát az eseménytér: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Mint véges eseményterek esetén általában, $\mathcal{A}$ az összes részhalmazt tartalmazza: az Ω minden részhalmaza véges sok elemi esemény egyesítése, tehát maga is esemény. Fontos megjegyezni, hogy ebben a szituációban még végtelen sok lehetőség van arra, hogy $\mathcal{A}$-n valószínűséget értelmezzünk, hiszen elegendő az $1, 2, 3, 4, 5, 6$ számokhoz tetszőleges nem negatív valós számokat hozzárendelni úgy, hogy összegük 1 legyen. A feladatban megfogalmazott "véletlenszerűen kihúzunk egy golyót" kifejezés azonban burkoltan azt fejezi ki, hogy feltételezzük, hogy bármely golyó kihúzásának valószínűsége egyforma. Ez tehát csupán egy feltételezés, amit az támaszt alá, hogy a feladat megfogalmazásában semmi nem utal arra, hogy ezzel a feltételezéssel ne éljünk, s e feltételezés nélkül a probléma határozatlan lenne. Tehát klasszikus valószínűségi mezővel van dolgunk, melyben az elemi események egyformán $\frac{1}{6}$ valószínűségűek. Jelölje A_i az i-es golyó kihúzását, $i = 1, 2, \ldots, 6$. Ezért az a) kérdésre a válasz: $P(A_4) = \frac{1}{6}$, a b) kérdésre: $P(A_1 + A_3 + A_5) = P(A_1) + P(A_3) + P(A_5) = 3 \cdot \frac{1}{6} = \frac{1}{2}$, s a c) kérdésre: $P(A_1 + A_2 + A_3 + A_4) = P(A_1) + P(A_2) + P(A_3) + P(A_4) = 4 \cdot \frac{1}{6} = \frac{2}{3}$.

5.2. Több kocka egyidejű feldobása

Három kockát dobunk fel egyszerre. Mennyi a valószínűsége, hogy a dobott számok összege 8 lesz?

Először azt kell eldöntenünk, hogy ennél a problémánál mik az elemi események. Mivel a három dobókockán egymástól függetlenül kijöhet az 1, 2, 3, 4, 5, 6 számok bármelyike, az elemi eseményeknek az olyan rendezett számhármasok tekinthetők, melyek elemei ezek a számok. Hány ilyen van? A kérdés: 6 különböző elemnek hány 3-ad osztályú ismétléses variációja van? A válasz: $V_6^{3,i} = 6^3 = 216$. Feltételezzük, hogy ezek mindegyike egyforma valószínűséggel fordulhat elő, ami tehát $\frac{1}{216}$. Most azt kell megvizsgálnunk, hogy az az esemény, hogy a dobott számok összege 8, tehát az, hogy egy adott fenti tulajdonságú rendezett számhármas komponenseinek összege 8, hány elemi eseményt tartalmaz. Ez ekvivalens a következővel: hányféleképpen lehet a 8-at három 1 és 6 közé eső egész szám összegeként előállítani, ha az összeadandók sorrendjét is tekintetbe vesszük? Ezek a lehetőségek könnyen összeszámolhatók. Ha két kockán 1-es van, akkor a harmadikon már csak 6-os lehet, s ez háromféleképpen fordulhat elő. Ha egy kockán van csak 1-es, akkor a másik kettőn vagy 2-es és 5-ös, vagy 3-as és 4-es lehet, melyek mindegyike hatféleképpen fordulhat elő, hiszen az 1-estől különböző két számot hatféleképpen lehet elhelyezni a fennmaradó két helyen. Ez eddig $3+6+6=15$ lehetőség. Most már csak a 8-nak háromtagú, 2 és 6 közé eső egész számok összegeként való előállításainak számát kell meghatározni. Ha a 2-es két helyen szerepel, akkor a harmadik helyen már csak 4-es lehet, amire három lehetőség van. Ha a 2-es csak egy helyen szerepel, akkor a másik két helyen 3-asnak kell állni, amire szintén három lehetőség van. Ez tehát újabb $3+3=6$ lehetőség. Több lehetőség nincs, hiszen 3 és 6 közé eső három szám összegeként nem lehet a 8-at előállítani. Így a kérdéses esemény $15+6=21$ elemi eseményt tartalmaz, tehát valószínűsége: $\frac{21}{216}$.

5.3. Egy mintavételi probléma

Legyen adott egy N elemű sokaság, mely s darab I. típusú és $N-s$ darab II. típusú elemet tartalmaz. Vegyünk e sokaságból véletlenszerűen egy n elemű mintát. Mi a valószínűsége, hogy e minta pontosan k darab I. típusú elemet tartalmaz? Ez a példa számos gyakorlati alkalmazásnál előfordul. Ha például a sokaság elemei bizonyos termékek, s az I. típusúak a selejtesek, a II. típusúak a kifogástalanok, akkor a kérdés nyilván úgy fogalmazható, hogy mi a valószínűsége annak, hogy egy véletlenszerűen választott n elemű minta pontosan k selejtes terméket tartalmaz? Egy másik lehetséges interpretáció a következő: legyenek a sokaság elemei vizsgatételek, melyek közül a vizsgázó s darabot tud, a többit nem, s a kérdés az, hogy ha véletlenszerűen húz n tételt, akkor mi a valószínűsége, hogy ezek közül pontosan k tételt tud? A kérdés megválaszolásához tisztáznunk kell, hogy a mintavétel milyen módját alkalmazzuk. Tekintsük először a visszatevés nélküli mintavételt. Ez azt jelenti, hogy az n elemet egyszerre emeljük ki a sokaságból, vagy, ami ugyanaz, a mintadarabokat egyenként emeljük ki a sokaságból, ezeket nem tesszük vissza, és a húzások sorrendjét is figyelembe vesszük. Természetesen feltételezzük, hogy bármely minta kiemelése azonos valószínűségű. Az összes lehetséges minták, tehát az összes elemi események száma megegyezik azzal a számmal, ahányféleképpen N elemből ki lehet választani n darabot, s ez éppen: $C_N^n = \binom{N}{n}$. Most összeszámoljuk, hogy az az esemény, hogy a minta pontosan k darab I. típusú elemet tartalmaz, hány elemi eseményből áll. Hogy egy adott minta ilyen legyen, ahhoz az s darab I. típusúból kell k darabot választanunk, ami $\binom{s}{k}$-féleképpen lehetséges, és minden egyes ilyen kiválasztáshoz $\binom{N-s}{n-k}$-féleképpen választhatunk $(n-k)$ darabot a II. típusból. Ez tehát összesen $\binom{s}{k} \cdot \binom{N-s}{n-k}$. A keresett valószínűség tehát visszatevés nélküli mintavétel esetén

$$P = \frac{\binom{s}{k}\binom{N-s}{n-k}}{\binom{N}{n}}.$$

Ezek után tekintsük a visszatevéses mintavétel esetét. Ebben az esetben a kiemelt mintadarabot minden húzás után visszahelyezzük a sokaságba. Ekkor nyilvánvaló, hogy az összes elemi események száma éppen az N elem n-ed osztályú ismétléses variációinak $V_N^{n,i}$ számával egyenlő, hiszen az n elemű minta bármely eleme lehet az N elem bármelyike. A pontosan k darab I. típusú elemet tartalmazó minták számát a következőképpen kapjuk meg: a minta elemeinek k helyére az s darab I. típusúból $V_s^{k,i}$-féleképpen kell választanunk, s ezek mindegyikéhez az $(N-s)$ darab II. típusúból V_{N-s}^{n-k}-féleképpen. Mivel az első típusúak k darab helyét $\binom{n}{k}$-féleképpen határozhatjuk meg, így az összes lehetőségek száma: $\binom{n}{k} V_s^{k,i} V_{N-s}^{n-k}$, s a keresett valószínűség

$$P = \frac{\binom{n}{k} V_s^{k,i} V_{N-s}^{n-k}}{V_N^{n,i}} = \frac{\binom{n}{k} s^k (N-s)^{n-k}}{N^n}.$$

Ezt a formulát egyszerűbb alakban is felírhatjuk, ha bevezetjük a $p = \frac{s}{N}$ "selejtarányt":

$$P = \binom{n}{k} p^k (1-p)^{n-k}.$$

Vegyük észre, hogy p éppen annak a valószínűsége, hogy az N elemű sokaságból egyetlen elemet véletlenszerűen kiválasztva, az I. típusú lesz.

5.4. Feladatok

5.1. Mennyi a valószínűsége annak, hogy egy kockával dobva
i) legalább négyet dobunk;
ii) legfeljebb négyet dobunk;
iii) párosat dobunk;
iv) ötnél kisebb páratlan számot dobunk;
v) kettőnél nagyobb páros, vagy ötnél kisebb páratlan számot dobunk?

5.2. Feldobunk egy érmét. Ha az eredmény fej, még egyszer dobunk, ha írás, még kétszer. Mennyi a valószínűsége, hogy összesen egy fejet dobunk?

5.3. Egy egyetem 500 hallgatója közül 300 olvas német nyelven, 200 olvas angolul, 50 olvas franciául, 20 olvas németül és franciául, 30 olvas angolul és franciául, 20 olvas németül és angolul, 10 olvas mindhárom nyelven. Ha találomra kiválasztunk a hallgatók közül egyet, akkor mennyi a valószínűsége annak, hogy az illető
i) mindhárom nyelven olvas;
ii) angolul olvas;
iii) az angolon kívül a másik két nyelven nem tud olvasni?

5.4. Egy kockát egymás után hatszor feldobunk. Mennyi a valószínűsége annak, hogy
i) az $1, 2, 3, 4, 5, 6$ számok mindegyike felülre kerül;
ii) az első dobás eredménye 6-os, a többi pedig ettől különböző;
iii) az első két dobás eredménye 6-os, a többi pedig 6-tól és egymástól is különböző;
iv) két dobás eredménye 6-os, a többi pedig ettől különböző?

5.5. Négy pénzdarabot dobunk fel egyszerre. Mennyi a valószínűsége annak, hogy
i) mind a négy fej lesz;
ii) kettő fej, kettő írás lesz;
iii) legalább az egyik fej lesz?

5.6. Három kockát dobunk fel egyszerre. Mennyi a valószínűsége, hogy a dobott számok összege 9 lesz?

5.7. Egy dobozban 3 piros golyó van. Hány fehér golyót kell hozzátenni, hogy annak valószínűsége, hogy először fehér golyót húzunk, nagyobb legyen 0.9-nél?

5.8. Egy urnában 4 piros, 3 fehér és 2 zöld golyó van. Mennyi annak a valószínűsége, hogy ha egyszerre két golyót kihúzunk, azok egyforma színűek lesznek?

5.9. Mennyi a valószínűsége annak, hogy n ember között nincs két olyan, akinek ugyanazon a napon van a születésnapja?

5.10. 50 termékből, melyek között 5 selejtes található, találomra kiválasztunk 5-öt.

i) Mennyi a valószínűsége, hogy ezek között 2 selejtest találunk, ha a mintavétel visszatevéssel történik?

ii) Megváltozik-e az előbbi valószínűség, ha a mintavétel 100 termékből történik, változatlan selejtarány mellett?

5.11. Mi annak a valószínűsége, hogy egy véletlenszerűen kitöltött ötöslottó-szelvényen $0, 1, 2, 3, 4, 5$ találatunk lesz?

5.12. Egy felvonó a földszintről 7 személlyel indul felfelé. Az épület 10 emeletes. Ha minden emeleten egyforma számú szoba van, és mindenkit lifttel szállítanak, akkor feltehető, hogy az utasok azonos valószínűséggel szállnak ki bármelyik emeleten. Mennyi a valószínűsége annak, hogy a hét utas közül egyik emeleten se száll ki egynél több?

5.13. Elhelyezünk három dobozba nyolc tárgyat úgy, hogy az egyes tárgyakat megkülönböztethetőnek tekintjük. Mennyi a valószínűsége, hogy az egyes dobozokba rendre $2, 4, 2$ tárgy kerül?

5.14. Egy üzletben három pénztárhoz véletlenszerűen 10 vásárló érkezik. Mennyi annak a valószínűsége, hogy az első pénztárhoz 4, a másodikhoz és a harmadikhoz pedig $3 - 3$ vásárló kerül?

5.15. Egy 9 tagú társaság felszáll a három kocsiból álló villamosra; a nagy tolongásban a társaság minden tagja csak azt nézi, hogy feljusson valamelyik kocsira, és nem törődik azzal, hogy társai melyik kocsiba szállnak. Mennyi a valószínűsége, hogy mind a három kocsiba a társaság $3 - 3$ tagja jut?

5.16. Egy vendéglő egyik asztalánál 9 vendég ül. Összesen három üveg sört, négy tésztát és két kávét rendelnek. (Mindegyik vendég csak egy ételt vagy italt rendel.) Pincérünk emlékszik arra, hogy miből mennyit kell hoznia, de teljesen elfelejtette, hogy mit kinek kell adnia. Találomra szétosztja, amit hozott; mennyi annak a valószínűsége, hogy mindenki azt kapja, amit kért?

5.17. Egy dobozban 10 szál gyufa van. Találomra kiveszünk közülük néhányat. Mennyi a valószínűsége annak, hogy a kivett gyufaszálak száma páros?

5.18. Egy három házaspárból álló társaság táncol. Az egyes párokra oszlás egyenlően valószínű. Mennyi a valószínűsége annak, hogy egy bizonyos pillanatban senki se táncol a saját feleségével?

5.19. Egy háziasszony négy vendégének feketekávét készít, és ízlésüknek megfelelően a kávékba $0, 1, 2$, illetve 3 szem cukrot tesz. Mire a felkevert kávékat beviszi, már elfelejti, melyik kié. Mennyi a valószínűsége, hogy egyik vendég sem jut a kívánt édességű kávéhoz?

5.20. Egy 1000 darabot tartalmazó alapsokaságból 50 darabból álló mintát veszünk. Mennyi a valószínűsége, hogy a mintában egy selejtes darabot találunk, ha a selejtes darabok aránya 2%?

5.21. Egy urnában 5 fehér, 2 fekete és 3 vörös golyó van. Mennyi a valószínűsége annak, hogy az urnából 6 golyót kihúzva azok között 2 fehér, 1 fekete és 3 vörös legyen? (A húzás egyszerre, vagy ami vele ekvivalens, egymás után visszatevés nélkül történik.)

5.22. Mennyi a valószínűsége annak, hogy egy véletlenszerűen kitöltött totószelvény - oszlop 10, 11, 12, illetve 13 találatos lesz?

5.23. Két fiú és három lány sakkversenyen vesz részt. A nyerés valószínűsége nemenként azonos, de minden fiú kétszer esélyesebb, mint bármelyik lány. Mennyi annak a valószínűsége, hogy a sakkversenyt lány nyeri?

5.24. Egy rekeszbe rendezetlenül lett behelyezve 20 nő és 10 férfi személyi nyilvántartási lapja. A férfiaknak is és a nőknek is a fele mérnök. Mi annak a valószínűsége, hogy egy véletlenül kiemelt lapon egy férfi, egy mérnök, illetve egy férfi mérnök adatai vannak?

6. Geometriai valószínűségek

6.1. Geometriai valószínűségek értelmezése

A valószínűség kiszámításának klasszikus képlete csak véges eseménytér esetében alkalmazható, hiszen ha az eseménytér elemi eseményeinek száma végtelen, akkor a képlet semmitmondó. Ezért a korábbiakban tárgyalt módszerekkel nem tudunk választ adni olyan *geometriai valószínűségekkel* kapcsolatos kérdésekre, mint a következő: valószínűsége annak, hogy a $(0, 1)$ intervallumban véletlenszerűen kiválasztott szám az intervallum első harmadába esik, vagy egy céltáblába csapódó lövedék a 6-os körön belülre talál, stb. Az ilyen jellegű problémák megválaszolásához természetesen valamilyen feltevéssel kell élnünk. Az első esetben például ésszerűnek látszik a feltételezés: annak valószínűsége, hogy a kiválasztott szám a $(0, 1)$ intervallum valamely részintervallumába esik, arányos e részintervallum hosszával. A második példában ez tűnne ésszerű feltételezésnek: annak valószínűsége, hogy a találat a céltábla valamely résztartományába esik, arányos a résztartomány területével. Általában ilyen esetekben tehát azzal a feltételezéssel élünk, hogy az illető esemény valószínűsége arányos az eseményt reprezentáló geometriai alakzat mértékével (hosszával, területével, térfogatával stb.). Természetesen az ilyen jellegű feltevés nem mindig állja meg a helyét a gyakorlatban. Például a céltáblára leadott célzott lövés esetében nem jogos feltételezni a valószínűség egyenletes eloszlását. Számos esetben azonban ezt az egyenletes eloszlást feltételezhetjük, s alkalmazásával jó eredményeket kaphatunk.

Összefoglalva, ha az eseményteret valamely Ω geometriai alakzat (intervallum, síkbeli, illetve térbeli tartomány stb.) reprezentálja, és feltételezzük, hogy az elemi eseményeket reprezentáló pontok egyenletesen helyezkednek el az alakzaton, akkor annak valószínűsége, hogy egy véletlenszerűen kiválasztott pont a tartomány valamely A résztartományába esik, egyenlő az A résztartomány és az Ω tartomány geometriai mértékeinek hányadosával. Természetesen az egyes problémák tárgyalása során ügyelni kell arra, hogy a probléma természetének legjobban megfelelő geometriai mértéket válasszuk, ugyanis más és más mérték választása esetén ugyanis más és más modellt kaphatunk. A következőkben ismertetünk néhány klasszikus problémát, melyek megoldása geometriai valószínűségek meghatározására vezet. Megjegyezzük, hogy a geometriai valószínűség fogalma szoros kapcsolatban van a későbbiekben tárgyalandó egyenletes eloszlással.

6.2. Háromszög szerkeszthetősége

A $(0, 1)$ intervallumban véletlenszerűen kiválasztunk két pontot. Mennyi a valószínűsége, hogy az így keletkezett három szakaszból háromszög alkotható? A két kiválasztott pontot jelölje x és y. Mivel ezek mindegyike a $(0, 1)$ intervallum tetszőleges pontja, ezért az elemi események tere a sík $0 < x < 1$ és $0 < y < 1$ feltételeknek eleget tevő (x, y) pontokból álló négyzettel reprezentálható. Vajon ezen négyzet mely pontjai alkotják azt a tartományt, mely a háromszög szerkeszthetősége szempontjából kedvező esetek halmazát reprezentálja? Három szakaszból pontosan akkor szerkeszthető háromszög, ha bármely kettő hosszának összege nagyobb a harmadik hosszánál. Két esetet különböztetünk meg: $x < y$, vagy $y \le x$. Az első esetben az előbbi feltétel azt jelenti, hogy teljesülni kell az $x < 1 - x$, $y - x < x + 1 - y$ és $1 - y < y$ egyenlőtlenségeknek, a második esetben pedig x-et és y-t fel kell cserélnünk. Az első egyenlőtlenségek a sík következő tartományát jellemzik: $x < \frac{1}{2}$, $y < x + \frac{1}{2}$, $y > \frac{1}{2}$, melynek területe nyilván $\frac{1}{8}$, s hasonlóan, a másik három egyenlőtlenség által jellemzett terület is $\frac{1}{8}$, s mivel az egységnégyzet területe 1, a keresett valószínűség $\frac{1}{8} + \frac{1}{8} = \frac{1}{4}$.

6.3. Találkozási probléma

Ketten megbeszélik, hogy du. 5 és 6 óra között egy meghatározott helyen találkoznak. Megállapodnak abban, hogy aki korábban érkezik, 20 percet vár a másikra, azután elmegy. Mennyi a találkozás valószínűsége, ha mindketten véletlenszerűen érkeznek? Ez a probléma ekvivalens azzal, hogy a $(0, 1)$ intervallumban találomra kiválasztva két pontot mennyi a valószínűsége annak, hogy azok távolsága legfeljebb $\frac{1}{3}$? Ha a kiválasztott két pontot itt is x és y jelöli, akkor a lehetséges pontok összessége ismét az egységnégyzet, míg a kedvező pontok összessége az

$$|x - y| \le \frac{1}{3}$$

egyenlőtlenség által jellemzett tartomány, melynek területe $\frac{5}{9}$, ami a keresett valószínűség.

6.4. A Bertrand[5]-féle paradoxon

Egy r sugarú körnek véletlenszerűen kiválasztjuk valamely húrját. Mennyi annak a valószínűsége, hogy ez a húr hosszabb a körbe írt szabályos háromszög oldalánál? Jelölje x és y azt a két ívhosszúságot, melyeket a két metszéspontig kapunk, a körvonal valamely rögzített pontjától az óramutató járásával ellenkező irányban haladva. Annak feltétele, hogy a húr hosszabb legyen, mint a beírt szabályos háromszög oldala, a következő egyenlőtlenség:

$$\frac{2r\pi}{3} < |x - y| < \frac{4r\pi}{3}.$$

Az összes eseteket a $0 \le x < 2r\pi$, $0 \le y < 2r\pi$ négyzet pontjai reprezentálják, míg a kedvező eseteket a fenti egyenlőtlenséggel leírt tartomány pontjai, melyek területeinek aránya $\frac{1}{3}$, a keresett valószínűség.

A feladattal kapcsolatban azonban a következőképpen is okoskodhatunk: mivel a húr helyzetét középpontja egyértelműen meghatározza, s a húr hossza akkor hosszabb a beírt szabályos háromszög oldalánál, ha középpontja az ebbe a szabályos háromszögbe írt kör belsejében fekszik, a keresett valószínűség a két kör területeinek aránya: $\frac{1}{4}$.

E két megoldáson kívül még számos további is adható, melyek különböző valószínűségekhez vezetnek. A látszólagos ellentmondást a következőképpen lehet feloldani: a két megoldás a kör összes húrjainak halmazán más – más valószínűségeloszlást tételez fel. Ha a húrnak a körrel vett metszéspontjait reprezentáló síkbeli (x, y) pont a sík $0 \le x < 2r\pi$, $0 \le y < 2r\pi$ négyzetében egyenlő területű tartományokba egyenlő valószínűséggel esik, akkor az első megoldás a helyes. Ha viszont a húr középpontja a kör belsejében egyenlő területű tartományokba egyenlő valószínűséggel esik, akkor a második megoldás helyes.

[5] Joseph Louis François Bertrand francia matematikus 1822-1900

6.5. A Buffon[6]-féle tűprobléma

Egy r hosszúságú tűt véletlenszerűen rádobunk egy párhuzamos egyenesekkel vonalkázott síklapra, ahol a szomszédos egyenesek távolsága egy állandó d szám. Ha feltesszük, hogy $r \leq d$, akkor mennyi a valószínűsége annak, hogy a tű valamelyik egyenest metszi? A határozottság kedvéért meg kell állapodnunk a következőkben. Rögzítsük az egyenesek irányát és jelölje φ a tű és az egyenesek hajlásszögét, y pedig a tű középpontjának a visszafelé legközelebb eső egyenestől való távolságát. Ekkor $0 \leq \varphi \leq \pi$, $0 \leq y \leq d$. Az ilyen (φ, y) párok reprezentálják az elemi eseményeket. Ha feltételezzük, hogy egy ilyen pont a $0 \leq \varphi \leq \pi$, $0 \leq y \leq d$ téglalap egyenlő területű részeibe egyenlő valószínűséggel esik, akkor a problémát egyértelművé tettük. A tű valamelyik egyenest akkor metszi, ha

$$y \leq \frac{r}{2} \sin \varphi,$$

vagy

$$y \geq d - \frac{r}{2} \sin \varphi.$$

A keresett valószínűség ezek alapján $\frac{2r}{d\pi}$, amit a görbék által körülzárt terület meghatározásával, integrálással kapunk.

[6]Georges Louis Leclerc Buffon francia természettudós, 1707-1788

6.6. Feladatok

6.1. A $(0,1)$ intervallumban véletlenszerűen kiválasztunk egy pontot, majd egy másikat ettől jobbra. Mennyi a valószínűsége annak, hogy az így keletkezett három szakaszból háromszög szerkeszthető?

6.2. A $(0,1)$ intervallumban véletlenszerűen választunk egy számot. Mennyi a valószínűsége annak, hogy a szám harmadik tizedesjegye 5?

6.3. Egy villamosmegállóhoz a villamosok 5 percenként érkeznek: órakor, óra 5-kor, stb. Mondjuk du. 5 és fél 6 között egy véletlenszerűen kiválasztott időpontban a megállóhoz megyünk. Mennyi annak a valószínűsége, hogy a várakozási idő ne legyen hosszabb 2 percnél?

6.4. Egy üzembe valamelyik napon déli 12 óra és délután 3 óra között két gépkocsi nyersanyagszállítmány érkezik. A gépkocsik érkezéséről csak annyit tudunk, hogy mindkettő érkezésének időpontja egyenletes valószínűségeloszlású a $(0,3)$ intervallumon (az időtengely 0 pontját déli 12 órához helyezzük), és a gépkocsik vezetői nem tudnak egymás érkezéséről. A szállítmány lerakása bármelyik gépkocsi esetén 1 órát vesz igénybe, s a rakodás után a gépkocsik rögtön visszaindulnak. Mi a valószínűsége, hogy a két kocsi az üzem területén találkozik?

6.5. Egy egységnyi hosszúságú szakaszon találomra választunk két pontot. Mi a valószínűsége annak, hogy ezek közelebb vannak egymáshoz, mint bármelyik végponthoz?

6.6. A $(0,1)$ intervallumban találomra kiválasztunk egy x pontot. Mi a valószínűsége annak, hogy az x, $1-x$ és $\frac{1}{2}$ hosszúságú szakaszokból háromszöget lehet szerkeszteni?

6.7. Egy ember elfelejtette felhúzni az óráját, s az megállt. Mi a valószínűsége annak, hogy a nagymutató a 3 és a 6 között állt meg?

6.8. Egy botot találomra kettétörünk, majd a nagyobbik darabot újra találomra kettétörjük. Mennyi a valószínűsége annak, hogy az így keletkezett három darabból háromszög alkotható?

6.9. Dobjunk a $(0,t)$ szakaszra találomra két pontot. Mennyi annak a valószínűsége, hogy mindkét pont a $(0,x)$ részintervallumba esik?

6.10. Az előbbi feladatban mennyi annak a valószínűsége, hogy a két pont távolsága x-nél kisebb?

7. A valószínűség alapvető összefüggései

7.1. Elemi tulajdonságok

A következőkben a valószínűség alapvető tulajdonságait foglaljuk össze, melyek egyrészt elméleti szempontból fontosak, másrészt gyakorlatilag könnyítik meg a valószínűséggel kapcsolatos számításokat. A továbbiakban feltételezzük, hogy adott egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és tételeinket erre fogalmazzuk meg.

7.1.1. Tétel *A lehetetlen esemény valószínűsége* 0*.*

Bizonyítás. Bármely A esemény esetén $A \cdot O = O$ és $A + O = A$, tehát

$$P(A) = P(A + O) = P(A) + P(O),$$

amiből következik, hogy $P(O) = 0$.

7.1.2. Tétel *Bármely A esemény esetén*

$$P(\overline{A}) = 1 - P(A).$$

Bizonyítás. Mivel A és $\overline{A}$ egymást kizárják, és összegük a biztos esemény, ezért

$$1 = P(I) = P(A + \overline{A}) = P(A) + P(\overline{A}),$$

amiből az állítás következik.

7.1.3. Tétel *Ha az A esemény maga után vonja a B eseményt, akkor*
i) $P(A) \leq P(B)$,
ii) $P(B - A) = P(B) - P(A)$.

Bizonyítás. Az állítás abból következik, hogy a tétel feltételei mellett A és $B-A$ egymást kizáró események, és összegük B, így

$$P(B) = P(A) + P(B-A).$$

7.2. Az additivitás és a szubadditivitás

7.2.1. Tétel *Tetszőleges A, B eseményekre érvényes*

$$P(A+B) = P(A) + P(B) - P(AB).$$

Bizonyítás. Az állítás az $A+B = A+(B-AB)$ és $A(B-AB) = O$ egyenlőségekből a valószínűség additivitása és a 7.1.3. ii) alapján következik.

Az előbbi tétel ismételt alkalmazásával tetszőleges – véges – számú esemény összegének valószínűsége kiszámítható. A megfelelő, úgynevezett Poincaré[7] -formulát a következő tétel tartalmazza.

7.2.2. Tétel *Tetszőleges $A_1, A_2, \dots, A_n$ események esetén fennáll*

$$P(A_1 + A_2 + \cdots + A_n) = \sum_{k=1}^{n} (-1)^{k-1} S_k^{(n)},$$

ahol

$$S_k^{(n)} = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} P(A_{i_1} A_{i_2} \dots A_{i_k}).$$

Ez a tétel teljes indukcióval könnyen bizonyítható, de a következő, Jordán Károlytól[8] származó tétel speciális eseteként is megkaphatjuk:

[7] Henri Poincaré francia matematikus, 1854-1912
[8] Jordán Károly magyar matematikus, 1871-1959

7.2.3. Tétel *Tetszőleges $A_1, A_2, \ldots, A_n$ események esetén jelölje Q_k annak a valószínűségét, hogy ezek közül legalább k számú bekövetkezik. Ekkor*

$$Q_k = \sum_{i=k}^{n} (-1)^{i-k} \binom{i-1}{k-1} S_i^{(n)}.$$

A következő tétel ugyancsak Jordán Károlytól származik.

7.2.4. Tétel *Tetszőleges $A_1, A_2, \ldots, A_n$ események esetén jelölje P_k annak a valószínűségét, hogy ezek közül pontosan k számú bekövetkezik. Ekkor*

$$P_k = \sum_{i=k}^{n} (-1)^{i-k} \binom{i}{k} S_i^{(n)}.$$

Ezek a formulák egymásból nyerhetők a

$$Q_k = P_k + \cdots + P_n;$$

$$P_k = Q_k - Q_{k-1}$$

összefüggések alapján. Az utóbbi két tételt nem bizonyítjuk.

7.2.5. Tétel *Tetszőleges $A_1, A_2, \ldots, A_n$ események esetén fennáll*

$$P(A_1 + A_2 + \cdots + A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n).$$

Bizonyítás. A bizonyítás teljes indukcióval végezhető el. Az $n = 2$ eset a 7.2.1. tétel következménye. Ha feltesszük, hogy az állítás n-re igaz, akkor

$$P(A_1 + A_2 + \cdots + A_n + A_{n+1}) \leq P(A_1 + A_2 + \cdots + A_n) + P(A_{n+1}) \leq$$

$$\leq P(A_1) + P(A_2) + \cdots + P(A_n) + P(A_{n+1}).$$

A tétel a valószínűség úgynevezett szubadditív tulajdonságát fejezi ki.

7.3. A valószínűség folytonossága

7.3.1. Tétel *Legyenek $A_1, A_2, \dots$ olyan események, melyekre teljesül $A_1 \supset A_2 \supset \dots$. Ekkor*

$$P(\prod_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(A_n).$$

Bizonyítás. Legyen $A = \prod_{i=1}^{\infty} A_i$ és $B_k = A_k - A_{k+1}$. Ekkor az A, $B_1, B_2, \dots$ események páronként kizárják egymást, és összegük A_1, így a valószínűség teljes additivitása alapján

$$P(A_1) = P(A) + \sum_{k=1}^{\infty} P(B_k) = P(A) + \lim_{n\to\infty} \sum_{k=1}^{n-1} P(B_k).$$

Másrészt, a fentiek alapján $P(B_k) = P(A_k) - P(A_{k+1})$, ezért

$$P(A_1) = P(A) + \lim_{n\to\infty} (P(A_1) - P(A_n)),$$

amiből az állítás következik.

A fenti tételben foglalt állítást a valószínűség folytonossági tulajdonságának szokás nevezni. Egyszerű következményként kapjuk a következő analóg állítást:

7.3.2. Tétel *Legyenek $A_1, A_2, \dots$ olyan események, melyekre teljesül $A_1 \subset A_2 \subset \dots$. Ekkor*

$$P(\sum_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(A_n).$$

A következő tétel a valószínűség szubadditivitási tulajdonságát általánosítja végtelen sok eseményre.

7.3.3. Tétel *Ha $A_1, A_2, \dots$ tetszőleges események sorozata, akkor*

$$P(\sum_{k=1}^{\infty} A_k) \le \sum_{k=1}^{\infty} P(A_k).$$

Bizonyítás. Legyen $B_n = \sum_{k=1}^n P(A_k)$; ekkor nyilván $B_n \subset B_{n+1}$ és

$$\sum_{k=1}^{\infty} P(B_k) = \sum_{k=1}^{\infty} P(A_k),$$

tehát a valószínűség folytonossági tulajdonsága alapján

$$\lim_{n\to\infty} \sum_{k=1}^{n} P(A_k) = \lim_{n\to\infty} P(B_n) = \sum_{k=1}^{\infty} P(A_k)$$

teljesül. Másrészt

$$P(B_n) = P(\sum_{k=1}^{n} A_k) \leq \sum_{k=1}^{n} P(A_k) \leq \sum_{k=1}^{\infty} P(A_k),$$

amiből az állítás következik.

7.4. Feladatok

7.1. Bizonyítsuk be a 7.2.2. tételt!

7.2. Bizonyítsuk be, hogy ha A, B tetszőleges események, akkor

$$P(A \circ B) = P(A) + P(B) - 2P(AB).$$

7.3. Bizonyítsuk be, hogy ha A, B, C tetszőleges események, akkor

$$P(A \circ C) \leq P(A \circ B) + P(B \circ C).$$

7.4. Bizonyítsuk be, hogy ha A, B tetszőleges események, és

$$\Delta(A, B) = \begin{cases} \frac{P(A \circ B)}{P(A+B)} & \text{ha } P(A+B) > 0 \\ 0 & \text{ha } P(A+B) = 0, \end{cases}$$

akkor érvényes:

$$\Delta(A, C) \leq \Delta(A, B) + \Delta(B, C).$$

7.5. Bizonyítsuk be: ahhoz, hogy adott p_1, p_2, p_{12} valós számokhoz található legyen olyan A és B esemény, hogy $P(A) = p_1$, $P(B) = p_2$, $P(AB) = p_{12}$, a következő négy egyenlőtlenség fennállása szükséges és elégséges:

$$p_1 + p_2 \leq 1 + p_{12},$$

$$p_{12} \leq p_1,$$

$$p_{12} \leq p_2,$$

$$p_{12} \geq 0.$$

7.6. Bizonyítsuk be a 7.3.2. tételt!

8. Feltételes valószínűség és függetlenség

8.1. A feltételes valószínűség értelmezése

A gyakorlati életben gyakran vetődik fel az a probléma, hogy valamely esemény bekövetkezése befolyásolja-e, s ha igen, milyen mértékben egy másik esemény bekövetkezését. Mivel a valószínűségelmélet minden kísérletet megadott körülmények között vizsgál, az események véletlen jellegét e körülménykomplexum határozza meg. Ha tehát azt kérdezzük, hogy mekkora az A esemény valószínűsége egy olyan kísérletben, amelyben a B esemény bekövetkezett, akkor ez a valószínűség általában különbözni fog az A esemény valószínűségétől, hiszen ekkor egy A-tól különböző eseményről van szó. Ezt az eseményt szokás $A|B$-vel jelölni, valószínűségét pedig $P(A|B)$-vel, amit az A eseménynek a B eseményre vonatkozó *feltételes valószínűségének* nevezünk. Ha mindezt a gyakoriságok, illetve relatív gyakoriságok nyelvén kívánjuk megfogalmazni, akkor a következőképpen gondolkodhatunk. Tekintsük azt a kísérletet, amely az A és B eseményekkel kapcsolatos, és ezt végezzük el n-szer egymástól függetlenül. Válasszuk ki azokat az eseteket, amelyekben a B esemény bekövetkezett. (Természetesen feltételezzük, hogy a B esemény valószínűsége zérustól különböző.) Ezek száma k_B: a B esemény gyakorisága. Ezen esetek között vannak olyanok, amelyekben az A esemény is bekövetkezett, melyek száma k_{AB}, az AB esemény gyakorisága. Mint általában a gyakoriságok hányadosai, a $\frac{k_{AB}}{k_B}$ hányados is stabilitási tendenciát mutat, ha a kísérletek számát növeljük. A fenti hányadost az A eseménynek a B eseményre vonatkozó *feltételes relatív gyakoriságának* nevezzük. Ha figyelembe vesszük, hogy a gyakoriságoknak és a kísérletek számának hányadosa bizonyos értelemben a valószínűséghez közeledik a kísérletek számának növelésével, akkor kézenfekvő, hogy a $\frac{k_{AB}}{k_B}$ hányados a $\frac{P(AB)}{P(B)}$ érték körül fog ingadozni, melyet az A eseménynek a B eseményre vonatkozó *feltételes valószínűségének* nevezünk, és $P(A|B)$-vel jelölünk.

8.2. A feltételes valószínűség tulajdonságai

A feltételes valószínűség rögzített B feltétel mellett egy valószínűségi függvény, amely tehát rendelkezik a valószínűség karakterisztikus tulajdonságaival. Ezt fejezi ki a következő tétel.

8.2.1. Tétel *Legyen $(\Omega, \mathcal{A}, P)$ valószínűségi mező, A és B két tetszőleges esemény, továbbá $P(B) > 0$, akkor az A esemény B-re vonatkozó $P(A|B)$ feltételes valószínűsége rendelkezik a következő tulajdonságokkal:*

i) $0 \leq P(A|B) \leq 1$;

ii) $P(B|B) = 1$;

iii) $P(\sum_{i=1}^{\infty} A_i|B) = \sum_{i=1}^{\infty} P(A_i|B)$, *ha $A_1, A_2, \ldots$ egymást páronként kizáró események sorozata.*

A tétel bizonyítása egyszerű számolás, a definíció és a valószínűség tulajdonságai alapján.

A feltételes valószínűséget definiáló formula a következőképpen is írható:

$$P(AB) = P(A|B)P(B).$$

Ezt a formulát szokás *szorzási szabálynak* nevezni. A szorzási szabályt gyakran két esemény szorzata valószínűségének kiszámítására lehet felhasználni.

Tekintsünk egy egyszerű példát. Egy urnában van 3 fehér és 6 fekete golyó. Egymás után kettőt kihúzunk. Mi a valószínűsége annak, hogy a második golyó fehér, ha az első fekete volt? Természetesen feltételezzük, hogy mindkét húzás úgy megy végbe, hogy az urnában levő golyók egyenlő valószínűséggel húzhatók ki. Jelölje B azt az eseményt, hogy az első kihúzott golyó fekete, A pedig azt, hogy a második fehér. A feladat kérdése tehát a $P(A|B)$ feltételes valószínűségre vonatkozik. Meg kell tehát határoznunk a $P(AB)$ és a $P(B)$ valószínűségeket. A húzások sorrendje is lényeges, ezért az összes lehetőségek száma $9 \cdot 8 = 72$. Ezek közül azok száma, amikor az első golyó fekete, a második pedig fehér: $6 \cdot 3 = 18$. Ezért

$$P(AB) = \frac{18}{72} = \frac{1}{4}.$$

A B esemény valószínűsége nyilván $\frac{6}{9} = \frac{2}{3}$, így a keresett feltételes valószínűség:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{1}{4} \cdot \frac{3}{2} = \frac{3}{8}.$$

A szorzási szabály alkalmazására tekintsük a következő példát: 50 termékből, amelyek között 10 selejtes, egymás után találomra kiveszünk kettőt visszatevés nélkül. Mennyi a valószínűsége annak, hogy az elsőre selejtest, a másodikra pedig jót húzunk? Jelentse A azt az eseményt, hogy elsőre selejtest húzunk, B pedig azt, hogy másodikra jót húzunk. Nyilván

$$P(A) = \frac{10}{50} = \frac{1}{5}.$$

A második húzásnál 49-ből húzunk és ezek közül 40 a kedvező, így

$$P(B|A) = \frac{40}{49}.$$

Ezért, a szorzási szabály alapján

$$P(AB) = P(B|A)P(A) = \frac{40}{49} \cdot \frac{1}{5} = \frac{8}{49}.$$

A szorzási szabály könnyen általánosítható n számú eseményre a következő módon:

8.2.2. Tétel *Legyenek $A_1, A_2, \ldots, A_n$ tetszőleges események, ekkor*

$$P(A_1 A_2 \cdot \ldots \cdot A_n) =$$

$$= P(A_n|A_1 A_2 \cdot \ldots \cdot A_{n-1}) P(A_{n-1}|A_1 A_2 \cdot \ldots \cdot A_{n-2}) \ldots P(A_2|A_1) P(A_1).$$

A bizonyítás a szorzási szabály ismételt alkalmazásával végezhető el.

A 8.2.2. tételben foglalt általános szorzási szabály alkalmazására tekintsük a következő példát: a 32 lapos magyar kártyából négy lapot húzunk ki egymás után, visszatevés nélkül. Mi a valószínűsége, hogy mind a négy lap piros lesz? Jelölje A_i azt az eseményt, hogy az i-edik húzás eredménye piros lap ($i = 1, 2, 3, 4$). Ekkor

$$P(A_1) = \frac{8}{32}, \quad P(A_2|A_1) = \frac{7}{31}, \quad P(A_3|A_1 A_2) = \frac{6}{30}, \quad P(A_4|A_1 A_2 A_3) = \frac{5}{29},$$

így a válasz a kérdésre a fentiek alapján:

$$P(A_1 A_2 A_3 A_4) = \frac{8}{32} \cdot \frac{7}{31} \cdot \frac{6}{30} \cdot \frac{5}{29}.$$

8.3. A teljes valószínűség tétele

A következő tétel a feltételes valószínűségekre vonatkozó egyik legfontosabb eredmény: a teljes valószínűség tétele.

8.3.1. Tétel *Legyen $B_1, B_2, \ldots$ pozitív valószínűségű eseményekből álló teljes eseményrendszer, A pedig tetszőleges esemény. Ekkor*

$$P(A) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i).$$

BIZONYÍTÁS. Mivel $B_1, B_2, \ldots$ teljes eseményrendszer, ezért

$$A = AI = A(\sum_{i}^{\infty} B_i) = \sum_{i}^{\infty} AB_i,$$

és a valószínűség teljes additivitása alapján

$$P(A) = P(\sum_{i}^{\infty} AB_i) = \sum_{i}^{\infty} P(AB_i).$$

Alkalmazva a szorzási szabályt az egyes tagokra, kapjuk, hogy

$$P(A) = \sum_{i}^{\infty} P(AB_i) = \sum_{i}^{\infty} P(A|B_i)P(B_i),$$

amit bizonyítani kellett.

Megjegyezzük, a tétel azt fejezi ki, hogy bármely esemény valószínűsége egyenlő az eseménynek egy teljes eseményrendszer eseményeire vonatkozó feltételes valószínűségeinek a teljes eseményrendszer elemei valószínűségeivel súlyozott számtani közepével.

Most tekintsük ennek a tételnek egy egyszerű alkalmazását. Tegyük fel, hogy három urnában fehér és fekete golyókat helyeztünk el, mégpedig az elsőbe 2 fehér és 3 fekete, a másodikba 4 fehér és 1 fekete, a harmadikba pedig 3 fehér és 7 fekete golyót. A kísérlet során $\frac{1}{2}$ valószínűséggel az első, $\frac{1}{3}$ valószínűséggel a második és $\frac{1}{6}$ valószínűséggel a harmadik urnát választva találomra kiveszünk abból egy golyót. Mennyi annak a valószínűsége, hogy ez a golyó fehér lesz? Jelölje A ezt az eseményt, B_1, B_2, B_3 pedig rendre azt, hogy az első, a második, illetve a harmadik urnát választottuk. Ekkor

$$P(B_1) = \frac{1}{2}, \;\; P(B_2) = \frac{1}{3}, \;\; P(B_3) = \frac{1}{6},$$

továbbá

$$P(A|B_1) = \frac{2}{5}, \;\; P(A|B_2) = \frac{4}{5}, \;\; P(A|B_3) = \frac{3}{10},$$

tehát, a teljes valószínűség tétele alapján

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) = \frac{31}{60}.$$

8.4. A Bayes[9]-tétel

Gyakori eset, hogy egy esemény bekövetkezésének ismerete mellett kívánunk információt kapni arra vonatkozóan, hogy egy teljes eseményrendszer eseményei mekkora valószínűséggel játszottak közre ebben. Erre vonatkozik a következő tétel: Bayes tétele.

8.4.1. Tétel *Legyen $B_1, B_2, \ldots$ pozitív valószínűségű eseményekből álló teljes eseményrendszer, A pedig tetszőleges pozitív valószínűségű esemény. Ekkor $k = 1, 2, \ldots$ esetén*

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^{\infty} P(A|B_i)P(B_i)}.$$

Bizonyítás. A feltételes valószínűség szorzási tétele alapján nyilván

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{P(A)}.$$

A teljes valószínűség tétele alapján a nevezőben álló $P(A)$ behelyettesíthető:

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^{\infty} P(A|B_i)P(B_i)},$$

[9]Thomas Bayes angol pap, 1702-1761

ami éppen az állítás.

A tétel szokásos elnevezése: az "okok valószínűségének tétele ", hiszen egy esemény bekövetkezéséből következtetünk a hipotézisek fennállásának valószínűségére. Egy lehetséges alkalmazást illusztrál a következő példa, mely a fentiekben tárgyalt példa megfordítása. Az ottani feltételek mellett kérdezzük azt: mennyi a valószínűsége annak, hogy a húzás az első urnából történt, ha a húzás eredménye fehér golyó? A Bayes-tétel alapján a korábban kapott értékeket behelyettesítve kapjuk, hogy

$$P(B_1|A) = \frac{12}{31},\ \ P(B_2|A) = \frac{16}{31},\ \ P(B_3|A) = \frac{3}{31},$$

tehát a kérdezett valószínűség $\frac{12}{31}$.

8.5. Függetlenség

A feltételes valószínűséggel kapcsolatos legfontosabb valószínűségelméleti fogalom a *függetlenség*. Hétköznapi értelemben két esemény függetlensége olyasmit jelent, hogy semmilyen befolyással nincsenek egymásra, tehát bármelyiknek a bekövetkezése esetén semmiféle információt nem kapunk a másik esemény bekövetkezésének valószínűségére nézve. Matematikailag ez a következőképpen fogalmazható meg: ha A, B két esemény, $P(B) > 0$, akkor az A-nak B-től való függetlensége azt jelenti, hogy az a tény, hogy a B esemény bekövetkezett, nem befolyásolja az A esemény bekövetkezésének valószínűségét, tehát

$$P(A|B) = P(A).$$

Ekkor azt mondjuk, hogy *A független B-től.* Ha ilyenkor A is pozitív valószínűségű, akkor

$$P(B|A) = \frac{P(BA)}{P(A)} = \frac{P(AB)}{P(A|B)} = P(B),$$

tehát B is független A-tól, így azt mondhatjuk, hogy A és B *függetlenek.* A szorzási szabály alapján következik az alábbi állítás.

8.5.1. Tétel *Az A és B események pontosan akkor függetlenek, ha*

$$P(AB) = P(A)P(B).$$

Megjegyezzük, hogy az előbbi formula alkalmazásához nincs szükség annak feltételezésére, hogy A vagy B pozitív valószínűségű legyen, így inkább ezt a formulát szokás tekinteni a függetlenség definíciójának, s a továbbiakban mi is ezt tesszük. Másrészt, ebből nyilvánvaló, hogy egy 0 valószínűségű esemény minden más eseménytől független.

Tekintsünk egy példát független eseményekre. Egy ládában 25 darab 110 voltos és 40 wattos, 25 darab 110 voltos és 60 wattos, 25 darab 220 voltos és 40 wattos és 25 darab 220 voltos és 60 wattos izzó van. Véletlenszerűen választunk egyet közülük. Jelölje A azt az eseményt, hogy a kiválasztott lámpa 110 voltos, B pedig azt, hogy 40 wattos. Ekkor $P(AB) = \frac{25}{100}$, $P(A) = \frac{50}{100}$, $P(B) = \frac{50}{100}$, így $P(AB) = P(A)P(B)$, tehát az A és B események függetlenek.

A függetlenség fogalma több eseményre is kiterjeszthető, a következő módon: események egy halmazát *függetlennek* nevezzük, ha közülük bármely véges sokat kiválasztva, azok szorzatának valószínűsége egyenlő valószínűségeik szorzatával. Ebből nyilván következik, hogy az események közül bármely kettő független egymástól, mással az események *páronként függetlenek*, ez a fogalom azonban ennél többet jelent, amint azt a következő példa mutatja: dobjunk fel két kockát, és jelölje A azt az eseményt, hogy az első kockával páros számot dobtunk, B azt, hogy a másodikkal páratlan számot dobtunk, C pedig azt, hogy vagy mindkét kockával páros, vagy mindkét kockával páratlan számot dobtunk. Ekkor nyilván

$$P(A) = P(B) = P(C) = \frac{1}{2},$$

valamint

$$P(AB) = P(AC) = P(BC) = \frac{1}{4}.$$

Ez azt jelenti, hogy az A, B, C események páronként függetlenek. Ugyanakkor $P(ABC) = 0 \neq P(A)P(B)P(C)$, ami azt jelenti, hogy az A, B, C események a fenti definíció értelmében nem függetlenek. Kettőnél több esemény esetében a függetlenség fenti fogalmát a páronkénti függetlenségtől való megkülönböztetés céljából *teljes függetlenségnek* nevezzük.

Könnyen igazolható a következő tétel.

8.5.2. Tétel *Legyenek az $A_1, A_2, \ldots, A_n$ események teljesen függetlenek. Ekkor azok az események is teljesen függetlenek, amelyeket úgy kapunk, hogy az $A_1, A_2, \ldots, A_n$ események közül tetszőleges számút kicserélünk az ellentettjére.*

A függetlenség fennállása esetén számos korábbi eredmény lényegesen egyszerűsödik. A következő két tétel könnyen adódik a 7.2.2. és 7.2.4. tételekből.

8.5.3. Tétel *Tetszőleges teljesen független $A_1, A_2, \ldots, A_n$ események esetén fennáll*

$$P(A_1 + A_2 + \cdots + A_n) = 1 - P(\overline{A_1})P(\overline{A_2}) \ldots P(\overline{A_n}).$$

8.5.4. Tétel *Legyenek $A_1, A_2, \ldots, A_n$ teljesen függetlenek, és legyen $P(A_i) = p$ $(i = 1, 2, \ldots, n)$. Jelölje P_k annak a valószínűségét, hogy ezek közül pontosan k számú következik be. Ekkor*

$$P_k = \binom{n}{k} p^k (1-p)^{n-k}.$$

Példaként tekintsük az úgynevezett Bernoulli[10]-féle problémát. Tekintsünk egy kísérletet, melynek két lehetséges kimenetele A és $\overline{A}$, továbbá legyen $P(A) = p$. Végezzük el a kísérletet n-szer egymás után, egymástól függetlenül. Mennyi annak a valószínűsége, hogy ebben a kísérletsorozatban az A esemény pontosan k-szor következik be? Nyilván feltételezzük, hogy $0 \le k \le n$. Jelölje A_i azt az eseményt, hogy az i-edik kísérletben az A esemény bekövetkezik. Ekkor nyilván $P(A_i) = p$ $(i = 1, 2, \ldots, n)$, és a kísérletek független elvégzése azt jelenti, hogy az A_i események teljesen függetlenek. Ezért a fenti tétel alapján a keresett valószínűség

$$P_k = \binom{n}{k} p^k (1-p)^{n-k}.$$

[10] Jacob Bernoulli svájci matematikus, 1654-1705

8.6. Feladatok

8.1. Bizonyítsuk be, hogy

$$P(A_1 + A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1A_2|B).$$

8.2. Mutassuk meg, hogy ha $P(B) > 0$, akkor érvényesek a következő összefüggések:
i) $0 \leq P(A|B) \leq 1$;
ii) $P(B|B) = 1$;
iii) $P(A_1 + A_2|B) = P(A_1|B) + P(A_2|B)$, ha $A_1A_2 = \emptyset$;
iv) $P(\overline{A}|B) = 1 - P(A|B)$.

8.3. Bizonyítsuk be, hogy ha B, C pozitív valószínűségű események, és $P(B|C) > 0$, akkor bármely A esemény esetén

$$P(A|BC) = \frac{P(AB|C)}{P(B|C)}.$$

8.4. Bizonyítsuk be, hogy ha B pozitív valószínűségű esemény, akkor bármely A esemény esetén

$$P(A|B) = P(AB|B).$$

8.5. Bizonyítsuk be, hogy ha A és B kizárják egymást, továbbá B pozitív valószínűségű esemény, akkor

$$P(A|B) = 0.$$

8.6. Egy játékkockát egyszer feldobunk. Legyen A az az esemény, hogy 6-nál kisebb számot dobunk, B pedig az, hogy páros számot dobunk. Számítsuk ki a $P(B)$ és $P(B|A)$ valószínűségeket!

8.7. Három kockával dobunk. Mennyi a valószínűsége annak, hogy az egyik kockával hatost dobunk, feltéve, hogy a dobott számok összege 12?

8.8. Egy önkiszolgáló üzletben 200 üveg tej friss, és 100 üveg másnapos. Két üveget kiveszünk. Mennyi a valószínűsége, hogy mindkettő friss?

8.9. 200 munkadarabból, melyek között 20 a selejtes, egymás után kiveszünk kettőt, visszatevés nélkül. Mennyi a valószínűsége annak, hogy
i) az elsőre selejtest, a másodikra jót húzunk;
ii) az elsőre jót, a másodikra selejtest húzunk;

iii) egyik alkalommal selejtest, másik alkalommal jó darabot húzunk?

8.10. Egy csomag magyar kártyából visszatevés nélkül kihúzunk 4 lapot. Mennyi annak a valószínűsége, hogy az első kettő piros, a másik kettő zöld?

8.11. Egy női konfekció boltban háromféle méretű ruha kapható: nagy-, közepes és kisméretű. Ezek megoszlása: nagyméretű az összes ruhák 26%-a, közepes 44%, kicsi a 30%-a. A ruhákat kívánságra alakra igazítják. Statisztikai adatokból ismeretes, hogy az alakra igazítás valószínűsége 0.2, 0.1 illetve 0.15. Mennyi a valószínűsége, hogy egy eladott ruhát alakítani kell?

8.12. Televízió-képcsövek kísérleti gyártását végzik egy gyárban. Három tétel készül el. Az első két tétel a teljes mennyiség egy-egy negyedét, a harmadik a felét adja. A vizsgálat során kiderül, hogy az előírt működési óraszámot az első tételnek csak a 90%-a, a másodiknak a 80%-a, a harmadiknak a 75%-a éri el. Mennyi a valószínűsége annak, hogy egy, a teljes mennyiségből találomra kiszemelt képcső az előírt ideig működik?

8.13. Négy termelőtől almát szállítanak egy üzletbe. Az első termelőtől származik a mennyiség $\frac{1}{10}$ része, melyből 40% elsőosztályú. A második termelőtől szállítják a tétel $\frac{1}{4}$ részét, amely 50%-ban elsőosztályú. A harmadiktól rendelték a mennyiség $\frac{2}{5}$ részét, ebből 20% elsőosztályú áru. A többi gyümölcs a negyedik termelőtől kerül az üzletbe, és mind elsőosztályú. Mennyi a valószínűsége annak, hogy az üzletben e szállítmányból találomra kiválasztva egy almát, az elsőosztályú?

8.14. Izzólámpákat 100 db-os csomagolásban szállítanak. Az előző megfigyelésekből ismert, hogy a 100 db-os tételek között azonos arányban fordul elő 0, 1, 2, 3 hibás darabot tartalmazó (3-nál több hibás darab nem fordul elő). Mennyi a valószínűsége annak, hogy egy tételből véletlenszerűen 3 égőt kiválasztva, mindhárom hibátlan lesz?

8.15. Egy rádiókészülékeket gyártó üzemben számos vizsgálat alapján megállapították, hogy minden 10 darab elkészült rádió között egyforma valószínűséggel lehet 0, 1, 2 vagy 3 selejtes, több nem. Egy ilyen 10 darab készüléket tartalmazó tételből a minőségellenőrzés során kiválasztottak 2 rádiót, azokat megvizsgálták, és mind a kettőt jónak találták. Mennyi a valószínűsége, hogy mind a 10 készülék hibátlan?

8.16. A görögdinnye-szállítmányokat általában úgy ellenőrzik, hogy néhány dinnyét meglékelnek, s a kapott eredményből következtetnek a szállítmányra. Menynyi annak a valószínűsége, hogy 100 görögdinnye közül legfeljebb 5 "rossz", ha közülük találomra kiválasztott 10 dinnye mindegyike "jó "-nak bizonyult? Az előzetes tapasztalatok alapján a 100 tételes szállítmányokban legfeljebb 20 darab lehet "rossz". A "rossz" dinnyék számának lehetséges értékei $(0, 1, 2, \ldots, 20)$ mind egyformán valószínűek.

8.17. Egy nagykereskedelmi raktárban az egyik konzervfajtából 500 darabból álló készlet van. Ebből 10 elemű mintát választottak. A minta átvizsgálásakor azt találták, hogy abban 2 darab nem felelt meg a minőségi előírásoknak. Mennyi a valószínűsége annak, hogy a selejtes konzervek száma az eredeti készletben is 2?

8.18. Ha az előző feladatban annak a valószínűsége érdekel bennünket, hogy a készletben a selejtes konzervek száma legfeljebb 3, hogyan tudnánk ezt meghatározni?

8.19. Egy tesztrendszerű vizsgánál minden vizsgakérdés egy vizsgalapra van felírva, és minden kérdéshez négy válasz van megadva, amelyek közül csak egy a helyes. A vizsgázónak ezt a lapot kell kitölteni a helyesnek vélt válasz megjelölésével. Tegyük fel, hogy egy vizsgázó p valószínűséggel tudja a helyes választ. Ha nem tudja a választ, akkor $\frac{1}{4}$ valószínűséggel jelöli meg a 4 lehetséges válasz közül az egyiket.

i) Mennyi a valószínűsége, hogy a vizsgázó helyesen válaszol?

ii) A vizsgalap átnézése során kiderül, hogy helyes a válasz. Mennyi a valószínűsége, hogy azért adott helyes választ, mert tudta a helyes eredményt?

8.20. Ketten asztaliteniszt játszanak. Mindkét játékosra nézve $\frac{1}{2}$ annak a valószínűsége, hogy egy pontot szerezzen, és az egyes pontok nyerései egymástól függetlenek. Mennyi annak a valószínűsége, hogy a játék valamelyik fél 23 : 21 arányú győzelmével ér véget?

8.21. Bizonyítsuk be a 8.5.3. tételt!

8.22. Bizonyítsuk be a 8.5.4. tételt!

9. Valószínűségi változók

9.1. A valószínűségi változó értelmezése

A gyakorlatban előforduló véletlen kísérletek elvégzése során a kísérletek eredményével, a bekövetkező elemi eseménnyel kapcsolatban többnyire valamilyen számérték is adódik. Tételezzük fel, hogy az elemi esemény ezt a számértéket egyértelműen meghatározza. Sok esetben magát az elemi eseményt egy számértékkel tudjuk jellemezni. Például, a kockadobási kísérletnél azok az elemi események, hogy a dobókocka egyes lapjai felülre kerülnek, rendre az $1, 2, 3, 4, 5, 6$ számértékekkel jellemezhetők. A pénzérme feldobásával kapcsolatos kísérletnél is tetszőleges módon hozzárendelhetünk egy – egy számértéket a fej, illetve írás dobáshoz. Mivel a kísérlet ismételt elvégzésekor más és más elemi események következhetnek be, így a figyelemmel kísért számérték is más és más lehet: ez az érték a véletlentől függ. Természetesen egy és ugyanazon kísérlethez tartozó elemi eseményekhez számos különböző módon rendelhetünk számértékeket. Az elemi események halmazán értelmezett függvényeket valószínűségi változóknak nevezzük. A változókat többnyire a görög ábécé kisbetűivel jelöljük. Valószínűségi változókkal kapcsolatban leggyakrabban a következő típusú kérdésekre kell felelni. Mi a valószínűsége annak, hogy a ξ valószínűségi változó értéke kisebb (nagyobb, kisebb vagy egyenlő, nagyobb vagy egyenlő), mint az adott x valós szám? Ahhoz, hogy erre a kérdésre válaszolni lehessen, fel kell tételeznünk, hogy azon elemi események összessége, melyekre a fenti feltétel teljesül, valószínűségelméleti értelemben esemény. Legyen tehát adott az $(\Omega, \mathcal{A}, P)$ valószínűségi mező, ahol Ω egy tetszőleges, nem üres halmaz, $\mathcal{A}$ az Ω részlamazainak egy σ-algebrája, P pedig egy valószínűség. A $\xi : \Omega \to \mathbb{R}$ valós függvényt *valószínűségi változónak* nevezzük, ha bármely valós x esetén az $\{\omega : \xi(\omega) < x\}$ halmaz $\mathcal{A}$-ba tartozik. Röviden, az eseménytéren értelmezett valós értékű függvények közül azokat nevezzük valószínűségi változóknak, melyek rendelkeznek azzal a tulajdonsággal, hogy bármely rögzített valós x esetén mindazon elemi események összessége, amelyeken ez a függvény az adott x-nél kisebb értéket vesz fel, esemény. Tekintsünk egy egyszerű példát, amely egy szabályos pénzérme egyszeri feldobásával kapcsolatos kísérletet írja le. Az eseménytér legyen $\Omega = \{0, 1\}$, ahol 0 reprezentálja a fej, 1 pedig az írás dobását. Az események az Ω összes részhalmazai, a P valószínűségeloszlás pedig a következőképpen van megadva: ha A tetszőleges esemény, akkor legyen

$$P(A) = \frac{1}{2},$$

ha $A = \{0\}$ vagy $A = \{1\}$; továbbá $P(O) = 0$ és $P(I) = 1$. Mivel Ω minden részhalmaza esemény, ezért minden valós értékű függvény valószínűségi változó. Legyen $\xi(\omega) = 0$, ha $\omega = 0$, és $\xi(\omega) = 1$, ha $\omega = 1$. Ezzel egy valószínűségi változót értelmeztünk, melynek értéke az ω elemi esemény bekövetkezésekor éppen az ω elemi eseményt reprezentáló számértékkel egyenlő. Legyen x tetszőleges valós szám, és vizsgáljuk meg az $A_x = \{\omega : \xi(\omega) < x\}$ eseményt. Világos, hogy $A_x = O$, ha $x \leq 0$ és $A_x = I$, ha $x > 1$. Végül, ha $0 < x \leq 1$, akkor $A_x = \{0\}$.

A továbbiakban egy adott ξ valószínűségi változó és adott x valós szám esetén az $\{\omega : \xi(\omega) < x\}$ halmazt röviden így jelöljük: $\{\xi < x\}$, valószínűségét pedig $P(\xi < x)$-szel. Hasonló rövidítést alkalmazunk akkor, ha $<$ helyett $=$, $\leq$, $>$ vagy $\geq$ áll.

Előfordulhat, hogy egy adott kísérlet során bekövetkező elemi eseményekhez nem egy, hanem több, de az elemi eseményektől független számú számérték rendelhető. Ezeket a számértékeket egy vektornak tekintve az elemi események halmazán egy vektor-értékű függvényt értelmezhetünk, melyet *valószínűségi vektorváltozónak* nevezünk, ha komponensei közönséges valószínűségi változók.

9.2. Eloszlásfüggvény

Legyen adott az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn a ξ valószínűségi változó, és legyen E a valós számok egy olyan részhalmaza, melyre $\{\omega : \xi(\omega) \in E\}$ esemény. (Ezt az eseményt a továbbiakban $\{\xi \in E\}$ jelöli, valószínűségét pedig $P(\xi \in E)$.) Például, ha E egy $(-\infty, x)$ alakú nyílt intervallum, akkor ez a valószínűségi változó definíciója alapján nyilván teljesül. Nem nehéz megmutatni, hogy ugyanez a helyzet akkor, ha E a számegyenes bármilyen (nyílt, zárt, félig nyílt, véges vagy végtelen) intervalluma. Azt is meg lehet mutatni, hogy az összes ilyen tulajdonságú E halmazok egy $\mathcal{B}$ σ-algebrát alkotnak a valós számok halmazában. Ezt a σ-algebrát – mely nyilván független a szóbanforgó valószínűségi változótól – *Borel*[11]*-féle halmaztestnek* szokás nevezni. Legyen bármely $E \in \mathcal{B}$ esetén

$$P_\xi(E) = P(\xi \in E).$$

[11]Émile Félix Eduard Borel francia matematikus, 1871-1956

Könnyű belátni, hogy az így értelmezett $P_\xi : \mathcal{B} \to \mathbb{R}$ függvény teljesíti a valószínűség axiómáit, azaz $(\mathbb{R}, \mathcal{B}, P_\xi)$ valószínűségi mező. A P_ξ valószínűségeloszlást a ξ valószínűségi változó *eloszlásának* nevezzük. Egy valószínűségi változó eloszlása tehát mindig egy, a valós számok $\mathcal{B}$ σ-algebráján értelmezett valószínűség. A fenti példában szereplő ξ valószínűségi változó eloszlását egyszerűen meghatározhatjuk, ugyanis nyilván $P_\xi(E)$ értéke 0, ha E nem tartalmazza sem a 0-t, sem az 1-et; $\frac{1}{2}$, ha a 0 és 1 közül pontosan az egyik eleme E-nek, és 1, ha a 0 és 1 mindegyike eleme E-nek.

A gyakorlatban a valószínűségeloszlás leírása eléggé körülményes lehet, hiszen ehhez meg kellene adnunk az összes $P_\xi(E)$ valószínűségeket, tetszőleges $E \in \mathcal{B}$ halmaz esetén. Be lehet azonban vezetni egy olyan egyszerűbb, és a gyakorlat számára rendkívül fontos fogalmat, amelyből ezek a valószínűségek mind származtathatók. Legyen tetszőleges valós x esetén

$$F_\xi(x) = P_\xi(\xi < x) = P(\{\omega : \xi(\omega) < x\}).$$

Ekkor $F_\xi : \mathbb{R} \to \mathbb{R}$ egy függvény, melyet a ξ valószínűségi változó *eloszlásfüggvényének* nevezünk. Például, a fenti példában szereplő ξ valószínűségi változó eloszlásfüggvénye a következő:

$$F_\xi(x) = \begin{cases} 0 & \text{ha } x \le 0, \\ \frac{1}{2} & \text{ha } 0 < x \le 1, \\ 1 & \text{ha } x > 1. \end{cases}$$

A továbbiakban – ha ez nem okoz félreértést – F_ξ helyett egyszerűen F-et írunk.

A következőkben összefoglaljuk az eloszlásfüggvény legfontosabb tulajdonságait.

1) Bármely valószínűségi változó eloszlásfüggvénye monoton növekvő függvény. Valóban, ha $x < y$ tetszőleges valós számok, akkor a $\{\xi < x\}$ esemény maga után vonja a $\{\xi < y\}$ eseményt, így $F(x) \le F(y)$.

2) Bármely valószínűségi változó eloszlásfüggvényének határértéke $-\infty$-ben 0, $+\infty$-ben 1. Bebizonyítjuk a második állítást; az első bizonyítása hasonlóan történik. Mivel F monoton növekvő, elegendő megmutatni, hogy ha $\{x_n\}$ monoton növekvő valós számsorozat, melyre $\lim_{n\to\infty} x_n = +\infty$, akkor $\lim_{n\to\infty} F(x_n) = 1$. Legyen $A_n = \{\xi < x_n\}$. Ekkor $A_1 \subset A_2 \subset A_3 \subset \ldots$ és $\sum_{n=1}^{\infty} A_n = I$. A valószínűség folytonossági tétele alapján

$$F(x_n) = P(\xi < x_n) = P(A_n) \to P(I) = 1$$

ha $n \to \infty$, ami az állítást bizonyítja.

3) Bármely valószínűségi változó eloszlásfüggvénye balról folytonos. A bizonyítás a fenti mintára végezhető el. Mivel F monoton növekvő, elegendő megmutatni, hogy ha $\{x_n\}$ egy monoton növekvő valós számsorozat, melyre $\lim_{n\to\infty} x_n = x$, akkor $\lim_{n\to\infty} F(x_n) = F(x)$. Legyen $A_n = \{\xi < x_n\}$ és $A = \{\xi < x\}$. Ekkor $A_1 \subset A_2 \subset A_3 \subset \ldots$ és $\sum_{n=1}^{\infty} A_n = A$. A valószínűség folytonossági tétele alapján

$$F(x_n) = P(\xi < x_n) = P(A_n) \to P(A) = F(x)$$

ha $n \to \infty$, ami az állítást bizonyítja.

Az eloszlásfüggvény fenti három tulajdonsága egyben jellemzi is az eloszlásfüggvényt, azaz bármely valós változós, valós értékű függvényhez – mely rendelkezik a fenti tulajdonságokkal – van olyan valószínűségi mező, és azon olyan valószínűségi változó, amelynek az adott függvény az eloszlásfüggvénye. Ennek a fontos állításnak a bizonyításával itt nem foglalkozunk.

Felvetődik a kérdés, hogy az F eloszlásfüggvény ismeretében hogyan lehet meghatározni egy tetszőleges $\mathcal{B}$-beli E valószínűségét? Könnyű belátni az alábbi összefüggéseket: bármely $a < b$ esetén

i) $P(a \le \xi < b) = F(b) - F(a)$;
ii) $P(\xi = a) = F(a+0) - F(a)$;
iii) $P(a < \xi < b) = F(b) - F(a+0)$;
iv) $P(a \le \xi \le b) = F(b+0) - F(a)$;
v) $P(a < \xi \le b) = F(b+0) - F(a+0)$.

Ilyen módon, ha E a számegyenes tetszőleges intervalluma, akkor valószínűsége F segítségével kiszámítható. Ha az E halmaz páronként diszjunkt intervallumok sorozatának egyesítése, akkor a valószínűség teljes additivitása alapján számítható ki a valószínűsége. Belátható, hogy ha E ezeknél bonyolultabb $\mathcal{B}$-beli halmaz, F akkor is meghatározza a valószínűségét, de ennek részleteivel nem foglalkozunk.

A fenti formulákból azonnal látható, hogy ha az eloszlásfüggvény folytonos, akkor bármely valós x esetén $P(\xi = x) = 0$, azaz ekkor a valószínűségi változó minden valós értéket 0 valószínűséggel vesz fel.

9.3. Diszkrét és folytonos eloszlások

A valószínűségi változóknak és eloszlásoknak a következő fontos típusait megkülönböztetjük meg.

1) Diszkrét valószínűségi változó. Egy valószínűségi változót *diszkrétnek*, vagy *diszkrét eloszlásúnak* nevezünk, eloszlását pedig *diszkrét eloszlásnak*, ha csak véges, vagy megszámlálhatóan végtelen sok értéket vesz fel. Ekkor a ξ értékei tehát egy véges, vagy megszámlálhatóan végtelen $x_1, x_2, \dots, x_n, \dots$ sorozatba rendezhetők, s az eloszlásfüggvény helyett célszerűbb a

$$p_k = P(\xi = x_k)$$

valószínűségeket használni. Nyilván a $\{p_k\}$ valószínűségek sorozata egyértelműen meghatározza az eloszlást és az eloszlásfüggvényt, ugyanis

$$P(\xi \in E) = \sum_{x_k \in E} p_k$$

és

$$F(x) = P(\xi < x) = \sum_{x_k < x} p_k.$$

Ilyenkor szokás a $\{p_k\}$ sorozatot a valószínűségi változó eloszlásának nevezni.

Diszkrét valószínűségi változóra példa egy tetszőleges eseményhez tartozó *indikátor változó.* Ez a következőképpen értelmezhető. Legyen A tetszőleges esemény, és legyen

$$\xi_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A. \end{cases}$$

Vizsgáljuk meg, hogy ξ_A valóban valószínűségi változó-e. Legyen x tetszőleges valós szám, ekkor nyilván

$$(\xi_A < x) = \begin{cases} \emptyset & x \leq 0, \\ \overline{A} & 0 < x \leq 1, \\ \Omega & 1 < x. \end{cases}$$

Látható, hogy ξ_A eleget tesz a valószínűségi változó definíciójában megfogalmazott követelménynek. A fentiek alapján a ξ_A eloszlásfüggvénye:

$$F_{\xi_A}(x) = P(\xi_A < x) = \begin{cases} 0 & x \leq 0, \\ 1 - P(A) & 0 < x \leq 1, \\ 1 & 1 < x. \end{cases}$$

2) Folytonos valószínűségi változó. A ξ valószínűségi változót és annak eloszlását *folytonosnak* nevezzük, ha van olyan $f : \mathbb{R} \to [0, +\infty)$ integrálható függvény, mellyel bármely $a < b$ esetén a ξ változó F eloszlásfüggvényére

$$F(b) - F(a) = \int_a^b f(x)dx$$

teljesül. (A folytonos valószínűségi változó tehát nem tévesztendő össze egy olyan valószínűségi változóval, amelynek az eloszlásfüggvénye folytonos!) Ebben az esetben az $a \to -\infty$ és $b = x$ helyettesítéssel kapjuk:

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

Az ilyen f függvényt a ξ valószínűségi változó *sűrűségfüggvényének* nevezzük. Nyilván

$$\int_{-\infty}^{+\infty} f(t)dt = 1,$$

azaz a sűrűségfüggvény integrálja az egész számegyenesen 1-gyel egyenlő. Folytonos valószínűségi változó minden értéket 0 valószínűséggel vesz fel. Továbbá érvényes

$$F'(x) = f(x),$$

tehát az eloszlásfüggvény a sűrűségfüggvényt bizonyos értelemben egyértelműen meghatározza. Ezzel kapcsolatban meg kell jegyeznünk a következőt. Mivel a sűrűségfüggvény definíciójában csupán annak integrálja szerepel különböző intervallumokon, ezért nem várható, hogy az eloszlásfüggvény pontonként egyértelműen meghatározza a, sűrűségfüggvényt hiszen ha két nem negatív integrálható függvény integrálja minden intervallumon egyenlő, abból még nem következik, hogy a két függvény egyenlő. Ennek a problémának a pontos tisztázása az úgynevezett Lebesgue-féle mérték- és integrálelméleten belül lehetséges, melynek tárgyalása meghaladja ezen jegyzet kereteit. Jegyezzük meg, hogy a sűrűségfüggvény csak annyiban egyértelmű, hogy tetszőleges olyan integrálható függvény hozzáadható, melynek integrálja minden intervallumon zérus. Így pl. a sűrűségfüggvényt véges sok pontban tetszőlegesen meg lehet változtatni.

Az eloszlásfüggvény tulajdonságaiból azonnal következnek a sűrűségfüggvény következő tulajdonságai: bármely valószínűségi változó sűrűségfüggvénye nem negatív integrálható függvény, melynek integrálja az egész számegyenesen 1. Megjegyezzük, hogy az ilyen függvények mindig tekinthetők valamilyen valószínűségi mezőn értelmezett valószínűségi változó sűrűségfüggvényének.

Folytonos valószínűségeloszlásra példa az *egyenletes eloszlás*, melyet a következőképpen értelmezhetünk. Legyenek $a < b$ tetszőleges valós számok, s válasszunk ki az (a, b) intervallumból tetszőlegesen egy x pontot. Ha a ξ valószínűségi változó értéke ez az x, akkor a geometriai valószínűséggel kapcsolatban mondottak alapján annak valószínűsége, hogy a ξ értéke az (a, b) intervallum egy (c, d) részintervallumába esik, arányos a részintervallum hosszával, azaz bármely $a < c < d < b$ esetén

$$P(c < \xi < d) = \frac{d - c}{b - a} = \int_{c}^{d} \frac{1}{b - a} dx.$$

Ezért ξ valószínűségi változó, melynek sűrűségfüggvénye

$$f(x) = \begin{cases} 0 & x < a, \\ \frac{1}{b-a} & a < x < b, \\ 0 & b < x. \end{cases}$$

(A sűrűségfüggvényről mondottak alapján az f-et az a és b pontokban tetszőlegesen értelmezhetjük.) Ennek alapján az eloszlásfüggvény

$$F(x) = \begin{cases} 0 & x \le a, \\ \frac{x-a}{b-a} & a < x \le b, \\ 1 & b < x. \end{cases}$$

A diszkrét és a folytonos eloszlásokon kívül vannak további eloszlások is, amelyek tanulmányozása kívül esik ezen jegyzet tárgykörén. A következőkben részletesen fogunk foglalkozni bizonyos nevezetes diszkrét és folytonos valószínűségeloszlásokkal. Megjegyezzük, hogy a valószínűségeloszlások tanulmányozása során általában a valószínűségi mezőtől és a valószínűségi változótól, annak konkrét értékeitől legtöbbször eltekinthetünk, csupán az eloszlásnak, az eloszlásfüggvénynek, illetve folytonos esetben a sűrűségfüggvénynek az ismerete szükséges. Látni fogjuk, hogy bizonyos, a gyakorlatban fontos szerepet játszó nevezetes eloszlások definícióját eleve pusztán az eloszlás-, illetve sűrűségfüggvény segítségével adjuk meg.

9.4. Feladatok

9.1. Két kockát dobunk fel egyszerre. A dobott számok összegét tekintsük valószínűségi változónak. Írjuk fel és ábrázoljuk a valószínűségi változó valószínűség-eloszlását és eloszlásfüggvényét!

9.2. Két kockát dobunk fel egyszerre. A dobott számok különbségének abszolút értékét tekintsük valószínűségi változónak. Írjuk fel és ábrázoljuk a valószínűségi változó valószínűség-eloszlását és eloszlásfüggvényét, továbbá számítsuk ki annak a valószínűségét, hogy a valószínűségi változó értéke 2-nél nagyobb, de a 4-et nem haladja meg!

9.3. Két kockát dobunk fel egyszerre. A dobott számok szorzatát tekintsük valószínűségi változónak. Írjuk fel a valószínűségi változó eloszlásfüggvényét!

9.4. Egy diszkrét valószínűségi változó lehetséges értékei -1 és 6, melyeket rendre $\frac{1}{3}$ és $\frac{2}{3}$ valószínűséggel vesz fel. Írjuk fel a valószínűségi változó eloszlásfüggvényét, számítsuk ki értékét az $x = 0$ helyen, és ábrázoljuk az eloszlásfüggvényt!

9.5. Legyen ξ olyan valószínűségi változó, amelynek lehetséges értékei az összes nem negatív egész számok. A valószínűségi változó a k értéket $p_k = \frac{2}{3^{k+1}}$ valószínűséggel veszi fel. Mutassuk meg, hogy a p_k számok valószínűségeloszlást alkotnak!

9.6. Legyen ξ olyan valószínűségi változó, amelynek lehetséges értékei az összes pozitív egész számok. A valószínűségi változó a k értéket $p_k = \frac{3}{4^k}$ valószínűséggel veszi fel. Mutassuk meg, hogy a p_k számok valószínűségeloszlást alkotnak!

9.7. Legyen ξ olyan valószínűségi változó, amelynek lehetséges értékei az összes pozitív egész számok. A valószínűségi változó a k értéket $p_k = \frac{1}{k(k+1)}$ valószínűséggel veszi fel. Mutassuk meg, hogy a p_k számok valószínűségeloszlást alkotnak!

9.8. Egy valószínűségi változó eloszlásfüggvénye

$$F(x) = \begin{cases} 0 & \text{ha } x < 1, \\ (x-1)^3 & \text{ha } 1 \le x < 2, \\ 1 & \text{ha } x \ge 2. \end{cases}$$

Határozzuk meg és ábrázoljuk a valószínűségi változó sűrűségfüggvényét!

9.9. Egy valószínűségi változó sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{2}{3} & \text{ha } 0 \le x < 1, \\ \frac{1}{3} & \text{ha } 1 \le x < 2, \\ 0 & \text{egyébként.} \end{cases}$$

Határozzuk meg és ábrázoljuk a valószínűségi változó eloszlásfüggvényét!

9.10. Egy ξ valószínűségi változó sűrűségfüggvénye

$$f(x) = \begin{cases} 0 & \text{ha } x < 0, \\ cx^2 & \text{ha } 0 \le x \le 2, \\ 0 & \text{egyébként.} \end{cases}$$

Határozzuk meg c értékét és írjuk fel ξ eloszlásfüggvényét! Mekkora valószínűséggel esik ξ az $(1, 3)$ intervallumba?

9.11. Egy ξ valószínűségi változó sűrűségfüggvénye

$$f(x) = \begin{cases} 0 & \text{ha } x \le 2, \\ \frac{a}{x^3} & \text{ha } x > 2. \end{cases}$$

Határozzuk meg a értékét és írjuk fel ξ eloszlásfüggvényét! Milyen x értékre adódik a $P(\xi > x)$ valószínűségre $\frac{1}{2}$?

9.12. Egy ξ valószínűségi változó sűrűségfüggvénye

$$f(x) = \begin{cases} 0 & \text{ha } x \le 0, \\ \frac{1}{\sqrt{x}} & \text{ha } 0 < x \le \frac{1}{4}, \\ 0 & \text{egyébként.} \end{cases}$$

Határozzuk meg ξ eloszlásfüggvényét! Mennyi annak a valószínűsége, hogy a ξ-nek a 0-tól való eltérése kisebb, mint 0.1?

9.13. Adjuk meg az ötös lottón kihúzott öt szám közül a legkisebb eloszlásfüggvényének értékét az $x = 25$ pontban!

9.14. Egységnyi hosszúságú szakaszt találomra választott pontjával két részre osztva mi a keletkezett szakaszok közül a rövidebbik hosszának az eloszlásfüggvénye?

9.15. A $(0, 1)$ intervallumban jelöljünk ki találomra három pontot. Határozzuk meg a középső pont abszcisszájának az eloszlásfüggvényét!

9.16. Válasszunk az egységnégyzetben egy pontot véletlenszerűen. Jelölje ξ a pontnak a négyzet legközelebbi oldalától való távolságát. Határozzuk meg ξ eloszlásfüggvényét!

9.17. Az alábbi számsorozatok közül melyek alkotnak valószínűségeloszlást?

i) $\binom{17}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{3}\right)^{17-k}$, $k = 1, 2, \ldots, 17$;

ii) $\frac{\binom{10}{k}\binom{10}{5-k}}{\binom{20}{5}}$, $k = 0, 1, 2, 3, 4, 5$;

iii) $p^3, 3p^2q, 3pq^2, q^3$, $(q = 1 - p, 0 \leq p \leq 1)$.

9.18. Az alábbi függvények közül melyek sűrűségfüggvények?

i)
$$f(x) = \begin{cases} \frac{\sin x}{2} & \text{ha } 0 < x < 1, \\ 0 & \text{egyébként;} \end{cases}$$

ii)
$$f(x) = \begin{cases} \frac{1}{x^2} & \text{ha } x > 1, \\ 0 & \text{egyébként;} \end{cases}$$

iii)
$$f(x) = \begin{cases} \frac{x}{x+1} & \text{ha } 0 < x < \infty, \\ 0 & \text{egyébként.} \end{cases}$$

9.19. Az alábbi függvények közül melyek sűrűségfüggvények?

i)
$$f(x) = \begin{cases} \frac{1}{3} & \text{ha } 0 < x < 1, \\ 0 & \text{egyébként;} \end{cases}$$

ii)
$$f(x) = \frac{1}{\pi}\frac{1}{(1 + x^2)};$$

iii)
$$f(x) = \begin{cases} \frac{1}{x^2} & \text{ha } x > 1 \\ 0 & \text{egyébként.} \end{cases}$$

9.20. Az alábbi függvények közül melyek eloszlásfüggvények?

i)

$$F(x) = \begin{cases} 0 & \text{ha } x < 0, \\ \frac{x}{x+1} & \text{ha } x \geq 0; \end{cases}$$

ii)

$$F(x) = \begin{cases} 0 & \text{ha } x \leq 0, \\ 1 & \text{ha } x > 0; \end{cases}$$

iii)

$$F(x) = \begin{cases} 0 & \text{ha } x < 0, \\ 2, & \text{ha } 0 \leq x < 1, \\ 1 & \text{ha } x \geq 1. \end{cases}$$

10. Eloszlások

10.1. Valószínűségi változók függvényeinek eloszlása

A gyakorlatban gyakran találkozunk a következő problémával: adott a ξ valószínűségi változó, valamint egy φ valós változós, valós értékű függvény, és meg akarjuk határozni az $\eta = \varphi(\xi)$ valószínűségi változó eloszlását, illetve eloszlásfüggvényét a ξ eloszlásfüggvényének ismeretében. A probléma matematikailag a következőképpen fogalmazható: legyen adott az $(\Omega, \mathcal{A}, P)$ valószínűségi mező és ezen a ξ valószínűségi változó, valamint adott a $\varphi : \mathbb{R} \to \mathbb{R}$ függvény. Az $\eta = \varphi(\xi)$ definícióval egy újabb véletlen függvényt definiálhatunk, melyről a φ függvényre tett egyszerű feltevések (pl. folytonosság, monotonitás, stb.) mellett következik, hogy ugyancsak valószínűségi változó. Kérdés: ha a ξ eloszlásfüggvénye F, akkor hogyan fejezhetjük ki η eloszlásfüggvényét? A válasz egyszerűen megfogalmazható a sűrűségfüggvények nyelvén.

10.1.1. Tétel *Legyen f a ξ valószínűségi változó sűrűségfüggvénye, melynek értékkészlete intervallum, továbbá legyen φ differenciálható, szigorúan monoton függvény azon az intervallumon, amelyből f az értékeit felveszi. Ekkor $\eta = \varphi(\xi)$ is valószínűségi változó, melynek sűrűségfüggvénye*

$$g(y) = f(\varphi^{-1}(y))\left|\frac{d\varphi^{-1}}{dy}\right|.$$

Bizonyítás. Azt az esetet tárgyaljuk, amikor φ szigorúan monoton növekvő. Ekkor az η eloszlásfüggvénye

$$G(y) = P(\eta < y) = P(\varphi(\xi) < y) = P(\xi < \varphi^{-1}(y) = F(\varphi^{-1}(y)),$$

amiből az állítás y szerinti differenciálással következik.

Tekintsük a következő egyszerű példát: legyen $\varphi(x) = ax+b$, ahol $a \neq 0$, b adott konstansok, tehát legyen
$$\eta = a\xi + b.$$
Ekkor η a ξ-nek lineáris függvénye. Nyilván
$$\varphi^{-1}(y) = \frac{y-b}{a},$$
így a fenti formula szerint az η sűrűségfüggvénye
$$g(y) = \frac{1}{|a|} f\Big(\frac{y-b}{a}\Big)$$
módon fejezhető ki a ξ változó f sűrűségfüggvényével.

10.2. Együttes eloszlás

Ha több valószínűségi változót egyidejűleg vizsgálunk, akkor többnyire nem elégedhetünk meg az egyes valószínűségi változók eloszlásainak ismeretével, hiszen általában a változók között különféle kapcsolatok állhatnak fenn, melyekre nézve az eloszlásaik külön-külön semmiféle felvilágosítást nem nyújtanak. Legyen például ξ és η két valószínűségi változó, melyek mindegyike az $1, 2, 3$ értékeket veszi fel, egyaránt $\frac{1}{3}$ valószínűséggel. Kérdés: mi annak a valószínűsége, hogy a ξ is és az η is pl. 1-et vesz fel értékül? Ha A_i, ill. B_j jelöli azt az eseményt, hogy a ξ értéke i, illetve az η értéke j $(i, j = 1, 2, 3)$, akkor nyilván $p_i = P(A_i) = q_j = B_j = \frac{1}{3}$ $(i, j = 1, 2, 3)$. A kérdés azonban az $A_1 \cdot B_1$ esemény valószínűsége. Ha bevezetjük az
$$r_{ij} = P(A_i \cdot B_j)$$
jelölést $i, j = 1, 2, 3$ esetén, akkor nyilván
$$r_{i1} + r_{i2} + r_{i3} = 1$$
és
$$r_{1j} + r_{2j} + r_{3j} = 1$$
hacsak $i, j = 1, 2, 3$. Ettől a két feltételtől eltekintve azonban az r_{ij} számok tetszőleges nem negatív számok lehetnek: a p_i és q_j valószínűségek nem határozzák meg egyértelműen értéküket. Megjegyezzük, hogy ha az A_i és B_j események függetlenek, akkor nyilván $r_{ij} = p_i q_j$ $(i, j = 1, 2, 3)$, ez azonban általában nem teljesül.

A két valószínűségi változó együttes vizsgálata helyettesíthető a (ξ, η) valószínűségi vektorváltozó vizsgálatával, melynek lehetséges értékei a sík pontjai. Ha E a sík egy részhalmaza, akkor a $P((\xi, \eta) \in E)$ számok összessége egy valószínűségeloszlást határoz meg a síkon. Ezt a két valószínűségi változó *együttes eloszlásának* nevezzük. A két valószínűségi változó *együttes eloszlásfüggvénye* ekkor az

$$F(x, y) = P(\xi < x; \eta < y)$$

módon értelmezett függvény. Az egyváltozós esethez hasonlóan itt is beszélhetünk *diszkrét valószínűségi vektorváltozóról* és *folytonos valószínűségi vektorváltozóról.* A diszkrét esetben, ha ξ lehetséges értékei az $x_1, x_2, \ldots$ számok, η lehetséges értékei pedig az $y_1, y_2, \ldots$ számok, akkor a valószínűségeloszlás az $r_{ij} = P(\xi = x_i; \eta = y_j)$ számokkal jellemezhető. Folytonosnak akkor nevezzük a (ξ, η) vektorváltozót, ha van olyan kétváltozós nem negatív h függvény, hogy bármely $a < b$, $c < d$ valós számok esetén

$$P(a \le \xi < b; c \le \eta < d) = \int_a^b \int_c^d h(x, y) dx dy$$

teljesül. Ekkor a $\{(\xi, \eta) \in E\}$ esemény valószínűsége

$$P((\xi, \eta) \in E) = \int \int_E h(x, y) dx dy.$$

A h függvényt *együttes sűrűségfüggvénynek* nevezzük.

A fenti fogalmak általánosítása kettőnél több valószínűségi változóra természetes módon végezhető el.

Míg a ξ és az η valószínűségi változók eloszlásai nem határozzák meg egyértelműen az együttes eloszlásfüggvényt, megfordítva más a helyzet: a H együttes eloszlásfüggvényből

$$F(x) = \lim_{y \to +\infty} H(x, y)$$

és

$$G(y) = \lim_{x \to +\infty} H(x, y)$$

módon kiszámítható a ξ és az η eloszlásfüggvénye. Ezeket az eloszlásokat *peremeloszlásnak* nevezzük. Folytonos eloszlású ξ és η esetén a sűrűségfüggvények is könnyen nyerhetők az együttes sűrűségfüggvényből az

$$f(x) = \int_{-\infty}^{+\infty} h(x, y) dy$$

és

$$g(y) = \int_{-\infty}^{+\infty} h(x, y) dx$$

formulák alapján. Legyen például a ξ és η együttes sűrűségfüggvénye $h(x,y) = \frac{6}{5}(x+y^2)$ ha $0<x<1$ és $0<y<1$, és 0 egyébként. Ekkor

$$f(x) = \int_{-\infty}^{+\infty} h(x,y)dy = \frac{6}{5}\int_0^1 (x+y^2)dy = \frac{6}{5}\Big(x+\frac{1}{3}\Big)$$

ha $0<x<1$ és 0 egyébként, továbbá

$$g(y) = \int_{-\infty}^{+\infty} h(x,y)dx = \frac{6}{5}\int_0^1 (x+y^2)dx = \frac{6}{5}\Big(\frac{1}{2}+y^2\Big)$$

ha $0<y<1$ és 0 egyébként.

10.3. Valószínűségi változók függetlensége

Az események függetlenségének fogalmához hasonlóan nagy jelentőséggel bír a valószínűségi változók függetlenségének fogalma. A ξ és η valószínűségi változókat *függetleneknek* nevezzük, ha együttes eloszlásfüggvényük egyenlő eloszlásfüggvényeik szorzatával. Ebben az esetben tehát a peremeloszlások egyértelműen meghatározzák az együttes eloszlást. Jelölje F, illetve G a ξ, illetve η eloszlásfüggvényét, H pedig az együttes eloszlásfüggvényt. Ekkor tehát a függetlenség egyenértékű a

$$H(x,y) = F(x)G(y)$$

feltétellel. Ha ξ és η diszkrét valószínűségi változók, lehetséges értékeik rendre az $x_1, x_2, \dots$, illetve $y_1, y_2, \dots$ számok, melyeket rendre $p_1, p_2, \dots$, illetve $q_1, q_2, \dots$ valószínűségekkel vesznek fel, akkor nyilván a (ξ,η) vektorváltozó is diszkrét, mely az (x_i, y_j) vektorértékeket veszi fel r_{ij} valószínűségekkel $(i,j = 1,2,\dots)$. A ξ és η függetlensége ebben az esetben nyilván ekvivalens azzal, hogy

$$r_{ij} = p_i q_j$$

teljesül minden $i,j = 1,2,\dots$-re. Ha ξ és η folytonos eloszlásúak f és g sűrűségfüggvényekkel, akkor együttes eloszlásfüggvényüket H-val jelölve nyilván

$$H(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{\partial^2 H(t,s)}{\partial t \partial s} dtds = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{\partial^2 F(t)G(s)}{\partial t \partial s} dtds =$$

$$= \int_{-\infty}^{x}\int_{-\infty}^{y} f(t)g(s)dtds,$$

ami azt jelenti, hogy az együttes eloszlás is folytonos, és sűrűségfüggvénye $h(x, y) = f(x)g(y)$. Megfordítva, ebből a feltételből integrálással megkapjuk, hogy következik a ξ és η függetlensége. Így két folytonos valószínűségi változó akkor és csak akkor független, ha együttes sűrűségfüggvényük egyenlő sűrűségfüggvényeik szorzatával.

A fentiek illusztrálására lássunk egy egyszerű példát. Tekintsünk egy négyzet alakú céltáblát, melyre véletlenszerűen adunk le lövéseket. Tegyük fel, hogy minden lövés eltalálja a céltáblát, és a céltábla egyenlő területű tartományait egyforma valószínűséggel találjuk el. A céltáblát reprezentálja a $[0, 1] \times [0, 1]$ egységnégyzet, és jelölje ξ, η a találati hely két koordinátáját. Ekkor

$$P(\xi < x) = x,$$

$$P(\eta < y) = y,$$

ha $0 < x \leq 1$ és $0 < y \leq 1$, továbbá

$$P(\xi < x; \eta < y) = xy,$$

ha $0 < x \leq 1$ és $0 < y \leq 1$. Tehát, ha x és y az egységnégyzetből van, akkor fennáll a $H(x, y) = F(x)G(y)$ egyenlőség. Könnyű belátni, hogy ez fennáll akkor is, ha x és y valamelyike az egységnégyzeten kívül van, tehát ξ és η függetlenek. Megjegyezzük, hogy kör alakú céltábla esetén a találati hely derékszögű koordinátái nem függetlenek.

10.4. Eloszlások kompozíciója

Az alkalmazásokban gyakran vetődnek fel olyan problémák, amelyeknél a vizsgált valószínűségi változó két független valószínűségi változó összege, tehát

$$\zeta = \xi + \eta,$$

ahol ξ és η függetlenek. Ilyen például két dobókockával való dobás esetén a dobott számok összege. Felmerül a kérdés, hogyan határozható meg ζ eloszlása a ξ és η eloszlásának ismeretében. Először tételezzük fel, hogy ξ és η diszkrét valószínűségi változók, melyek az egész számokat veszik fel értékül, továbbá legyen

$$P(\xi = i) = p_i, \qquad P(\eta = k) = q_k, \qquad P(\zeta = n) = r_n,$$

ahol $i, k, n = 0, \pm 1, \pm 2, \ldots$. Ekkor nyilvánvaló, hogy

$$r_n = P(\zeta = n) = \sum_{k=-\infty}^{+\infty} P(\xi = n - k, \eta = k).$$

Mivel a ξ és η függetlenek, így fennáll a

$$P(\xi = n-k, \eta = k) = P(\xi = n-k)P(\eta = k) = p_{n-k}q_k$$

egyenlőség minden n, k esetén. Ezért

$$r_n = \sum_{k=-\infty}^{+\infty} p_{n-k}q_k.$$

Általánosabban, ha $\{p_i\}$ és $\{q_k\}$ tetszőleges diszkrét valószínűségeloszlások $(i, k = 0, \pm 1, \pm 2, \dots)$, és

$$r_n = \sum_{k=-\infty}^{+\infty} p_{n-k}q_k,$$

akkor könnyű belátni, hogy $\{r_n\}$ is egy valószínűségeloszlás, melyet a $\{p_i\}$ és $\{q_k\}$ eloszlások *kompozíciójának* nevezünk. Ha $p_i = q_k = 0$ minden negatív i és k esetén, akkor

$$r_n = \sum_{k=0}^{n} p_{n-k}q_k.$$

A fentiek analógiájára értelmezhető két folytonos eloszlás kompozíciója is. Anélkül, hogy a részletekbe bocsátkoznánk, ha a ξ és η folytonos eloszlású valószínűségi változók sűrűségfüggvényei rendre f és g, akkor a

$$h(x) = \int_{-\infty}^{+\infty} f(x-y)g(y)dy$$

definícióval egy h sűrűségfüggvényt kapunk, melyhez tartozó eloszlást a ξ és η eloszlásai kompozíciójának nevezünk. Ha negatív x-re f és g eltűnik, akkor

$$h(x) = \int_{0}^{x} f(x-y)g(y)dy.$$

Ha ξ és η függetlenek, akkor h a $\xi + \eta$ sűrűségfüggvénye.

10.5. Feltételes eloszlás

A feltételes valószínűség általánosításának tekinthető a feltételes eloszlás fogalma. Legyen adott egy ξ valószínűségi változó, és egy pozitív valószínűségű

A esemény. A ξ valószínűségi változónak az A eseményre vonatkozó *feltételes eloszlásfüggvényének* nevezzük az

$$F_A(x|A) = P(\xi < x|A)$$

módon értelmezett $F_A : \mathbb{R} \to \mathbb{R}$ függvényt. Valójában a ξ valószínűségi változó feltételes eloszlásfüggvénye úgy is felfogható, mint közönséges eloszlásfüggvénye az $(\Omega, \mathcal{A}, P_A)$ valószínűségi mezőn, ahol $P_A = P(\cdot|A)$, az A esemény által indukált feltételes valószínűség. Ezek alapján látható, hogy a feltételes eloszlásfüggvény a közönséges eloszlásfüggvényhez hasonló tulajdonságokkal rendelkezik, ezeket itt nem soroljuk fel. Megjegyezzük, hogy ha $A_1, A_2, \ldots$ pozitív valószínűségű eseményekből álló teljes eseményrendszer, akkor a teljes valószínűség tételéből következik, hogy fennáll a következő azonosság:

$$F(x) = \sum_{i=1}^{\infty} F_{A_k}(x|A_k)P(A_k),$$

ahol F jelöli a ξ (közönséges) eloszlásfüggvényét.

Ha az F_A feltételes eloszlásfüggvény differenciálható, és egyenlő integrálható deriváltjának integrálfüggvényével, akkor deriváltját a ξ valószínűségi változó A eseményre vonatkozó *feltételes sűrűségfüggvényének* nevezzük:

$$f_A(x|A) = F'_A(x|A).$$

Ekkor ismét a teljes valószínűség tétele alapján fennáll

$$f(x) = \sum_{i=1}^{\infty} f_{A_k}(x|A_k)P(A_k),$$

ahol f jelöli a ξ (közönséges) sűrűségfüggvényét.

Az eddigiekben a feltételes eloszlást csak olyan A eseményre vonatkozóan értelmeztük, amely pozitív valószínűségű. Bizonyos esetekben azonban szükség van ennek a definíciónak a kiterjesztésére. Különösen fontos az az eset, amikor az A esemény az, hogy egy η valószínűségi változó egy adott y értéket vesz fel. Legyenek tehát ξ és η valószínűségi változók, y valós szám. A ξ valószínűségi változó $\eta = y$ *feltétel melletti feltételes eloszlásfüggvényét* a következő módon értelmezzük:

$$F_\eta(x|y) = \lim_{\Delta y \to 0} P(\xi < x | y \leq \eta < y + \Delta y),$$

feltéve, hogy a jobb oldalon álló határérték létezik. Ez a helyzet például akkor, amikor ξ és η együttes eloszlása folytonos. Ekkor természetesen ξ és η mindegyike folytonos eloszlású. Jelölje H az együttes eloszlásfüggvényt, h az együttes sűrűségfüggvényt, továbbá jelölje F, illetve G a ξ, illetve η eloszlásfüggvényét, f, illetve g pedig a ξ, illetve η sűrűségfüggvényét. Ekkor a következő számítást végezhetjük el:

$$P(\xi < x | y \le \eta < y + \Delta y) = \frac{P(\xi < x, y \le \eta < y + \Delta y)}{P(y \le \eta < y + \Delta y)} = \frac{H(x, y + \Delta y) - H(x, y)}{G(y + \Delta y) - G(y)}.$$

Ha a számlálót és a nevezőt elosztjuk Δy-nal, és elvégezzük a $\Delta y \to 0$ határátmenetet, akkor a következőt kapjuk:

$$F_\eta(x|y) = \frac{\frac{\partial}{\partial y} H(x, y)}{g(y)}.$$

A feltételekből következően a *feltételes sűrűségfüggvény* is létezik, és

$$f_\eta(x|y) = \frac{\partial}{\partial x} F_\eta(x|y).$$

Nyilván fennáll

$$f_\eta(x|y) = \frac{h(x, y)}{g(y)}.$$

Ha az η változó $\xi = x$ feltétel melletti feltételes sűrűségfüggvényét $g_\xi(y|x)$ jelöli, akkor hasonlóan

$$g_\xi(y|x) = \frac{h(x, y)}{f(x)}.$$

Ha ezekből az egyenlőségekből kifejezzük h-t, akkor a következőt kapjuk:

$$f(x) = \int_{-\infty}^{\infty} f_\eta(x|y) g(y) dy,$$

és

$$g(y) = \int_{-\infty}^{\infty} g_\xi(y|x) f(x) dx,$$

ami tekinthető a teljes valószínűség tétele "folytonos" változatának.

10.6. Véletlen bolyongás

Ezt a fejezetet egy klasszikus problémának, a *véletlen bolyongás* problémájának vizsgálatával zárjuk. Legyen adott egy pont, amely a számegyenesen az origóból kiindulva véletlenszerűen egyforma valószínűséggel jobbra vagy balra haladva a következő egész koordinátájú pontba megy. Kérdés: mennyi annak a valószínűsége, hogy n lépés megtétele után a k koordinátájú pontban lesz?

Mivel páros számú lépéssel csak páros koordinátájú pontba juthatunk, páratlan számú lépéssel pedig páratlan koordinátájúba, ezért a keresett valószínűség 0, ha n és k különböző paritású. Tegyük fel tehát, hogy $n-k$ páros. Nyilván azt is feltehetjük, hogy $k \geq 0$, hiszen k és $-k$ esetén a keresett valószínűség ugyanannyi. A lehetséges utak, tehát a kedvező esetek száma 2^n, hiszen minden lépésnél két lehetőség közül lehet választani. A kedvező utak számát a következő módon határozhatjuk meg. Ha n lépésben a k pontba jutottunk, akkor k-val több esetben mentünk jobbra, mint balra. Ebből könnyen adódik, hogy a jobbra történő lépések száma $k + \frac{n-k}{2} = \frac{n+k}{2}$, a balra történő lépések száma pedig $\frac{n-k}{2}$. Az n lépés közül ezt az $\frac{n-k}{2}$ számú balra történő lépést

$$\binom{n}{\frac{n-k}{2}}$$

különböző módon választhatjuk ki, ennyi tehát a kedvező esetek száma, s így a keresett valószínűség

$$p_k = \binom{n}{\frac{n-k}{2}} \frac{1}{2^n},$$

ha $n-k$ páros szám. Speciálisan, ha n páros, akkor az origóba való visszaérkezés valószínűsége

$$p_0 = \binom{n}{\frac{n}{2}} \frac{1}{2^n}.$$

A problémát más megközelítésben is tárgyalhatjuk a valószínűségi változók segítségével. Legyenek a ξ_i valószínűségi változók értékei -1 vagy 1 aszerint, hogy az i-edik ugrás balra, vagy jobbra történt. Ekkor

$$P(\xi = \pm 1) = \frac{1}{2}, \quad i = 1, 2, \ldots,$$

továbbá a ξ_i valószínűségi változók függetlenek. Az n-edik lépés után a bolyongó pont helyét az

$$\eta_n = \xi_1 + \xi_2 + \cdots + \xi_n$$

valószínűségi változó adja meg. A független valószínűségi változók összegének eloszlására nyert eredményekből levezethető, hogy a $P(\eta_n = k)$ valószínűség éppen az, amit fentebb kiszámoltunk.

10.7. Feladatok

10.1. A ξ valószínűségi változó a -3, -2, -1, 0, 1, 2, 3 értékeket veszi fel, rendre 0.08, 0.14, 0.19, 0.27, 0.17, 0.09, 0.06 valószínűségekkel.
i) Határozzuk meg ξ^2 eloszlását!
ii) Határozzuk meg $|\xi - 1| + |\xi + 1|$ eloszlását!

10.2. Legyen a ξ valószínűségi változó F eloszlásfüggvénye folytonos és szigorúan monoton.
i) Milyen eloszlású $F(\xi)$?
ii) Milyen eloszlású $\ln \frac{1}{F(\xi)}$?

10.3. Legyen a ξ valószínűségi változó eloszlásfüggvénye F, sűrűségfüggvénye f. Írjuk fel az alábbi valószínűségi változók eloszlás- és sűrűségfüggvényeit:
i) $\xi + c$ (c konstans);
ii) $c\xi$ ($c \neq 0$ konstans);
iii) e^{ξ}.

10.4. Legyen ξ és η együttes sűrűségfüggvénye a következő: $h(x, y) = x + y$, ha $0 \leq x \leq 1$, $0 \leq y \leq 1$ és $h(x, y) = 0$ egyébként. Határozzuk meg $\xi + \eta$ sűrűségfüggvényét!

10.5. Legyen ξ sűrűségfüggvénye $\lambda e^{-\lambda x}$, ha $x > 0$ és 0 egyébként, η sűrűségfüggvénye $\mu e^{-\mu x}$, ha $x > 0$ és 0 egyébként, legyenek továbbá ξ és η függetlenek. Határozzuk meg $\frac{\xi}{\eta}$ sűrűségfüggvényét!

10.6. Bizonyítsuk be, hogy ha ξ és η független, azonos eloszlású, folytonos valószínűségi változók, akkor $P(\xi < \eta) = \frac{1}{2}$!

10.7. Bizonyítsuk be, hogy két esemény akkor és csak akkor független, ha indikátorváltozóik függetlenek!

10.8. Legyen a ξ és η együttes sűrűségfüggvénye

$$h(x, y) = \frac{4}{5}(x + xy + y),$$

ha $0 < x < 1$, $0 < y < 1$, egyébként pedig 0. Határozzuk meg a peremeloszlásokat!

10.9. Legyen a ξ és η együttes sűrűségfüggvénye

$$h(x, y) = f(x)g(y),$$

ahol f sűrűségfüggvény. Határozzuk meg a $\min(\xi, \eta)$ és $\max(\xi, \eta)$ együttes eloszlás-, és sűrűségfüggvényét!

10.10. Legyen a ξ és η együttes sűrűségfüggvénye

$$h(x, y) = \frac{1}{8\pi^2} \exp\Big(-\frac{x^2 + \pi^2 y^2}{8\pi^2}\Big).$$

Állapítsuk meg, hogy ξ és η függetlenek-e, s határozzuk meg ξ és η várható értékét és szórását!

10.11. Legyen ξ olyan valószínűségi változó, hogy adott valós K esetén a $(\xi < K)$ esemény pozitív valószínűségű. Határozzuk meg a ξ feltételes eloszlásfüggvényét a $\xi < K$ feltétel mellett! Feltéve, hogy ξ folytonos eloszlású, határozzuk meg feltételes eloszlásfüggvényét a $\xi = K$ feltétel mellett!

11. A várható érték

11.1. A várható érték értelmezése

A valószínűségi változókkal kapcsolatos legtöbb lényeges kérdés megválaszolható az eloszlás, illetve az eloszlásfüggvény ismeretében, így ezek a változó véletlen ingadozásainak jellemzésére is elegendőek. A gyakorlatban azonban gyakran szükség lehet olyan számadatokra, amelyek további, pontosabb felvilágosítást nyújtanak a valószínűségi változók viselkedéséről. Az egyik ilyen alapvető fontosságú adat a *várható érték*. A várható érték tulajdonképpen a valószínűségi változó értékeinek súlyozott átlaga. Ez persze nem tekinthető egzakt matematikai definíciónak, de például egy olyan valószínűségi változó esetében, amely véges sok értéket vesz fel, a várható érték valóban az értékek súlyozott átlaga, s a súlyok éppen a felvételi valószínűségek.

A következőkben megadjuk a várható érték pontos definícióját diszkrét és folytonos valószínűségi változók esetén.

Tekintsük először a diszkrét valószínűségi változók esetét. Tegyük fel, hogy a ξ valószínűségi változó az $x_1, x_2, \ldots$ értékeket veszi fel rendre a $p_1, p_2, \ldots$ valószínűségekkel. Ekkor tehát $\sum_{i=1}^{\infty} p_i = 1$. Ha a $\sum x_i p_i$ sor abszolút konvergens, akkor az

$$M(\xi) = \sum_{i=1}^{\infty} x_i p_i$$

számot a ξ valószínűségi változó *várható értékének* nevezzük. Ha az előbbi sor nem abszolút konvergens, akkor azt mondjuk, hogy a valószínűségi változónak nem létezik várható értéke. A fentiekben mondottakkal összhangban, ha például ξ véges sok értéket vesz fel, akkor mindig létezik várható értéke és ez éppen a fent említett súlyozott átlag.

Lássunk egy egyszerű példát. Tegyük fel, hogy két játékos, A és B két kockával a következő játékot játssza. Feldobják a két kockát, s ha legalább az egyiken hatos lesz, akkor A fizet B-nek 5 forintot, de ha egyiken se lesz hatos, akkor B fizet A-nak 3 forintot. Melyik játékos feltételei előnyösebbek? Ennek a kérdésnek a megválaszolásához számítsuk ki az A és a B játékos várható veszteségét. Jelentse ξ az A játékos veszteségét egy játszmában, ekkor ξ a 0 és az 5 értékeket veszi fel. A 0 értéket akkor veszi fel, ha egyik kockán se dobtak hatost, s ennek valószínűsége nyilván $\frac{25}{36}$, míg az 5 értéket az ellentett esemény bekövetkezésekor veszi fel, annak valószínűsége tehát $\frac{11}{36}$. Így a veszteség várható értéke

$$M(\xi) = 0 \cdot \frac{25}{36} + 5 \cdot \frac{11}{36} \approx 1.53.$$

Hasonlóan kiszámíthatjuk a B játékos várható veszteségét is, amely nyilván egyenlő az A játékos várható nyereségével. Ha a B veszteségét egy játékban η jelöli, akkor η lehetséges értékei 0 és 3, melyeket rendre $\frac{11}{36}$ és $\frac{25}{36}$ valószínűségekkel vesz fel, így várható értéke

$$M(\eta) = 0 \cdot \frac{11}{36} + 3 \cdot \frac{25}{36} \approx 2.08.$$

Ez azt mutatja, hogy az A játékos van kedvezőbb helyzetben, hiszen tiszta nyereségének várható értéke közelítőleg $2.08 - 1.53 = 0.55$ játszmánként.

Egy kétszemélyes játékot akkor nevezünk *méltányosnak*, ha a két játszó fél várható vesztesége (nyeresége) egyenlő. Az ilyen jellegű problémáknak számos alkalmazása van pl. a biztosítási matematikában.

Van olyan diszkrét valószínűségi változó is, amelynek nem létezik várható értéke. Legyen $P(\xi = (-1)^k \frac{2^k}{k}) = \frac{1}{2^k}$ $(k = 1, 2, \dots)$, ekkor a

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

sor konvergens, de nem abszolút konvergens, hiszen

$$\sum_{k=1}^{\infty} \frac{2^k}{k} \frac{1}{2^k} = \sum_{k=1}^{\infty} \frac{1}{k} = +\infty.$$

Így a ξ-nek nem létezik várható értéke.

A következőkben folytonos valószínűségi változók várható értékét értelmezzük. Legyen ξ folytonos valószínűségi változó, melynek sűrűségfüggvénye f. Ha az $\int x f(x) dx$ improprius integrál abszolút konvergens, akkor az

$$M(\xi) = \int_{-\infty}^{+\infty} x f(x) dx$$

számot a ξ várható értékének nevezzük. Ha a fenti improprius integrál nem abszolút konvergens, akkor azt mondjuk, hogy ξ-nek nem létezik várható értéke.

A korábbiakban szerepelt az (a, b) intervallumban egyenletes eloszlású ξ valószínűségi változó, melynek sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{ha } a < x < b \\ 0 & \text{egyébként.} \end{cases}$$

Ekkor a fenti improprius integrál valójában közönséges integrál egy véges intervallumon, tehát a várható érték létezik és értéke

$$M(\xi) = \int_{-\infty}^{+\infty} x f(x) dx = \int_a^b \frac{x}{b-a} dx =$$

$$= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

Tekintsünk most egy úgynevezett *Cauchy*[12]*-eloszlású* ξ valószínűségi változót, melynek sűrűségfüggvénye

$$f(x) = \frac{1}{\pi(1+x^2)} \quad -\infty < x < +\infty.$$

(Mutassuk meg, hogy ez valóban sűrűségfüggvény!) Ekkor a megfelelő improprius integrál

$$\int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)} dx = +\infty,$$

tehát ennek a valószínűségi változónak nem létezik várható értéke.

11.2. A várható érték tulajdonságai

A következőkben néhány, a várható érték legfontosabb tulajdonságaival kapcsolatos tételt ismertetünk. Ezeket általában vagy a diszkrét, vagy a folytonos esetre bizonyítjuk be.

[12] Augustin Louis Cauchy francia matematikus, 1789-1857

11.2.1. Tétel *Ha ξ korlátos valószínűségi változó, azaz $k \leq \xi \leq K$, akkor létezik várható értéke és*

$$k \leq M(\xi) \leq K.$$

Bizonyítás. A tételt a folytonos esetre bizonyítjuk be. A feltételből következik, hogy a ξ változó f sűrűségfüggvénye 0-val egyenlő a (k, K) intervallumon kívül. Ezért

$$k = \int_k^K kf(x)dx \leq \int_k^K xf(x)dx = M(\xi) \leq \int_k^K kf(x)dx = K,$$

amit bizonyítani kellett.

11.2.2. Tétel *Ha a ξ várható értéke létezik, akkor bármely c konstans esetén létezik $c\xi$ várható értéke is, és*

$$M(c\xi) = cM(\xi).$$

Bizonyítás. Ha $c = 0$, akkor az állítás nyilvánvaló. Ha $c \neq 0$, akkor a folytonos esetben a ξ sűrűségfüggvényét f-el jelölve, a $c\xi$ sűrűségfüggvénye

$$g(x) = \frac{1}{|c|} f\left(\frac{x}{c}\right).$$

Ebből adódóan

$$M(c\xi) = \int_{-\infty}^{+\infty} xg(x)dx = \int_{-\infty}^{+\infty} \frac{x}{|c|} f\left(\frac{x}{c}\right)dx = c\int_{-\infty}^{+\infty} f(u)du = cM(\xi).$$

11.2.3. Tétel *Ha $\xi \geq 0$ és $M(\xi) = 0$, akkor $P(\xi = 0) = 1$.*

Bizonyítás. A diszkrét esettel foglalkozunk. Ha a ξ változó összes különböző értékei az $x_1, x_2, \ldots$ nem negatív, valós számok, rendre $p_1, p_2, \ldots$ valószínűségekkel, akkor a

$$\sum_k x_k p_k = 0$$

összefüggés alapján látható, hogy ha $x_k > 0$ valamely k-ra, akkor szükségképpen $p_k = 0$, s mivel $\sum_k p_k = 1$, ezért valamely k-ra $p_k > 0$, ekkor persze $x_k = 0$, így $p_k = 1$, ami éppen az állítás.

11.2.4. Tétel *Ha $P(\xi = c) = 1$ (c állandó), akkor $M(\xi) = c$.*

Bizonyítás. Az állítás mind a diszkrét, mind a folytonos esetre triviális.

11.2.5. Tétel *Ha létezik a ξ és η várható értéke, akkor létezik $\xi + \eta$ várható értéke is, és*

$$M(\xi + \eta) = M(\xi) + M(\eta).$$

Bizonyítás. Tekintsük a diszkrét eset bizonyítását. Legyenek a ξ lehetséges értékei az $x_1, x_2, \ldots$ számok $p_1, p_2, \ldots$ valószínűségekkel, s az η értékei $y_1, y_2, \ldots$ a q_1, $q_2, \ldots$ valószínűségekkel. Ekkor a $\zeta = \xi + \eta$ változó lehetséges értékei az $x_i + y_j$ alakban előállítható számok. Ha A_{jk} azt az eseményt jelenti, hogy $\xi = x_j$ és $\eta = y_k$, akkor tetszőleges z esetén

$$zP(\xi + \eta = z) = z \sum_{x_j + y_k = z} P(A_{jk}) = \sum_{x_j + y_k = z} (x_j + y_k) P(A_{jk}).$$

Ebből adódóan

$$M(\zeta) = \sum_j \sum_k (x_j + y_k) P(A_{jk}) = \sum_j x_j P(\xi = x_j) +$$

$$+ \sum_k y_k P(\eta = y_k) = M(\xi) + M(\eta),$$

amit bizonyítani kellett.

A tétel állítása teljes indukcióval kiterjeszthető kettő helyett tetszőleges véges számú összeadandóra, s a 11.2.2. tétellel kombinálva kapjuk a következő eredményt.

11.2.6. Tétel *Ha a $\xi_1, \xi_2, \ldots, \xi_n$ valószínűségi változóknak létezik várható értéke és $c_1, c_2, \ldots, c_n$ tetszőleges konstansok, akkor a $c_1\xi_1 + c_2\xi_2 + \cdots + c_n\xi_n$ valószínűségi változónak is létezik várható értéke, és*

$$M(c_1\xi_1 + c_2\xi_2 + \cdots + c_n\xi_n) = c_1 M(\xi_1) + c_2 M(\xi_2) + \cdots + c_n M(\xi_n).$$

A 11.2.6. tételben foglalt állítást a várható érték *linearitásának* szokás nevezni. Egy speciális esetben a 11.2.4. tételből kapjuk a következőt.

11.2.7. Következmény *Ha a ξ valószínűségi változónak létezik várható értéke, és c tetszőleges konstans, akkor a $\xi + c$-nek is létezik várható értéke, és*

$$M(\xi + c) = M(\xi) + c.$$

11.2.8. Tétel *Ha ξ és η független valószínűségi változók, melyeknek létezik várható értéke, akkor a $\xi\eta$ valószínűségi változónak is létezik várható értéke, és*

$$M(\xi\eta) = M(\xi)M(\eta).$$

Bizonyítás. A bizonyítást a diszkrét esetre végezzük el. Legyenek a ξ lehetséges értékei az $x_1, x_2, \ldots$ számok $p_1, p_2, \ldots$ valószínűségekkel, s az η értékei $y_1, y_2, \ldots$ a q_1, $q_2, \ldots$ valószínűségekkel. Ekkor a $\zeta = \xi \cdot \eta$ változó lehetséges értékei az $x_i \cdot y_j$ alakban előállítható számok. Ha A_{jk} azt az eseményt jelenti, hogy $\xi = x_j$ és $\eta = y_k$, akkor tetszőleges z esetén

$$zP(\xi \cdot \eta = z) = z \sum_{x_j \cdot y_k = z} P(A_{jk}) = \sum_{x_j \cdot y_k = z} (x_j \cdot y_k) P(A_{jk}).$$

Másrészt, a függetlenség miatt $P(A_{jk}) = P(\xi = x_j)P(\eta = y_k) = p_j q_k$. Ebből adódóan

$$M(\zeta) = \sum_j \sum_k (x_j \cdot y_k) P(A_{jk}) = (\sum_j x_j p_j)(\sum_k y_k q_k) = M(\xi)M(\eta),$$

amit bizonyítani kellett. Itt felhasználtuk, hogy két abszolút konvergens sor összeszorzásával kapott sor is abszolút konvergens.

Az előbbi tétel két független valószínűségi változóról tetszőleges véges számúra is kiterjeszthető teljes indukcióval, amennyiben azok teljesen függetlenek.

A következő egyenlőtlenség lényegében az analízisből ismert *Cauchy-Bunyakovszkij*[13]*-Schwarz*[14]*-féle egyenlőtlenség.*

11.2.9. Tétel *Ha a ξ és η olyan valószínűségi változók, melyek négyzetének létezik várható értéke, akkor a $\xi\eta$ változó várható értéke is létezik, és*

$$|M(\xi\eta)| \leq \sqrt{M(\xi^2)M(\eta^2)}.$$

---

[13] Viktor Jakovlevics Bunyakovszkij orosz matematikus, 1804-1889

[14] Hermann Amadeus Schwarz német matematikus, 1843-1921

Bizonyítás. Tekintsük tetszőleges, rögzített valós λ mellett a $\zeta = (\xi - \lambda\eta)^2$ valószínűségi változót. Mivel $0 \leq \zeta \leq 2\xi^2 + 2\lambda^2\eta^2$, ezért ζ várható értéke létezik. Másrészt

$$M(\zeta) = M(\xi^2) - 2M(\xi\eta) + \lambda^2 M(\eta^2) \geq 0,$$

a korábbi tételek alapján. Tehát $M(\zeta)$ a λ másodfokú polinomja, amely nem negatív. Ez csak úgy fordulhat elő, ha diszkriminánsa nem pozitív, tehát

$$4M(\xi\eta)^2 - 4M(\xi^2)M(\eta^2) \leq 0,$$

ami átrendezve éppen az állítást adja.

11.3. Valószínűségi változók függvényeinek várható értéke

A gyakorlatban időnként szükség van adott valószínűségi változók függvényeinek várható értékére. Ennek meghatározására szolgálnak a következő tételek.

11.3.1. Tétel *Ha $\xi_1, \xi_2, \ldots, \xi_n$ diszkrét valószínűségi változók, melyek lehetséges értékei $x_1^{(i)}, x_2^{(i)}, \ldots$ és φ egy n változós függvény, akkor az $\eta = \varphi(\xi_1, \xi_2, \ldots, \xi_n)$ valószínűségi változó várható értéke*

$$M(\eta) = \sum_{i_1, i_2, \ldots, i_n} \varphi(x_{i_1}^{(1)}, x_{i_2}^{(2)}, \ldots, x_{i_n}^{(n)}) \times$$

$$\times P(\xi_1 = x_{i_1}^{(1)}, \xi_2 = x_{i_2}^{(2)}, \ldots, \xi_n = x_{i_n}^{(n)}),$$

feltéve, hogy a jobb oldalon álló sor abszolút konvergens.

11.3.2. Tétel *Ha $\xi_1, \xi_2, \ldots, \xi_n$ folytonos valószínűségi változók, melyek együttes sűrűségfüggvénye h, és φ egy n változós függvény, melyre az $\eta = \varphi(\xi_1, \xi_2, \ldots, \xi_n)$ valószínűségi változó, akkor η várható értéke*

$$M(\eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \varphi(x_1, x_2, \ldots, x_n) \times$$

$$\times h(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n,$$

feltéve, hogy a jobb oldalon álló improprius integrál abszolút konvergens.

A két utóbbi tételt nem bizonyítjuk, helyette alkalmazásukra mutatunk két egyszerű példát.

Legyen először ξ diszkrét valószínűségi változó, mely az $x_1, x_2, \dots$ értékeket veszi fel rendre $p_1, p_2, \dots$ valószínűségekkel, és legyen φ valós függvény. Ekkor a 11.3.1. tétel alapján az $\eta = \varphi(\xi)$ valószínűségi változó várható értéke

$$M(\eta) = \sum_{k=1}^{\infty} \varphi(x_k)p_k,$$

feltéve, hogy a jobb oldalon álló sor abszolút konvergens. Hasonlóan, ha ξ folytonos változó az f sűrűségfüggvénnyel, akkor η várható értéke a 11.3.2. tétel alapján

$$M(\eta) = \int_{-\infty}^{+\infty} \varphi(x)f(x)dx,$$

feltéve, hogy a jobb oldalon álló improprius integrál abszolút konvergens.

Legyenek most ξ, η diszkrét valószínűségi változók, melyeknek létezik várható értéke, és legyen $\varphi(x, y) = x + y$. Ekkor a 11.3.1. tételből kapjuk

$$M(\xi + \eta) = M(\varphi(\xi, \eta)) = \sum_j \sum_k (x_j + y_k)P(\xi = x_j; \eta = y_k) = M(\xi) + M(\eta),$$

ami éppen a 11.2.5. tételben megismert formula. Hasonlóan kapjuk ezt az állítást a folytonos esetben is a 11.3.2. tételből.

11.4. A feltételes várható érték

A feltételes várható érték a feltételes eloszlásból származtatható. Ha tehát ξ diszkrét, és lehetséges értékei $x_1, x_2, \dots$, akkor

$$M(\xi|A) = \sum_{i=1}^{\infty} x_i P(\xi = x_i|A),$$

s ha az A eseményre vonatkozó feltételes eloszlás folytonos, akkor

$$M(\xi|A) = \int_{-\infty}^{\infty} x f_A(x|A)dx,$$

ahol f_A a feltételes sűrűségfüggvény.

Ha a ξ és η együttes eloszlása folytonos, akkor adott valós y esetén a ξ-nek az $\eta = y$ feltételre vonatkozó *feltételes várható értéke*

$$M(\xi|\eta = y) = \int_{-\infty}^{\infty} x f_\eta(x|y) dx.$$

Felhasználva a várható érték definícióját, könnyen adódik az

$$M(\xi) = \int_{-\infty}^{\infty} M(\xi|\eta = y) g(y) dy$$

egyenlőség, ahol g az η sűrűségfüggvénye. Természetesen a fenti számítások érvényességéhez fel kell teteleznünk, hogy a szereplő végtelen sorok, illetve improprius integrálok abszolút konvergensek.

Rögzített ξ és η esetén a valós y változó

$$h(y) = M(\xi|\eta = y)$$

módon értelmezett h függvényét a ξ valószínűségi változó η-ra vonatkozó *regressziójának* nevezzük. A ξ változó η-ra vonatkozó *feltételes várható értéke* alatt a következő valószínűségi változót értjük:

$$M(\xi|\eta) = h(\eta),$$

ahol h a ξ-nek az η-ra vonatkozó regressziója. Ezzel kapcsolatos a következő tétel, amely a fenti számításokból következik.

11.4.1. Tétel *Ha ξ és η folytonos valószínűségi változók, továbbá ξ-nek és $h(\eta)$-nak létezik várható értéke, akkor*

$$M[M(\xi|\eta)] = M(\xi).$$

11.5. Feladatok

11.1. Egy sorsjátékban, amelyben ezer sorsjegy vesz részt, 1 db 1000 forintos, 10 darab 100 forintos és 100 darab 20 forintos nyereményt sorsolnak ki. Számítsuk ki a kisorsolt nyeremény várható értékét!

11.2. Két személy, A és B két kockával játssza a következő játékot. Ha mindkét kockával páratlan számot dobnak, akkor az A játékos fizet a B-nek x összeget, ha pedig az egyik kockával páros, a másikkal páratlan számot dobnak, akkor B fizet A-nak ugyanennyit. Milyen x mellett lesz a játék igazságos?

11.3. Egy szabályos pénzérmét feldobunk. Ha az első dobás eredménye fej, akkor még kétszer dobunk, ha írás, akkor még egyszer. Mennyi a fej dobások számának várható értéke?

11.4. Egy játékos feldob egy kockát. Ha páratlan számot dob, veszít 1 Ft-ot, ha 6-ost dob, nyer 4 Ft-ot, ha 2-est vagy 4-est dob, újból dobhat. A második dobásnál 1 Ft-ot nyer, ha párost dob és 2 Ft-ot veszít, ha páratlant dob. Állapítsuk meg, hogy a játék a játékos számára előnyös, méltányos vagy hátrányos!

11.5. Egy dohányos ember mindennap 10 vagy 11 cigarettát szív p, illetve $1-p$ valószínűséggel. Mennyi az egy hónap (30 nap) alatt elszívott cigaretták számának várható értéke?

11.6. Egy valószínűségi változó sűrűségfüggvénye

$$f(x) = \begin{cases} 2-2x & \text{ha } 0 < x < 1, \\ 0 & \text{egyébként.} \end{cases}$$

Számítsuk ki a valószínűségi változó várható értékét!

11.7. Számítsuk ki az $f(x) = \frac{1}{2}e^{-|x|}$ $(-\infty < x < +\infty)$ sűrűségfüggvényű valószínűségi változó várható értékét!

11.8. Két ember asztaliteniszt játszik. A győztesnek három játszmát kell nyernie. Legyen p és q annak a valószínűsége, hogy egy játszmát az első játékos nyerjen, illetve veszítsen. Tegyük fel, hogy az egyes játszmák eredményei egymástól függetlenek. Mennyi a szükséges játszmák számának várható értéke?

11.9. Egy urnában n piros és egy fehér golyó van. Egymás után húzunk mindig egy golyót, és a kihúzottat minden egyes húzás után visszatesszük, továbbá minden egyes húzás után még egy piros golyót teszünk az urnába. Jelölje ξ annak a húzásnak a sorszámát, amikor először húzunk fehér golyót. Határozzuk meg a ξ várható értékét!

11.10. Bizonyítsuk be a diszkrét és folytonos eloszlások esetére, hogy ha a ξ valószínűségi változónak létezik várható értéke, és eloszlásfüggvénye F, akkor

$$M(\xi) = \int_0^{\infty} (1 - F(x))dx - \int_{-\infty}^{0} F(x)dx.$$

11.11. Egy ξ valószínűségi változó eloszlását szimmetrikusnak nevezzük az m számra nézve, ha bármely $\epsilon > 0$ esetén

$$P(m - \epsilon \leq \xi \leq m) = P(m \leq \xi \leq m - \epsilon).$$

Bizonyítsuk be diszkrét és folytonos eloszlások esetére, hogy ha ξ eloszlása szimmetrikus és létezik várható értéke, akkor $M(\xi) = m$!

11.12. Egy kockával addig dobunk, míg 6-os nem lesz az eredmény. Mennyi az addigi dobásszám várható értéke, ha az utolsó dobást is beleszámítjuk?

12. A szórás és a korrelációs együttható

12.1. A szórás értelmezése

Egy valószínűségi változó várható értéke csupán korlátozott mértékű felvilágosítást ad a változó ingadozásáról: azt a számot adja meg, amely körül az értékek ingadoznak, de az ingadozás mértékére vonatkozóan nem tartalmaz információt. Mivel a várható érték bizonyos értelemben a valószínűségi változó felvett értékeinek súlyozott átlaga, hasznos lenne tudni, hogy milyen az értékek szóródása ezen érték körül. Ezt a szóródást kézenfekvő lenne a valószínűségi változónak a saját várható értékétől való eltérés várható értékével mérni. Azonban, ha itt "eltérés" alatt különbséget értünk, akkor ez mindig 0 lenne, ha pedig az eltérést a különbség abszolút értékével mérnénk, akkor analitikusan kényelmetlenül kezelhető kifejezést kapnánk. A megfelelő fogalom a következő: a ξ valószínűségi változó *szórásának* nevezzük a

$$D(\xi) = \sqrt{M\Big((\xi - M(\xi))^2\Big)}$$

számot, amennyiben ξ-nek és $(\xi - M(\xi))^2$-nek létezik várható értéke. Ellenkező esetben azt mondjuk, hogy a ξ-nek nem létezik szórása.

Gyakran a fenti mennyiség négyzete, az úgynevezett *szórásnégyzet* kényelmesebben kezelhető. Ennek könnyebb kiszámításáról szól a következő tétel.

12.1.1. Tétel *A ξ valószínűségi változó szórása akkor és csak akkor létezik, ha létezik ξ-nek és ξ^2-nek várható értéke, s ekkor*

$$D^2(\xi) = M(\xi^2) - M(\xi)^2.$$

Bizonyítás. Az első állítás a

$$(\xi - M(\xi))^2 = \xi^2 - 2\xi M(\xi) + M(\xi)^2$$

azonosságból következik, továbbá

$$D^2(\xi) = M\Big((\xi - M(\xi))^2\Big) = M(\xi^2 - 2\xi M(\xi) + M(\xi)^2) = M(\xi^2) - M(\xi)^2.$$

Tekintsük például a kockadobási kísérletet, s jelölje ξ a dobás értékét. Ekkor nyilván

$$M(\xi) = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6},$$

továbbá

$$M(\xi^2) = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6},$$

amiből a ξ valószínűségi változó szórása

$$D(\xi) = \frac{91}{6} - \frac{21}{6}^2 = \sqrt{\frac{105}{36}}.$$

Speciálisan, diszkrét valószínűségi változó esetén

$$D^2(\xi) = \sum_{i=1}^{\infty} x_i^2 p_i - (\sum_{i=1}^{\infty} x_i p_i)^2,$$

folytonos esetben pedig

$$D^2(\xi) = \int_{-\infty}^{+\infty} x^2 f(x)dx - (\int_{-\infty}^{+\infty} xf(x)dx)^2,$$

feltéve, hogy a jobb oldalakon álló sorok, illetve improprius integrálok abszolút konvergensek.

A következő egyszerűen igazolható állítás a mechanikából ismert *Steiner*[15]-*formula* analogonja, mely azt fejezi ki, hogy egy valószínűségi változó tetszőleges állandótól való eltérése négyzetének a várható értéke akkor a legkisebb, ha az állandó éppen a várható érték.

12.1.2. Tétel *Ha létezik a ξ valószínűségi változó szórása, akkor tetszőleges m esetén*

$$M[(\xi - m)^2] = D^2(\xi) + (M(\xi) - m)^2.$$

[15] Jacob Steiner svájci matematikus, 1796-1863

12.2. A szórás tulajdonságai

12.2.1. Tétel *Ha létezik a ξ valószínűségi változó szórása, és a, b tetszőleges valós számok, akkor*

$$D^2(a\xi + b) = a^2 D^2(\xi).$$

Ebből a tételből következik, hogy egy konstans hozzáadásával a valószínűségi változó szórása nem változik, és konstanssal való szorzáskor a valószínűségi változó szórása a konstans abszolút értékével szorzódik.

12.2.2. Tétel *Ha $\xi_1, \xi_2, \ldots, \xi_n$ független valószínűségi változók, melyeknek létezik szórása, akkor összegüknek is létezik szórása és*

$$D^2(\xi_1 + \xi_2 + \cdots + \xi_n) = D^2(\xi_1) + D^2(\xi_2) + \cdots + D^2(\xi_n).$$

Bizonyítás. Ha ξ_i és ξ_j függetlenek, akkor fennáll

$$M[(\xi_i - M(\xi_i)(\xi_j - M(\xi_j))] = M[\xi_i - M(\xi_i)]M[\xi_j - M(\xi_j)] = 0,$$

ha $i \neq j$. Ennek felhasználásával

$$D^2(\xi_1 + \xi_2 + \cdots + \xi_n) =$$

$$= M[(\xi_1 + \xi_2 + \cdots + \xi_n - M(\xi_1 + \xi_2 + \cdots + \xi_n))^2] =$$

$$= M[(\xi_1 - M(\xi_1) + \xi_2 - M(\xi_2) + \cdots + \xi_n - M(\xi_n))^2] =$$

$$= \sum_{i=1}^{n} M[(\xi_i - M(\xi_i))^2] + 2\sum_{i<k} M[(\xi_i - M(\xi_i)(\xi_k - M(\xi_k))] = \sum_{i=1}^{n} D^2(\xi_i),$$

amit bizonyítani kellett.

12.3. A kovariancia és a korrelációs együttható

A 12.2.2. tétel bizonyításából látható, hogy az állítás érvényességéhez valójában nincs szükség a függetlenségre, csak annyit kell feltételeznünk, hogy

$$M[(\xi_i - M(\xi_i)(\xi_j - M(\xi_j))] = M[\xi_i - M(\xi_i)]M[\xi_j - M(\xi_j)] = 0,$$

ha $i \neq j$. Könnyű belátni: ez ekvivalens azzal, hogy

$$M(\xi_i \xi_j) = M(\xi_i)M(\xi_j),$$

ha $i \neq j$. Ez függetlenség esetén mindig fennáll, de annál gyengébb feltevés. Ezzel kapcsolatban bevezetjük a következő fontos fogalmat. A

$$c(\xi, \eta) = M[(\xi - M(\xi))(\eta - M(\eta))]$$

várható értéket – feltéve, hogy létezik – a ξ és η valószínűségi változók *kovarianciájának* nevezzük. Az $\eta = \xi$ esetben nyilván

$$c(\xi, \xi) = D^2(\xi),$$

tehát a kovariancia tekinthető bizonyos értelemben a szórásnégyzet általánosításának. Ha a ξ és az η nem állandó 1 valószínűséggel, akkor szórásuk pozitív, és ekkor az

$$r(\xi, \eta) = \frac{c(\xi, \eta)}{D(\xi)D(\eta)} = \frac{M[(\xi - M(\xi))(\eta - M(\eta))]}{D(\xi)D(\eta)}$$

számot a ξ és η *korrelációs együtthatójának* nevezzük. Mind a kovariancia, mind a korrelációs együttható a két valószínűségi változó függőségére nézve ad felvilágosítást. Független ξ és η esetén $c(\xi, \eta) = r(\xi, \eta) = 0$, ez azonban fordítva nem igaz. A 11.2.9. Tételből kapjuk a következőt:

12.3.1. Tétel *Ha létezik a ξ és η valószínűségi változók korrelációs együtthatója, akkor*

$$|r(\xi, \eta)| \leq 1.$$

Meg lehet mutatni, hogy $|r(\xi, \eta)| = 1$ pontosan akkor teljesül, ha ξ és η között lineáris függvénykapcsolat áll fenn 1 valószínűséggel. Ennek részleteivel itt nem foglalkozunk.

Mind a kovariancia, mind a korrelációs együttható képlete tovább alakítható, a következő módon:

$$c(\xi, \eta) = M(\xi\eta) - M(\xi)M(\eta),$$

és

$$r(\xi, \eta) = \frac{M(\xi\eta) - M(\xi)M(\eta)}{D(\xi)D(\eta)}.$$

Tekintsük a következő példát. Legyen a (ξ, η) valószínűségi vektorváltozó a következő módon értelmezve:

$$P((\xi, \eta) = (-1, 0)) = P((\xi, \eta) = (0, 1)) =$$

$$= P((\xi, \eta) = (0, -1)) = P((\xi, \eta) = (1, 0)) = \frac{1}{4}.$$

Ebből következik, hogy ξ és η a $-1, 0, 1$ értékeket veszik fel a következő valószínűségekkel:

$$P(\xi = -1) = P(\eta = -1) = \frac{1}{4},$$

$$P(\xi = 0) = P(\eta = 0) = \frac{1}{2},$$

$$P(\xi = 1) = P(\eta = 1) = \frac{1}{4}.$$

Innen adódik, hogy

$$M(\xi) = M(\eta) = 0,$$

tehát $c(\xi, \eta) = r(\xi, \eta) = 0$. Másrészt

$$0 = P(\xi = 0, \eta = 0) \neq P(\xi = 0)P(\eta = 0) = \frac{1}{4},$$

tehát ξ és η nem függetlenek.

Ha $r(\xi, \eta) = 0$, akkor azt mondjuk, hogy ξ és η *korrelálatlanok*. Ezek szerint független valószínűségi változók mindig korrelálatlanok, fordítva azonban ez nem igaz.

A fentiekben bevezetett várható érték, illetve szórás mellett a valószínűségi változók tanulmányozásának fontos segédeszközei az úgynevezett *momentumok*. Ha k természetes szám, akkor a ξ valószínűségi változó *k-adik momentumának* nevezzük a ξ^k valószínűségi változó várható értékét, amennyiben létezik. Így például az első momentum maga a várható érték, a szórásnégyzet pedig szoros kapcsolatban áll a második momentummal, hiszen a második momentumnak és a várható érték négyzetének különbsége.

12.4. Feladatok

12.1. Bizonyítsuk be, hogy egy diszkrét valószínűségi változó szórása akkor és csak akkor zérus, ha a valószínűségi változó 1 valószínűséggel konstans!

12.2. Határozzuk meg az (a, b) intervallumban egyenletes eloszlású valószínűségi változó szórását!

12.3. A ξ valószínűségi változó eloszlása legyen a következőképpen adott:

$$P(\xi = 1) = P(\xi = -1) = \frac{3}{8}, \;\; P(\xi = 2) = P(\xi = -2) = \frac{1}{8}.$$

i) Határozzuk meg a ξ szórását!
ii) Határozzuk meg a ξ és ξ^2 korrelációs együtthatóját!

12.4. A ξ valószínűségi változó sűrűségfüggvénye legyen

$$f(x) = \begin{cases} \frac{2x}{9} & \text{ha } 0 \le x < 3, \\ 0 & \text{egyébként.} \end{cases}$$

Határozzuk meg a ξ változó szórását!

12.5. Bizonyítsuk be a 12.3.1. tételt!

12.6. Bizonyítsuk be, hogy ha ξ és η valószínűségi változók, akkor $|r(\xi, \eta)| = 1$ akkor és csak akkor teljesül, ha vannak olyan $a \ne 0$ és b számok, hogy

$$P(\eta = a\xi + b) = 1.$$

12.7. Legyenek ξ és η független valószínűségi változók, melyeknek létezik a harmadik momentuma. Bizonyítsuk be, hogy

$$M[(\xi + \eta - m_1 - m_2)^3] = M[(\xi - m_1)^3] + M[(\eta - m_2)^3],$$

ahol $m_1 = M(\xi)$ és $m_2 = M(\eta)$.

13. Nevezetes valószínűségeloszlások

Ebben a fejezetben néhány speciális valószínűségeloszlással ismerkedünk meg, amelyek különösen fontosak a gyakorlati alkalmazások szempontjából. Egyik - másik közülük már előfordult a korábbiakban. Az alábbiakban összefoglaljuk ezen eloszlások legfontosabb tulajdonságait, esetenként meghatározzuk várható értéküket és szórásukat.

13.1. A binomiális eloszlás

A ξ valószínűségi változó eloszlását n-edrendű, p paraméterű *binomiális eloszlásnak* nevezzük, ha a valószínűségi változó a $k = 0, 1, \dots, n$ természetes számokat rendre

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \quad (k = 0, 1, \dots, n)$$

valószínűségekkel veszi fel. Gyakran használjuk a $q = 1 - p$ jelölést. Például a Bernoulli-féle kísérletsorozatban azoknak a kísérleteknek a száma, amelyekben az adott esemény bekövetkezik, binomiális eloszlású valószínűségi változó.

Azt a tényt, hogy a fenti $\{p_k\}$ $(k = 0, 1, \dots, n)$ számok valóban valószínűség-eloszlást alkotnak, könnyen igazolhatjuk a binomiális tétel alapján.

A binomiális eloszlás várható értéke a következő módon számítható ki a definíció alapján:

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

$$= np \sum_{k=1}^{n-1} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} =$$

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

A binomiális eloszlás szórása ennek alapján:

$$D^2(\xi) = M(\xi^2) - M(\xi)^2 = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} - (np)^2 =$$

$$= p^2 n(n-1) \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{n-k} + np - (np)^2 =$$

$$= p^2 n(n-1) \sum_{l=0}^{n-2} \binom{n-2}{l} p^{l} (1-p)^{n-2-l} + np - (np)^2 =$$

$$= p^2 n(n-1) + np - (np)^2 = np(1-p),$$

tehát

$$D(\xi) = \sqrt{np(1-p)} = \sqrt{npq}.$$

13.2. A hipergeometrikus eloszlás

A ξ valószínűségi változó eloszlását *hipergeometrikus eloszlásnak* nevezzük, ha a valószínűségi változó a $k = 0, 1, \ldots, n$ természetes számokat rendre

$$p_k = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \quad (k = 0, 1, \ldots, n)$$

valószínűségekkel veszi fel, ahol $0 < n \leq \min(M, N-M)$ és $M < N$.

A hipergeometrikus eloszlás a visszatevés nélküli mintavétel problémájához kapcsolódik. Ha az N elemű sokaságban M darab selejtes van, akkor az n elemű mintában található selejtes darabok száma hipergeometrikus eloszlású valószínűségi változó. Azt, hogy a fenti $\{p_k\}$ számok valószínűségeloszlást alkotnak, egyszerű számolással igazolhatjuk, ha az

$$(1+x)^M (1+x)^{N-M} = (1+x)^N$$

azonosság mindkét oldalát a binomiális tétel segítségével kifejtjük, majd az azonosság két oldalán összehasonlítjuk x^n együtthatóit.

A hipergeometrikus eloszlás várható értékét a következő módon kaphatjuk meg:

$$M(\xi) = \sum_{k=0}^{n} \frac{k\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} =$$

$$= \frac{M\binom{N-1}{n-1}}{\binom{N}{n}} \sum_{l=0}^{n-1} \frac{\binom{M-1}{l}\binom{(N-1)-(M-1)}{n-1-l}}{\binom{N-1}{n-1}} = \frac{Mn}{N}.$$

Ha a visszatevés nélküli mintavételnél az $\frac{M}{N}$ selejtarány jelölésére a p-t használjuk, azaz

$$p = \frac{M}{N},$$

akkor tehát $M(\xi) = np$, csakúgy, mint a visszatevéses mintavételnél.

Most meghatározzuk a hipergeometrikus eloszlás szórását. Mivel

$$D^2(\xi) = M(\xi^2) - M(\xi)^2 = \sum_{k=0}^{n} \frac{k^2\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} - \left[n\frac{M}{N}\right]^2,$$

először a jobb oldal első tagjával foglalkozunk:

$$M(\xi^2) = M\sum_{k=1}^{n}[(k-1)+1]\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\binom{N}{n}} =$$

$$= \frac{Mn}{N}\Big(1 + \frac{(M-1)(n-1)}{N-1}\Big).$$

Ebből adódóan

$$D^(\xi) = \frac{Mn}{N}\Big(1 + \frac{(M-1)(n-1)}{N-1}\Big) = \frac{Mn}{N}\Big(1 - \frac{M}{N}\Big)\Big(1 - \frac{n-1}{N-1}\Big).$$

Ismét alkalmazva a $p = \frac{M}{N}$ és $q = 1 - p$ jelölést, kapjuk

$$D(\xi) = \sqrt{npq\Big(1 - \frac{n-1}{N-1}\Big)}.$$

13.3. A Poisson[16]-eloszlás

A ξ valószínűségi változó eloszlását λ paraméterű *Poisson-eloszlásnak* nevezzük, ha a valószínűségi változó a $k = 0, 1, \ldots$ természetes számokat rendre

$$p_k = \frac{\lambda^k e^{-\lambda}}{k!} \qquad (k = 0, 1, \ldots)$$

valószínűségekkel veszi fel, ahol $\lambda > 0$ állandó.

A Poisson-eloszlás a következő típusú problémáknál játszik szerepet: radioaktív anyagban adott idő előtt elbomló atomok száma; egy telefonközpontba adott idő alatt befutó hívások száma; egy nyári éjszakán adott idő alatt észlelhető csillaghullások száma stb. Azt, hogy a fenti $\{p_k\}$ számok valószínűségeloszlást alkotnak, egyszerű számolással igazolhatjuk, ha felhasználjuk az exponenciális függvény sorfejtését.

A Poisson-eloszlás várható értékét a következő módon kaphatjuk meg:

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

Tehát a Poisson-eloszlás paramétere egyben a várható érték.

Most számítsuk ki a Poisson-eloszlás szórását.

$$D^2(\xi) = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k e^{-\lambda}}{k!} - \lambda^2 =$$

$$= \lambda \sum_{k=1}^{\infty} [(k-1)+1] \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!} - \lambda^2 =$$

$$= \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} - \lambda^2 =$$

$$= \lambda^2 e^{-\lambda} \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} + \lambda e^{-\lambda} \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} - \lambda^2 =$$

$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda e^{-\lambda} e^{\lambda} - \lambda^2 = \lambda,$$

[16]Simèon-Denis Poisson francia matematikus, 1781-1840

tehát
$$D(\xi) = \sqrt{\lambda}.$$

Így a λ paraméterű Poisson-eloszlás szórásnégyzete számszerűleg egyenlő a várható értékkel, ami éppen az eloszlás paramétere.

A Poisson-eloszlásnak – egyebek mellett – megvan az a fontos tulajdonsága, hogy jól közelíti a binomiális eloszlást. Pontosabban, ha az n-edrendű p paraméterű binomiális eloszlásnál n értéke minden határon túl nő, miközben p úgy tart 0-hoz, hogy az $np = \lambda$ szorzat állandó, akkor a határeloszlás a λ paraméterű Poisson - eloszlás. Ezt a következőképpen láthatjuk be.

Vezessük be a binomiális eloszlás tagjainak jelölésére a
$$P_k^{(n)} = \binom{n}{k} p^k (1-p)^{n-k} \qquad (k = 0, 1, \dots, n)$$
formulát. Megmutatjuk, hogy rögzített k és $np = \lambda$ esetén
$$\lim_{n\to\infty} P_k^{(n)} = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, \dots).$$

Az $np = \lambda$ egyenlőség alapján
$$P_k^{(n)} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$
$$= \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)-k =$$
$$= \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \frac{n(n-1)(n-2)\dots(n-k+1)}{n \cdot n \cdot \dots \cdot n} \left(1 - \frac{\lambda}{n}\right)^{-k}.$$
Mivel
$$\frac{n(n-1)(n-2)\dots(n-k+1)}{n \cdot n \cdot \dots \cdot n} = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{k-1}{n}\right) \to 1,$$
és
$$\left(1 - \frac{\lambda}{n}\right)^{-k} \to 1,$$
ha $n \to \infty$, valamint
$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda},$$
ezért a fentiekből állításunk következik.

Tegyük fel például, hogy n golyót N számú számozott urnába egymás után véletlenszerűen elhelyezünk. Annak a valószínűsége, hogy egy meghatározott urnába k golyó kerüljön

$$p_k = \binom{n}{k} \frac{1}{N^k} \Big(1 - \frac{1}{N}\Big)^{n-k}.$$

Ha az urnák számát és ezzel együtt a golyók számát minden határon túl növeljük, de közben

$$\frac{n}{N} = \lambda > 0$$

teljesül, akkor p_k határértéke a λ paraméterű Poisson-eloszlás k-adik tagja. Ez tehát annak a valószínűsége, hogy nagyon sok urnába nagyon sok golyót elhelyezve egy adott urnába k golyó kerüljön, ahol λ az egy urnára eső golyók száma.

A fenti eredmény lehetőséget ad arra, hogy a binomiális eloszlás tagjait a Poisson-eloszlás tagjaival közelítsük. Tehát nagy n esetén $\binom{n}{k}p^k(1-p)^{n-k}$ egy lehetséges közelítése $\frac{(np)^k}{k!}e^{-np}$.

13.4. A negatív binomiális eloszlás

A ξ valószínűségi változó eloszlását r-edrendű *negatív binomiális eloszlásnak* nevezzük, ha a valószínűségi változó a $k+r$ $(k = 0, 1, \dots)$ természetes számokat rendre

$$p_k = \binom{k+r-1}{r-1} p^r (1-p)^k \qquad (k = 0, 1, \dots)$$

valószínűségekkel veszi fel, ahol r nem negatív egész.

Tekintsünk egy kísérletet, amelynek két lehetséges kimenetele van: A és $\overline{A}$, továbbá $P(A) = p$ és $q = 1 - p$. Ha ezt a kísérletet többször egymás után végrehajtjuk, és $A_k^{(r)}$ azt az eseményt jelenti, hogy az A esemény r-edszerre a $k+r$-edik kísérletben következett be, akkor

$$P(A_k^{(r)}) = \binom{k+r-1}{r-1} p^r (1-p)^k \qquad (k = 0, 1, \dots).$$

A negatív binomiális eloszlás várható értékét a következő módon kaphatjuk meg:

$$M(\xi) = \sum_{k=0}^{\infty} (k+r) \binom{k+r-1}{r-1} p^r (1-p)^k = \frac{r}{p} \sum_{k=0}^{\infty} (k+r) \binom{k+r}{r} p^{r+1} (1-p)^k = \frac{r}{p}.$$

Speciálisan, ha $r = 1$, akkor az elsőrendű negatív binomiális eloszlás várható értéke a fentiek alapján $\frac{1}{p}$.

Most számítsuk ki a negatív binomiális eloszlás szórását:

$$D^2(\xi) = \sum_{k=0}^{\infty}(k+r)^2\binom{k+r-1}{r-1}p^r(1-p)^k - \left(\frac{r}{p}\right)^2.$$

A jobb oldal első tagját a következőképpen alakíthatjuk át:

$$\sum_{k=0}^{\infty}(k+r)^2\binom{k+r-1}{r-1}p^r(1-p)^k =$$

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

$$= r\sum_{k=0}^{\infty}\frac{(k+r+1)!}{r!k!}p^r(1-p)^k - \frac{r}{p}\sum_{k=0}^{\infty}\binom{k+r}{r}p^{r+1}(1-p)^k =$$

$$= \frac{r(r+1)}{p^2}\sum_{k=0}^{\infty}\binom{k+r+1}{r+1}p^{r+2}(1-p)^k - \frac{r}{p} = \frac{r(r+1)}{p^2} - \frac{r}{p}.$$

Ezért

$$D^2(\xi) = \frac{r(r+1)}{p^2} - \frac{r}{p} - \frac{r^2}{p^2} = \frac{(1-p)r}{p^2} = \frac{qr}{p^2}.$$

13.5. A geometriai eloszlás

Az elsőrendű negatív binomiális eloszlást *geometriai eloszlásnak* nevezzük. Ekkor a valószínűségi változó a $k = 1, 2, \ldots$ természetes számokat rendre

$$p_k = p(1-p)^{k-1} \qquad (k = 1, 2, \ldots)$$

valószínűségekkel veszi fel.

A geometriai eloszlás alkalmazására tekintsük a következő példát. Egy urnában N golyó van, ezek közül M piros. Az urnából visszatevéssel húzunk golyókat, találomra. Jelölje A_k azt az eseményt, hogy piros golyót először a k-adik alkalommal húzunk ($k = 1, 2, \ldots$). Bevezetve a $p = \frac{M}{N}$ jelölést, könnyű látni, hogy

$$P(A_k) = p(1-p)^{k-1},$$

tehát a piros golyó első alkalommal való kihúzásának sorszáma geometriai eloszlású.

A geometriai eloszlás várható értékét és szórását a fentiekben tárgyalt negatív binomiális eloszlás speciális eseteként kapjuk:

$$M(\xi) = \frac{1}{p},$$

és

$$D(\xi) = \frac{\sqrt{1-p}}{p}.$$

Az eddigiekben speciális diszkrét valószínűségeloszlásokkal foglalkoztunk, most pedig áttérünk néhány nevezetes folytonos valószínűségeloszlás tárgyalására. A folytonos valószínűségeloszlásokat általában sűrűségfüggvényükkel adjuk meg.

13.6. Az egyenletes eloszlás

Az egyenletes eloszlással már találkoztunk a korábbiakban, a 9. fejezetben. Most emlékeztetünk az értelmezésére. A ξ valószínűségi változó eloszlását az (a, b) intervallumban *egyenletes eloszlásnak* nevezzük, ha sűrűségfüggvénye

$$f(x) = \begin{cases} 0 & \text{ha } x < a, \\ \frac{1}{b-a} & \text{ha } a < x < b, \\ 0 & \text{ha } b < x. \end{cases}$$

Ekkor a ξ eloszlásfüggvénye

$$F(x) = \begin{cases} 0 & \text{ha } x \leq a, \\ \frac{x-a}{b-a} & \text{ha } a < x \leq b, \\ 1 & \text{ha } b < x. \end{cases}$$

Az egyenletes eloszlás várható értékét a 11. fejezetben már kiszámoltuk:

$$M(\xi) = \frac{a+b}{2}.$$

A szórás a következőképpen számítható ki:

$$D^2(\xi) = \frac{1}{b-a}\int_a^b x^2 dx - \Big(\frac{a+b}{2}\Big)^2 = \frac{(b-a)^2}{12},$$

tehát

$$D(\xi) = \frac{b-a}{2\sqrt{3}}.$$

13.7. Az exponenciális eloszlás

A ξ valószínűségi változó eloszlását λ paraméterű *exponenciális eloszlásnak* nevezzük, ha sűrűségfüggvénye

$$f(x) = \lambda e^{-\lambda x} \qquad (x \geq 0),$$

ahol $\lambda > 0$ pozitív állandó. Ekkor a ξ eloszlásfüggvénye

$$F(x) = 1 - e^{-\lambda x}, \text{ ha } x \geq 0.$$

Az exponenciális eloszlás a gyakorlatban általában véletlen hosszúságú időtartamok eloszlásaként fordul elő. Az exponenciális eloszlás várható értékét parciális integrálással kapjuk:

$$M(\xi) = \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}.$$

A szórás hasonló módon számítható ki:

$$D^2(\xi) = \lambda \int_0^\infty x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

A λ paraméterű exponenciális eloszlás szórása tehát számszerűleg megegyezik az eloszlás várható értékével.

13.8. A normális eloszlás

A ξ valószínűségi változó eloszlását (m, σ) paraméterű *normális eloszlásnak* nevezzük, ha sűrűségfüggvénye

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \qquad (-\infty < x < +\infty),$$

ahol m tetszőleges, $\sigma > 0$ pozitív állandó. Ekkor a ξ eloszlásfüggvénye

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}} dt.$$

Azt, hogy az f függvény valóban sűrűségfüggvény, úgy igazolhatjuk, hogy felhasználjuk a következő ismert összefüggést:

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

A normális eloszlás központi szerepet játszik a valószínűségszámításban, a matematikai statisztikában és ezek gyakorlati alkalmazásaiban. Ennek oka az, hogy nagy számú független valószínűségi változó összege mindig közelítően normális eloszlású. A normális eloszlást hibaeloszlásnak is szokás nevezni, mivel a mérési eredmények véletlen hibája általában normális eloszlású.

A normális eloszlás sűrűségfüggvényét *Gauss*[17]*-Laplace*[18]*-függvénynek* szokás nevezni, grafikonját pedig *haranggörbének*.

Az (m, σ) paraméterű normális eloszlást $\mathcal{N}(m, \sigma)$ eloszlásnak is nevezzük. Az $\mathcal{N}(0, 1)$ eloszlást *standard normális eloszlásnak* nevezzük, eloszlásfüggvényét, illetve sűrűségfüggvényét Φ, illetve φ jelöli. Tehát

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

és

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Mivel a normális eloszlás eloszlásfüggvényének definíciójában szereplő integrálfüggvény nem fejezhető ki elemi függvények segítségével, így értékei csak közelítőleg

[17] Carl Friedrich Gauss német matematikus, 1777-1855

[18] Pierre Simon Laplace francia matematikus, 1749-1827

határozhatók meg. A standard normális eloszlás eloszlásfüggvényének közelítő értékeit táblázatokba foglalták, tetszőleges paraméterekkel rendelkező normális eloszlás eloszlásfüggvényének értékeit pedig ezek segítségével a következő tétel alapján lehet kiszámítani.

13.8.1 Tétel *Az $\mathcal{N}(m,\sigma)$ eloszlás F eloszlásfüggvényére teljesül*

$$F(x) = \Phi\Big(\frac{x-m}{\sigma}\Big).$$

Bizonyítás. Az állítást az $u = \frac{t-m}{\sigma}$ helyettesítéssel kapjuk a következő módon:

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}}\,dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-m}{2}} e^{-\frac{u^2}{2}}\,du.$$

Most kiszámítjuk az $\mathcal{N}(m,\sigma)$ eloszlás várható értékét. A definíció alapján

$$M(\xi) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} xe^{-\frac{(x-m)^2}{2\sigma^2}}\,dx =$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\Big(\frac{x-m}{\sigma}\Big)e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx + \frac{m}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx = m,$$

hiszen a jobboldal első tagja 0, a második pedig az $\mathcal{N}(m,\sigma)$ eloszlás sűrűségfüggvénye integráljának az m-szerese.

Tehát a normális eloszlás m paraméterének jelentése: az eloszlás várható értéke.

Most meghatározzuk az $\mathcal{N}(m,\sigma)$ eloszlás szórását. A definíció értelmében

$$D^2(\xi) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}(x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}}\,dt,$$

amit a $t = \frac{x-m}{\sigma}$ helyettesítéssel kapunk. Innen parciális integrálással adódik:

$$D^2(\xi) = \sigma^2,$$

tehát

$$D(\xi) = \sigma,$$

azaz, a normális eloszlás σ paramétere éppen a szórás. Ezért az $\mathcal{N}(m,\sigma)$ eloszlás m várható értékű és σ szórású normális eloszlás. Speciálisan, a standard normális eloszlás várható értéke 0, szórása 1.

A gyakorlati feladatok megoldása során gyakran szükség van a normális eloszlásból származtatott t-eloszlásra, illetve a χ^2-eloszlásra. Ezek definíciója a következő.

Ha $\xi_1, \xi_2, \ldots, \xi_n$ és η független, $\mathcal{N}(0,1)$ eloszlású valószínűségi változók, akkor a

$$t = \frac{\sqrt{n}\eta}{\sqrt{\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2}}$$

valószínűségi változó eloszlását n *szabadsági fokú t-eloszlásnak* nevezzük.

Ha $\xi_1, \xi_2, \ldots, \xi_n$ független, $\mathcal{N}(0,1)$ eloszlású valószínűségi változók, akkor a

$$\chi^2 = \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2$$

valószínűségi változó eloszlását n *szabadsági fokú χ^2-eloszlásnak* nevezzük.

A t-eloszlás és a χ^2-eloszlás eloszlás-, illetve sűrűségfüggvényeinek értékeit táblázatok tartalmazzák. Alkalmazásukra a 17. fejezetben fogunk példát látni.

13.9. Feladatok

13.1. Jelölje ξ_A az A esemény indikátor változóját. Határozzuk meg a ξ_A eloszlásának várható értékét és szórását!

13.2. Adott 200 darab termék közül 60 darab elsőosztályú. Találomra kiválasztunk 6 darabot egymás után, visszatevéssel. Legyen a ξ valószínűségi változó a kiválasztott elsőosztályú darabok száma. Írjuk fel és ábrázoljuk a valószínűségi változó eloszlását, továbbá határozzuk meg várható értékét és szórását!

13.3. A tapasztalat alapján 1000 újszülöttből átlagosan 516 a fiú és 484 a lány. Mekkora a valószínűsége, hogy egy 6 gyermekes családban legfeljebb egy lány van?

13.4. Öt katona lő egy céltáblára. Mindegyikük 100 lövése közül átlag 30 talál. Mekkora a valószínűsége, hogy ha egyszerre leadnak egy-egy lövést, akkor legfeljebb három találat lesz a céltáblán?

13.5. Valamely 4000 darabból álló szállítmányban 400 darab selejt található. Visszatevéses módszerrel 10 elemű mintát veszünk. Mennyi annak a valószínűsége, hogy a mintában legfeljebb 3 selejt van?

13.6. Egy binomiális eloszlású valószínűségi változó várható értéke 4, szórása pedig $0.8\sqrt{5}$. Mekkora az n, p és q értéke?

13.7. Határozzuk meg, hogy milyen k esetén lesz maximuma a

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}$$

binomiális eloszlásnak!

13.8. A 32 lapos magyar kártyából találomra egyszerre 4 lapot húzunk. Határozzuk meg a kihúzott lapok számának eloszlását, várható értékét és szórását!

13.9. Legyen ξ $\lambda = 4$ paraméterű Poisson-eloszlású valószínűségi változó. Határozzuk meg várható értékét és szórását! Milyen valószínűséggel esik ξ a $(2, 5)$ intervallumba? Milyen valószínűséggel vesz fel ξ a várható értékénél kisebb értéket?

13.10. Egy telefonközpontba percenként átlagosan 4 hívás fut be. Mennyi annak a valószínűsége, hogy 1 perc alatt 6 hívás fut be?

13.11. Egy 400 oldalas nyomdai korrektúrában átlagosan 400 sajtóhiba található. Mennyi a valószínűsége annak, hogy egy találomra választott oldalon legalább három sajtóhiba van, ha feltételezzük, hogy a sajtóhibák száma Poisson - eloszlású?

13.12. Határozzuk meg a $(3,7)$ intervallumban egyenletes eloszlású valószínűségi változó eloszlás-, és sűrűségfüggvényét, valamint várható értékét és szórását!

13.13. Egy egyenletes eloszlású valószínűségi változó várható értéke 4, szórásnégyzete $\frac{4}{3}$. Írjuk fel az eloszlásfüggvényt!

13.14. Egy rádióállomás minden órában közli a pontos időt. Valaki bekapcsolja a rádiót. Mennyi a valószínűsége, hogy legfeljebb 10 percet kell várnia az időjelzésre, ha feltételezzük, hogy a rádió bekapcsolásának időpontja egyenletes eloszlású?

13.15. Egy gép működési időtartama az első meghibásodásig exponenciális eloszlású valószínűségi változó. Írjuk fel eloszlás- és sűrűségfüggvényét, ha várható értéke 2 év!

13.16. Egy benzinkútnál a tapasztalatok szerint a várakozási idő átlagosan 4 perc. Ha a várakozási idő exponenciális eloszlású, akkor mennyi annak a valószínűsége, hogy egy alkalommal 3 percnél többet, de 4 percnél kevesebbet kell várakozni?

13.17. Egy munkapadról kikerülő alkatrész hossza normális eloszlású, $m = 30$ cm várható értékkel és $\sigma = 0.2$ cm szórással.

i) Írjuk fel és ábrázoljuk a valószínűségi változó eloszlás-, és sűrűségfüggvényét!
ii) Mennyi annak a valószínűsége, hogy egy alkatrész hossza 29.7 cm és 30.3 cm közé esik?
iii) Mennyi annak a valószínűsége, hogy a várható értéktől való eltérés abszolút értéke 1 cm-nél kevesebb?
iv) Milyen pontosságot biztosíthatunk 0.9 valószínűséggel a munkadarabok hosszára?

13.18. A liszt csomagolásánál a csomagológép 1 kg várható súlyú csomagokat készít 2.5 dkg szórással. Feltehető, hogy a csomagolt mennyiség súlya normális eloszlású. Mennyi annak a valószínűsége, hogy egy csomagban 95 dkg-nál kevesebb liszt lesz?

13.19. Legyen ξ egy λ paraméterű exponenciális eloszlású valószínűségi változó. Határozzuk meg $\sqrt{\xi}$ sűrűségfüggvényét!

13.20. Legyen ξ $\mathcal{N}(m,\sigma)$ eloszlású valószínűségi változó. Határozzuk meg e^{ξ} sűrűségfüggvényét!

13.21. Legyen ξ $\mathcal{N}(0,1)$ eloszlású valószínűségi változó. Határozzuk meg ξ^2 sűrűségfüggvényét!

13.22. Legyen ξ 1 paraméterű exponenciális eloszlású valószínűségi változó. Határozzuk meg $|\xi - 2|$ eloszlásfüggvényét, sűrűségfüggvényét és várható értékét!

13.23. Legyen ξ $\mathcal{N}(3, 2)$ eloszlású valószínűségi változó. Határozzuk meg a $P(-2 < \xi < 1)$ és a $P(0.1 < |\xi|)$ valószínűségeket!

13.24. Határozzuk meg m és σ értékét, ha tudjuk, hogy a ξ valószínűségi változó $\mathcal{N}(m, \sigma)$ eloszlású, és $P(\xi > 0) = 0.5$, $P(-1 < \xi < 1) = 0.1$!

13.25. Határozzuk meg a b értékét, ha tudjuk, hogy a ξ valószínűségi változó $\mathcal{N}(0, 1)$ eloszlású, és $P(\xi \geq b) = 0.65$!

13.26. Legyen ξ $\mathcal{N}(3, 2)$ eloszlású valószínűségi változó. Határozzuk meg a $|\xi - 2|$ sűrűségfüggvényét, várható értékét, szórását, valamint a $P(0.1 < |\xi - 2|)$ valószínűséget!

13.27. Határozzuk meg a standard normális eloszlású valószínűségi változó összes momentumait!

13.28. Határozzuk meg a λ paraméterű exponenciális eloszlású valószínűségi változó összes momentumait!

13.29. Legyenek ξ és η λ, illetve μ paraméterű exponenciális eloszlású valószínűségi változók. Határozzuk meg a $\xi + \eta$ sűrűségfüggvényét!

13.30. Határozzuk meg a 13.20. feladatban szereplő e^{ξ} valószínűségi változó várható értékét!

13.31. Legyenek ξ, η független $\mathcal{N}(0, 1)$ eloszlású valószínűségi változók. Határozzuk meg $|\xi + \eta|$ sűrűségfüggvényét!

13.32. Legyenek $\xi_1, \xi_2, \ldots, \xi_n$ független λ paraméterű exponenciális eloszlású valószínűségi változók. Adjuk meg $\min(\xi_1, \xi_2, \ldots, \xi_n)$, $\max(\xi_1, \xi_2, \ldots, \xi_n)$ és $\xi_1 + \xi_2 + \cdots + \xi_n$ eloszlás- és sűrűségfüggvényét!

13.33. Legyenek ξ, η független egyenletes eloszlású valószínűségi változók a $(0, 1)$, illetve a $(0, 2)$ intervallumon. Határozzuk meg az $\frac{1}{\xi}$ és a $\frac{\xi}{1+\xi}$ eloszlásfüggvényét, valamint a $\xi + \eta$ sűrűségfüggvényét!

13.34. Legyenek a ξ, η, ζ valószínűségi változók függetlenek, s rendre $\mathcal{N}(2, 1)$, $\mathcal{N}(3, \sqrt{2})$ és $\mathcal{N}(4, \sqrt{3})$ eloszlásúak. Határozzuk meg a következő valószínűségeket:

i) $P(\xi \leq \eta)$;
ii) $P(3\xi - 2\eta > 1)$;
iii) $P(\xi + \eta \leq 2\zeta - 4)$;

iv) $P(\xi \leq \eta; \zeta < 5)$.

13.35. Legyenek $\xi_1, \xi_2, \ldots, \xi_n$ független, egyenletes eloszlású valószínűségi változók a $(0,1)$ intervallumon. Határozzuk meg a nagyság szerinti k-adik sűrűségfüggvényét, várható értékét és szórását!

13.36. Egy nagy egyetem diákjainak 30%-a albérletben él. Mennyi a valószínűsége, hogy 200 találomra kiválasztott diák közül 50 és 75 között van az albérletben élők száma?

13.37. Legyen ξ λ paraméterű Poisson-eloszlású valószínűségi változó. Határozzuk meg a $P(5 \leq \xi \leq 14)$ valószínűséget!

13.38. Egy automata gép 2 kg lisztet rak egy-egy zacskóba, de véletlen ingadozás következtében a zacskókban levő liszt mennyisége $\mathcal{N}(m, \sigma)$ eloszlású. Előzetes megfigyelésekből tudni lehet, hogy $\sigma = 0.002$, valamint, hogy annak a valószínűsége, hogy egy zacskóban kevesebb van 2 kg-nál 0.01. Határozzuk meg m értékét! Milyen σ mellett biztosíthatjuk, hogy a fenti valószínűség 0.001 legyen?

13.39. Budapesten meg akarják állapítani a dohányosok arányát. Ehhez megkérdeznek n egyént úgy, hogy minden választásnál mindenki ugyanakkora eséllyel jöhet szóba (visszatevéses mintavétel). Milyen nagyra kell az n-t választani, hogy a megkérdezettek között a dohányosok aránya legalább 0.95 valószínűséggel 0.005-nél nem nagyobb hibával közelítse meg a dohányosok valódi arányát? Közelítsünk normális eloszlással!

13.40. Egy csokoládégyár egyik termékének a tömege (grammban mérve) $\mathcal{N}(8; 0.5)$ eloszlású valószínűségi változó. A termékeket egyesével, elkészülésük sorrendjében rakják zacskókba. Minimálisan hány darabot kell egy zacskóba tenni ahhoz, hogy legalább 0.999 valószínűséggel a csomag tiszta tömege legalább 200 g legyen?

13.41. Egy alkatrész működési ideje (órában mérve) $\mathcal{N}(20000; 1500)$ eloszlású. Ha az alkatrész 15000 óránál rövidebb ideig működik, akkor gyárilag selejtesnek minősül. Az alkatrészeknek átlagosan hány százaléka lesz selejtes?

13.42. Legyenek ξ és η független $\mathcal{N}(0,1)$ eloszlású valószínűségi változók. Milyen eloszlású a $\xi^2 + \eta^2$ valószínűségi változó?

13.43. Bizonyítsuk be, hogy ha két normális eloszlású valószínűségi változó korrelálatlan, akkor függetlenek!

14. Generátorfüggvények

14.1. A generátorfüggvény értelmezése

A fentiekben láttuk, hogy a diszkrét valószínűségeloszlások vizsgálata valójában egyenértékű olyan nem negatív tagú valós számsorozatok vizsgálatával, amelyekből képzett sor összege 1. Az ilyen sorozatok hatékonyabb vizsgálatához jól felhasználhatók az úgynevezett generátorfüggvények. Mint látni fogjuk, a generátorfüggvények módszerének többek között az az előnye, hogy segítségével a diszkrét sorozatokra vonatkozó egyes problémák analitikus függvényekre vonatkozó problémákra fogalmazhatók át, így lehetővé válik a matematikai analízis eszköztárának felhasználása.

Egy $\{a_k\}$ valós számsorozat *generátorfüggvényén* a $\sum a_k x^k$ hatványsor összegfüggvényét értjük a hatványsor konvergenciaintervallumában. Az analízisből ismeretes, hogy egy ilyen hatványsor mindig konvergens egy 0 körüli nyílt intervallumban, ha

$$\limsup_k (a_k)^{\frac{1}{k}}$$

véges. Ha ez az érték pozitív, akkor reciprokát R-rel jelölve, a $(-R, R)$ intervallumban a hatványsor konvergens, összegfüggvénye analitikus függvény, melynek tetszőleges deriváltját a hatványsor megfelelő számú tagonkénti differenciálásával kapjuk. Ha a fenti határérték zérus, akkor a hatványsor mindenütt konvergens.

Számunkra azok az esetek lesznek érdekesek, amikor az $\{a_k\}$ sorozat tagjai valószínűségeloszlást alkotnak. Ekkor a fenti határérték legfeljebb 1, így a megfelelő hatványsor biztosan konvergens a $(-1, 1)$ intervallumban.

Legyen ξ diszkrét valószínűségi változó, melynek eloszlása $\{p_k\}$ $(k = 0, 1, 2, \dots)$. A ξ valószínűségi változó, illetve a $\{p_k\}$ eloszlás *generátorfüggvénye* a

$$G(s) = \sum_{k=0}^{\infty} p_k s^k$$

függvény a $(0,1)$ intervallumon. Látható, hogy azonos eloszlású valószínűségi változók generátorfüggvényei egyenlők, így joggal beszélhetünk egy diszkrét valószínűségeloszlás generátorfüggvényéről. Mivel a fenti definícióból azonnal adódik, hogy

$$p_k = \frac{G^{(k)}(0)}{k!} \qquad (k = 0, 1, 2, \dots),$$

ezért a generátorfüggvény egyértelműen meghatározza az eloszlást.

14.2. A generátorfüggvény alkalmazásai

A generátorfüggvény segítségével könnyen kiszámíthatók a valószínűségeloszlások fontosabb mérőszámai. Erről szól a következő tétel.

14.2.1. Tétel *Ha létezik a ξ valószínűségi változó várható értéke, akkor G generátorfüggvénye az $s = 1$ helyen balról differenciálható, és*

$$M(\xi) = G'(1).$$

Ha létezik ξ szórásnégyzete, akkor G az $s = 1$ helyen balról kétszer differenciálható, és

$$D^2(\xi) = G''(1) + G'(1) - (G'(1))^2.$$

A tételt nem bizonyítjuk. Megjegyezzük, hogy a tétel általánosabban is érvényes: ha létezik a ξ k-adik momentuma (s ekkor minden k-nál alacsonyabb rendű is), akkor G az $s = 1$ pontban balról k-szor differenciálható, és $G^{(k)}(1)$ kifejezhető a ξ legfeljebb k-adrendű momentumai segítségével.

Legyen például ξ_A egy A esemény indikátor változója, ekkor a megfelelő eloszlás: $p_0 = P(\xi_A = 0) = 1 - p$, $p_1 = P(\xi_A = 1) = p$, ahol p az A esemény valószínűsége. A ξ generátorfüggvénye:

$$G(s) = p_0 \cdot s^0 + p_1 \cdot s^1 = 1 - p + ps.$$

A fenti tételnek megfelelően

$$M(\xi) = P(A) = p = G'(1),$$

és

$$D^2(\xi) = p - p^2 = G''(1) + G'(1) - (G'(1))^2.$$

Mivel a generátorfüggvény csupán az eloszlástól függ, magától a valószínűségi változótól, illetve annak értékeitől nem, mindig feltételezhetjük egy adott eloszlással kapcsolatban, hogy ez egy olyan valószínűségi változó eloszlása, amely a természetes számokat veszi fel értékül, és

$$p_k = P(\xi = k) \qquad (k = 0, 1, 2, \dots).$$

Ekkor bármely rögzített $(-1, 1)$ intervallumbeli s esetén az $\eta = s^\xi$ valószínűségi változó az $1, s, s^2, \dots$ értékeket veszi fel ugyanezekkel a valószínűségekkel és várható értéke

$$M(\eta) = M(s^\xi) = \sum_{k=0}^{\infty} s^k p_k = G(s),$$

tehát a ξ valószínűségi változó generátorfüggvénye az s helyen éppen az s^ξ valószínűségi változó várható értékével egyenlő. Ebből kapjuk a következő fontos tételt.

14.2.2. Tétel *Ha $\xi_1, \xi_2, \dots, \xi_n$ nem negatív, egész értékeket felvevő, független valószínűségi változók, melyek generátorfüggvényei G_1, G_2, $\dots$, G_n, akkor az $\eta = \xi_1 + \xi_2 + \cdots + \xi_n$ valószínűségi változó generátorfüggvénye $G_1 \cdot G_2 \cdot \ldots \cdot G_n$.*

Bizonyítás. Ha $\xi_1, \xi_2, \dots, \xi_n$ függetlenek, akkor ugyancsak függetlenek az s^{ξ_1}, s^{ξ_2}, $\dots$, s^{ξ_n} valószínűségi változók is, ahol s a $(-1, 1)$ intervallum tetszőleges eleme. Ezért a várható érték tulajdonságai alapján az η valószínűségi változó G_η generátorfüggvénye

$$G_\eta(s) = M(s^\eta) = M(s^{\xi_1+\xi_2+\cdots+\xi_n}) =$$

$$= M(s^{\xi_1} s^{\xi_2} \dots s^{\xi_n}) = M(s^{\xi_1}) M(s^{\xi_2}) \dots M(s^{\xi_n}) = G_1(s) G_2(s) \dots G_n(s),$$

amit bizonyítanunk kellett.

A binomiális eloszlás generátorfüggvényét az előbbi tétel segítségével a következő módon határozhatjuk meg. Ha η binomiális eloszlású valószínűségi változó, akkor összege n független, $\xi_1, \xi_2, \dots, \xi_n$ valószínűségi változónak, melyek csak a 0 és 1 értékeket veszik fel, mégpedig az 1-et p valószínűséggel. Ezek azonos eloszlásúak, így generátorfüggvényük a definíció alapján

$$G(s) = s^0 \cdot (1 - p) + s^1 \cdot p = 1 + p(s - 1).$$

Ezért az η generátorfüggvénye

$$h(s) = [G(s)]^n = [1 + p(s - 1)]^n.$$

Ebből a 14.2.1. tétel alapján ismételten megkaphatjuk a binomiális eloszlás várható értékét és szórását.

A Poisson-eloszlás generátorfüggvényét is könnyen kiszámíthatjuk a definíció alapján:

$$G(s) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} s^k =$$

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

Ha ξ és η független Poisson-eloszlású valószínűségi változók rendre λ, illetve μ paraméterrel, akkor összegük generátorfüggvénye a 14.2.2. tétel és az előzők alapján

$$h(s) = e^{\lambda(s-1)} e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)},$$

ami a $\lambda+\mu$ paraméterű Poisson-eloszlás generátorfüggvénye. Mivel a generátorfüggvény az eloszlást egyértelműen meghatározza, ezzel bebizonyítottuk a következő tételt:

14.2.3. Tétel *Két független Poisson-eloszlású valószínűségi változó összege is Poisson-eloszlású, s paramétere egyenlő az összeadandók paramétereinek összegével.*

Most határozzuk meg az r-edrendű negatív binomiális eloszlás generátorfüggvényét. Mivel ez r független geometriai eloszlás összegeként kapható, így előbb a geometriai eloszlás generátorfüggvényét számítjuk ki:

$$G(s) = \sum_{k=0}^{\infty} pq^{k-1} s^k = \frac{ps}{1-qs},$$

s ebből az r-edrendű negatív binomiális eloszlás generátorfüggvénye

$$G(s) = \left(\frac{ps}{1-qs}\right)^r.$$

14.3. Feladatok

14.1 Egy urnában két piros és egy fehér golyó van. Kétszer húzunk visszatevéssel. Határozzuk meg a piros golyók számának generátorfüggvényét!

14.2. Két kockával addig dobunk, amíg mind a két kockán hatost nem kapunk. Adjuk meg a szükséges dobások számának generátorfüggvényét!

14.3 Legyen a ξ valószínűségi változó generátorfüggvénye G. Határozzuk meg $\xi + 1$ és 2ξ generátorfüggvényét!

14.4. Határozzuk meg két kockadobás összegének generátorfüggvényét!

14.5. Adjuk meg az ötös lottón kihúzott öt szám közül a legkisebbnek a generátorfüggvényét!

14.6. Adjuk meg két kockadobás szorzatának generátorfüggvényét, és ennek segítségével határozzuk meg a várható értéket és a szórást!

14.7. Jelölje G a ξ valószínűségi változó generátorfüggvényét. Írjuk fel G segítségével a következő számsorozatok generátorfüggvényét ($n = 0, 1, \dots$):

i) $P(\xi \leq n)$;
ii) $P(\xi < n)$;
iii) $P(\xi \geq n)$;
iv) $P(\xi > n + 1)$;
v) $P(\xi = 2n)$.

15. Nevezetes egyenlőtlenségek és alkalmazásaik

15.1. A Markov[19]- és a Csebisev[20]-egyenlőtlenség

A korábbiakban láttuk, hogy egy valószínűségi változó a saját várható értéke körül ingadozik, s az ingadozásának egyik lehetséges mértéke a szórás. A következőkben olyan egyenlőtlenségeket tárgyalunk, melyek még pontosabb felvilágosítást nyújtanak a valószínűségi változó saját várható értéke körüli ingadozásának mértékéről.

15.1.1. Tétel *(Markov-egyenlőtlenség) Ha ξ nem negatív valószínűségi változó, és létezik várható értéke, akkor bármely k pozitív szám esetén*

$$P(\xi \geq kM(\xi)) \leq \frac{1}{k}.$$

Bizonyítás. A diszkrét esetben

$$M(\xi) = \sum_{i=1}^{\infty} x_i p_i \geq \sum_{x_i \geq k}^{\infty} x_i p_i \geq k \sum_{x_i \geq k}^{\infty} p_i = kP(\xi \geq k),$$

a folytonos esetben pedig

$$M(\xi) = \int_0^{\infty} xf(x)dx \geq \int_k^{\infty} xf(x)dx \geq k \int_k^{\infty} f(x)dx = kP(\xi \geq k).$$

[19] Andrej Andrejevics Markov orosz matematikus, 1865-1922

[20] Pafnutyij Lvovics Csebisev orosz matematikus,1821-1894

15.1.2. Tétel *(Csebisev-egyenlőtlenség) Ha a ξ valószínűségi változónak létezik a várható értéke és szórása, továbbá k tetszőleges pozitív szám, akkor*

$$P(|\xi - M(\xi)| \geq k) \leq \frac{D^{(}\xi)}{k^2}.$$

Bizonyítás. Alkalmazzuk a Markov-egyenlőtlenséget ξ helyett a $(\xi - M(\xi))^2$ valószínűségi változóra és k helyett k^2-re. Ekkor

$$P(|\xi - M(\xi)| \geq k) = P((\xi - M(\xi))^2 \geq k^2) \leq$$
$$\leq \frac{M[(\xi - M(\xi))^2]}{k^2}.$$

A Csebisev-egyenlőtlenséget a következő alakban is kimondhatjuk:

15.1.3 Következmény *Ha a ξ valószínűségi változónak létezik várható értéke és szórása, továbbá k tetszőleges pozitív szám, akkor*

$$P(|\xi - M(\xi)| \geq kD(\xi)) \leq \frac{1}{k^2}.$$

A bizonyítás abból áll, hogy a 15.1.2. tételben a k helyett $kD(\xi)$-t írunk. A tétel állítása így fogalmazható: annak a valószínűsége, hogy a valószínűségi változó saját várható értékétől abszolút értékben a szórás k-szorosánál többel térjen el, legfeljebb $\frac{1}{k^2}$. A Csebisev-egyenlőtlenséget sokszor az alábbi formában alkalmazzuk:

$$P(M(\xi) - kD(\xi) < \xi < M(\xi) + kD(\xi)) = P(|\xi - M(\xi)| < kD(\xi)) =$$
$$= 1 - P(|\xi - M(\xi)| \geq kD(\xi)) \geq 1 - \frac{1}{k^2}.$$

Tekintsünk egy példát a fentiek alkalmazására. Legyen a ξ valószínűségi változó várható értéke 12, szórása 2. A fenti egyenlőtlenség alapján például válaszolhatunk arra a kérdésre, hogy milyen valószínűséggel esik ξ a $(9, 15)$ intervallumba. Ugyanis

$$P(9 < \xi < 15) = P(12 - 3 < \xi < 12 + 3) =$$
$$= P(12 - \frac{3}{2} \cdot 2 < \xi < 12 + \frac{3}{2} \cdot 2) \geq 1 - \frac{1}{(3/2)^2} \approx 0.55.$$

Ha például ξ binomiális eloszlású, akkor $np = 12$ és $npq = 4$ alapján $p = \frac{2}{3}$, $q = \frac{1}{3}$ és $n = 18$, így a keresett valószínűség

$$P(9 < \xi < 15) = \sum_{k=10}^{14} \binom{18}{k} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^{18-k} \approx 0.78.$$

15.2. A nagy számok törvénye

A következőkben a Csebisev-egyenlőtlenség alkalmazásaként egy alapvető fontosságú tételt bizonyítunk be. A tétel, mely a "nagy számok törvénye" nevet viseli, lényegében a következőt állítja: ha eléggé nagy számú független kísérletet végzünk egy A esemény bekövetkezése valószínűségének megállapítása céljából, akkor "gyakorlatilag biztos", hogy a megfigyelésekből adódó relatív gyakoriság az A esemény valószínűségét tetszőlegesen megközelíti. A "gyakorlatilag biztos" azt jelenti, hogy "1-hez tetszőlegesen közeli valószínűséggel". A tétel matematikailag pontosan megfogalmazva a következő:

15.2.1. Tétel *(A nagy számok törvénye) Legyenek $\xi_1, \xi_2, \ldots$ páronként független, azonos eloszlású valószínűségi változók, melyeknek létezik az M várható értékük és a szórásuk. Ekkor bármely $\epsilon > 0$ esetén*

$$\lim_{n\to\infty} P\Big(\Big|\frac{\xi_1+\xi_2+\cdots+\xi_n}{n} - M\Big| \geq \epsilon\Big) = 0.$$

Bizonyítás. Jelölje D a valószínűségi változók közös szórását. Legyen

$$\eta_n = \frac{\xi_1+\xi_2+\cdots+\xi_n}{n},$$

ekkor $M(\eta_n) = M$ és $D^2(\eta_n) = \frac{D^2}{n}$. Alkalmazzuk a Csebisev-egyenlőtlenséget a $k = \frac{\epsilon\sqrt{n}}{D}$ választással, ekkor kapjuk

$$P(|\eta_n - M| \geq \frac{\epsilon\sqrt{n}}{D} \cdot \frac{D}{\sqrt{n}}) \leq \frac{1}{(\frac{\epsilon\sqrt{n}}{D})^2} = \frac{D^2}{n\epsilon^2},$$

amiből az állítás következik.

A tétel speciális eseteként kapjuk Bernoulli következő tételét.

15.2.2. Tétel *(Bernoulli-tétel) Tekintsünk egy kísérletet, melynek egyik lehetséges kimenetele A. Jelentse η_n az A esemény bekövetkezéseinek számát a kísérlet n független végrehajtása során. Ekkor bármely $\epsilon > 0$ esetén*

$$\lim_{n\to\infty} P\Big(\Big|\frac{\eta_n}{n} - P(A)\Big| \geq \epsilon\Big) = 0.$$

Megjegyezzük, hogy a Bernoulli - tétel nem azt állítja, hogy a kísérletek számát növelve a relatív gyakoriság konvergál a valószínűséghez, hanem csak azt, hogy bármely előre adott számnál nagyobb eltérés egyre valószínűtlenebb. Ezen a ponton célszerű tisztázni a "konvergencia" pontos jelentését. Akkor mondjuk, hogy a $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ valószínűségi változók sorozata *sztochasztikusan konvergál* az η valószínűségi változóhoz, ha bármely $\epsilon > 0$ esetén

$$\lim_{n\to\infty} P(|\xi_n - \eta| \geq \epsilon) = 0.$$

Így a Bernoulli-tétel a következőképpen fogalmazható meg: független kísérletek sorozatában az A esemény relatív gyakorisága sztochasztikusan konvergál az A esemény valószínűségéhez.

A tételnek számos általánosítása és finomítása ismeretes. Ezek összefoglaló elnevezése: a "nagy számok törvényei". Egy ilyen általánosítás Hincsin nevéhez fűződik, aki megmutatta, hogy a 15.2.1. Tételben nem szükséges a valószínűségi változók szórásának létezését feltételezni.

15.3. A Moivre-Laplace-tétel

Ebben a fejezetben még egy nagy fontosságú tételt ismertetünk: a Moivre-Laplace-tételt. Ez a tétel az úgynevezett centrális határeloszlás-tételek egy speciális esete, melyek lényegében azt fejezik ki, hogy nagy számú független valószínűségi változó összege közelítőleg normális eloszlású, feltéve, hogy az összeg tagjai az egész összeghez képest nagy valószínűséggel kicsinyek. A Moivre-Laplace-tétel arról a speciális esetről szól, amikor a valószínűségi változók a 0 és 1 értékeket veszik fel, tehát összegük binomiális eloszlású.

15.3.1. Tétel *(Moivre-Laplace-tétel) Ha η_n n-edrendű $p > 0$ paraméterű binomiális eloszlású valószínűségi változó, akkor bármely valós x esetén*

$$\lim_{n\to\infty} P\left(\frac{\eta_n - np}{\sqrt{np(1-p)}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Tetszőleges ξ valószínűségi változó esetén, melynek létezik várható értéke és (zérustól különböző) szórása, a $\xi^* = \frac{\xi - M(\xi)}{D(\xi)}$ valószínűségi változónak nyilván 0 a várható értéke, és 1 a szórása. Ezt a valószínűségi változót szokás a ξ *standardizáltjának* nevezni. Például egy p valószínűségű A esemény ξ_A indikátor változójának standardizáltja

$$\xi_A^* = \frac{\xi - p}{\sqrt{pq}},$$

ahol $q = 1 - p$. Az $\mathcal{N}(m, \sigma)$ normális eloszlás standardizáltja az $\mathcal{N}(0, 1)$ eloszlás. A Moivre-Laplace-tételben szereplő $\frac{\eta_n - np}{\sqrt{np(1-p)}}$ valószínűségi változó az η standardizáltja. Ezért a tételt úgy is fogalmazhatjuk, hogy az n-edrendű binomiális eloszlás standardizáltjának eloszlásfüggvénye pontonként konvergál a standard normális eloszlás eloszlásfüggvényéhez, ha $n \to \infty$. Tehát, a standard binomiális eloszlás határeloszlása a standard normális eloszlás.

15.4. Feladatok

15.1. Legyen a ξ valószínűségi változó várható értéke 8, szórása 2. Legalább mennyi a valószínűsége annak, hogy a valószínűségi változó az $(5, 11)$ intervallumba esik?

15.2. A dohányzó lakosság napi átlagos cigarettafogyasztásának várható értéke 20 darab, szórása 6 darab. Legalább mennyi annak a valószínűsége, hogy a tényleges fogyasztás 11 és 29 darab közé essék? Legfeljebb mennyi annak a valószínűsége, hogy a tényleges fogyasztás legalább 30, vagy legfeljebb 10? Legalább mennyi annak a valószínűsége, hogy a tényleges fogyasztás 15 és 25 darab közé essék?

15.3. Egy automata élettartamának várható értéke 5 év, szórása 0.5 év. Legfeljebb mennyi annak a valószínűsége, hogy az automata élettartama legalább 1 évvel eltér a várható értéktől?

15.4. Egy raktárba naponta beérkező rendelések számának várható értéke 500, szórása 25. Legalább mekkora annak a valószínűsége, hogy egy napon a rendelések száma 400 és 600 közé esik?

16. A matematikai statisztika elemei

16.1. Statisztikai mező

A valószínűségelmélet alapjainak tárgyalásakor hangsúlyoztuk, hogy e tárgy véletlen tömegjelenségek tanulmányozására szolgál. A valószínűségelmélet gyakorlati alkalmazásai során azonban gyakori eset, hogy a modell nincs teljes egészében meghatározva, vannak bizonyos ismeretlen paraméterek, melyek nélkül az illető jelenség leírása csak korlátozottan, esetleg közelítőleg írható le. Leggyakrabban az az eset fordul elő, hogy a szóban forgó jelenséggel kapcsolatos valószínűségi változók eloszlása ismeretlen, csak annyit lehet tudni, hogy egy bizonyos valószínűségcsaládban keresendő. A matematikai statisztika az ilyen jellegű véletlen jelenségekkel foglalkozik, s bizonyos véletlen események bekövetkezése alapján igyekszik az ismeretlen valószínűségekre következtetni.

A matematikai statisztika tárgya úgynevezett statisztikai sokaságok vizsgálata. A statisztikai sokaság bizonyos egyedek, elemek meghatározott halmaza, mely egyedek bizonyos – általában számszerű – jellemzőkkel rendelkeznek. A statisztikában éppen az ilyen jellemzők vizsgálata az alapvető feladat.

Matematikailag a statisztika természetes közege a *statisztikai mező*. Statisztikai mezőnek nevezzük az $(\Omega, \mathcal{A}, \mathcal{P})$ rendezett hármast, ahol Ω tetszőleges, nem üres halmaz, $\mathcal{A}$ az Ω részhalmazainak σ-algebrája, $\mathcal{P}$ pedig az $\mathcal{A}$-n értelmezett valószínűségek egy családja. Általában a $\mathcal{P}$ valószínűségek egyparaméteres seregeként realizálódik, tehát $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta\}$.

Az Ω halmaz elemei az elemi eseményeket reprezentálják, $\mathcal{A}$ pedig a kísérlettel kapcsolatban megfigyelhető események halmaza. A $\mathcal{P}$ halmaz elemei a lehetséges valószínűségek. Vegyük észre, hogy bármely $\mathcal{P}$-beli P esetén $(\Omega, \mathcal{A}, P)$ valószínűségi mező, tehát a statisztikai mező úgy is felfogható, mint olyan valószínűségi mezők serege, melyekben az események azonosak.

Az $\mathbf{X} = (X_1, X_2, \ldots, X_n) : \Omega \to \mathbb{R}^n$ valószínűségi vektorváltozót *mintának*, komponenseit *mintaelemeknek*, az n számot pedig *mintanagyságnak* nevezzük. Ha az $\mathbf{X}$ komponensei függetlenek, illetve azonos eloszlásúak, akkor *független*, illetve *azonos eloszlású* mintáról szokás beszélni. A minta értékkészletét *mintatérnek* szokás nevezni.

A minta tehát egy valószínűségi vektorváltozó. A mintához gyakorlatilag mintavétellel lehet hozzájutni, ami azt jelenti, hogy egy valószínűségi változó értékeire megfigyeléseket végzünk, s az aktuális értékekből képzett vektor a minta. Mivel ezek az értékek maguk is a véletlentől függnek, ezért a minta maga is valószínűségi változó. Természetesen az alkalmazásokban általában feltételezik, hogy a minta független és azonos eloszlású, ami elég természetes követelmény, hiszen komponenseit ugyanarra a valószínűségi változóra vonatkozó megfigyelésekből kapjuk. A mintaelemek esetenként lehetnek függő valószínűségi változók is, például véges alapsokaságból való visszatevés nélküli mintavétel során.

Adott minta esetén a mintaelemek segítségével kívánunk következtetni a valószínűségi változó ismeretlen paramétereire, vagyis arra, hogy a $\mathcal{P}$ valószínűségcsaládnak melyik P_ϑ eleme írja le ténylegesen a jelenséget, tehát a statisztikai mezők seregéből melyik az aktuális valószínűségi mező. Ha $\mathbf{X} : \Omega \to \mathbb{R}^n$ egy minta és $T : \mathbb{R}^n \to \mathbb{R}^k$ egy olyan függvény, hogy a $T(\mathbf{X})$ összetett függvény valószínűségi változó, akkor T-t *statisztikának* nevezzük. A statisztika tehát egy, a mintatéren értelmezett vektor értékű függvény. Amennyiben értékei egydimenziós vektorok, akkor tehát egy, a mintatéren értelmezett, valós értékű függvényről van szó. Ekkor a statisztika a mintaelemekből "kiszámít" egy értéket, melyet arra kívánunk felhasználni, hogy a valószínűségi változó ismeretlen paraméterét ezzel közelítsük, becsüljük. Megjegyezzük, hogy gyakran nemcsak a T függvényt, hanem a $T(\mathbf{X})$ valószínűségi változót is statisztikának fogjuk nevezni, ami nem okozhat félreértést.

16.2. Gyakran használt statisztikák

A következőkben $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ egy rögzített mintát jelöl, és értelmezünk néhány gyakran használt statisztikát.

1. Legyen

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

tehát $\overline{X}$ a mintaelemek számtani közepe, átlaga. Ezt a statisztikát *mintaközépnek* vagy *empirikus várható értéknek* (*tapasztalati várható értéknek*) nevezzük.

2. Legyen
$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$
Az S^2 statisztikát *empirikus (tapasztalati) szórásnégyzetnek* nevezzük. Könnyű belátni, hogy ez a következő módon is kifejezhető:
$$S^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \overline{X}^2.$$

Az empirikus szórásnégyzet mintájára képezhetők az *empirikus k-adik momentumok.*

3. Legyen $(i_1, i_2, \ldots, i_n)$ az $(1, 2, \ldots, n)$ számok azon permutációja, melyre $X_{i_1} \leq X_{i_2} \leq \cdots \leq X_{i_n}$ teljesül, és legyen $\mathbf{X}^* = (X_{i_1}, X_{i_2}, \ldots, X_{i_n})$. Az $\mathbf{X}^*$ statisztikát *rendezett mintának* nevezzük. Ez tehát egy vektor értékű statisztika. Az $X_{i_n} - X_{i_1}$ statisztikát *mintaterjedelemnek* nevezzük.

Mivel a fentiekben értelmezett statisztikák maguk is valószínűségi változók, kiszámíthatjuk ezek várható értékét, szórását, amennyiben ezek léteznek. A továbbiakban tegyük fel, hogy a rögzített $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ minta független és azonos eloszlású, továbbá a mintaelemeknek létezik az m várható értéke és a σ szórása. Megjegyezzük, hogy m és σ valójában függ a ϑ paramétertől, hiszen kiszámításukhoz a $\mathcal{P}$ család bármelyik P_ϑ elemét felhasználhatjuk. Pillanatnyilag azonban ennek a függésnek nincs jelentősége: számításaink bármely P_ϑ esetén érvényesek.

Először határozzuk meg a mintaközép várható értékét:
$$M(\overline{X}) = M(\frac{1}{n}\sum_{i=1}^{n}X_i) = \frac{1}{n}\sum_{i=1}^{n}M(X_i) = \frac{1}{n}\sum_{i=1}^{n}m = m,$$
a várható érték tulajdonságai alapján. A mintaközép várható értéke tehát megegyezik az elméleti várható értékkel.

Következő példánk az empirikus várható érték szórása. Ezt a következő módon kapjuk:
$$D^2(\overline{X}) = D^2(\frac{1}{n}\sum_{i=1}^{n}X_i) = \frac{1}{n^2}\sum_{i=1}^{n}D^2(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n},$$
s ebből
$$D(\overline{X}) = \frac{\sigma}{\sqrt{n}}.$$

Például ha az alapsokaság $\mathcal{N}(m,\sigma)$ normális eloszlású, akkor a mintaközép $\mathcal{N}(m,\frac{\sigma}{\sqrt{n}})$ eloszlást követ. (A mintaközép normalitása abból adódik, hogy független, normális eloszlású változók összege is normális eloszlású.)

Most kiszámítjuk az empirikus szórásnégyzet várható értékét. Ehhez először átalakítjuk S^2 kifejezését a következő módon:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 - (\overline{X} - m)^2.$$

Mindkét oldal várható értékét véve kapjuk:

$$M(S^2) = \frac{1}{n}\sum_{i=1}^{n} M[(X_i - m)^2] - M[(\overline{X} - m)^2] = \frac{1}{n}\sum_{i=1}^{n}\sigma^2 - D^2(\overline{X}) = \frac{n-1}{n}\sigma^2.$$

Tehát az empirikus szórásnégyzet várható értéke nem egyezik meg az elméleti szórásnégyzettel. Ez a tény indokolja a *korrigált empirikus szórásnégyzet* bevezetését:

$$S^{*2} = \frac{n}{n-1}S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

melynek várható értéke már megegyezik az elméleti szórásnégyzettel.

A statisztikai vizsgálatok során gyakran használatos az *empirikus eloszlásfüggvény*, melyet a következőképpen értelmezünk:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\chi_{(X_i<x)} = \begin{cases} 0 & ha \quad x \le X_{i_1} \\ \frac{k}{n} & ha \quad X_{i_k} < x \le X_{i_{k+1}} \\ 1 & ha \quad X_{i_n} < x \end{cases}.$$

Világos tehát, hogy rögzített n és x mellett $F_n(x)$ valószínűségi változó. Az empirikus eloszlásfüggvény alapvető tulajdonságát fejezi ki a következő tétel, mely a nagy számok törvényének fontos következménye.

16.2.1. Tétel *Legyenek $X_1, X_2, \ldots$ független, azonos eloszlású valószínűségi változók F eloszlásfüggvénnyel. Az $(X_1, X_2, \ldots, X_n)$ minta empirikus eloszlásfüggvénye legyen F_n. Ekkor az $\{F_n\}$ függvénysorozat 1 valószínűséggel egyenletesen konvergál F-hez, azaz*

$$P(\lim_{n\to\infty}\sup_x |F_n(x) - F(x)| = 0) = 1.$$

A tételt nem bizonyítjuk.

16.3. Feladatok

16.1. Határozzuk meg az empirikus szórásnégyzet és a korrigált empirikus szórásnégyzet szórásnégyzetét!

16.2. Véletlen számtáblázat segítségével válasszunk egy 50 elemű mintát a $[0,1]$ intervallumban. Készítsük el a rendezett mintát, határozzuk meg a tapasztalati eloszlásfüggvényt! Határozzuk meg a mintaközepet és a korrigált tapasztalati szórásnégyzetet!

16.3. Egy üzlet dolgozóinak alapbéréről az alábbi adatokat ismerjük:

alapbér (Ft)	létszám (fő)
15500	2
17500	4
19000	5
22000	1
25500	4
30000	1

Számítsuk ki a mintaközepet és az empirikus szórásnégyzetet, valamint a korrigált empirikus szórásnégyzetet! Határozzuk meg az empirikus eloszlásfüggvényt!

16.4. Mutassuk meg, hogy ha $x_1, x_2, \dots, x_n$ és a valós számok, akkor

$$\frac{1}{n}\sum_{i=1}^{n}(a-x_i)^2 \geq \frac{1}{n}\sum_{i=1}^{n}(\overline{x}-x_i)^2,$$

ahol $\overline{x}$ jelöli az $x_1, x_2, \dots, x_n$ számtani közepét! Mutassuk meg továbbá, hogy a fenti egyenlőtlenségben egyenlőség akkor és csak akkor áll fenn, ha $a = \overline{x}$!

16.5. Tekintsük a következő mintát: $1250, 1340, 1120, 2100, 2450, 2000, 1400,$ $2100,$ 1100. Határozzuk meg a mintaközepet, az empirikus szórásnégyzetet, a korrigált empirikus szórásnégyzetet, a rendezett mintát, a mintaterjedelmet, és ábrázoljuk az empirikus eloszlásfüggvényt!

17. Statisztikai becslések

17.1. A statisztikai becslés fogalma

A statisztikai becslés célja a valószínűségi változó eloszlásában szereplő ismeretlen paraméterek közelítő meghatározása. Ez a feladat a valószínűségi változóhoz tartozó minta segítségével oldható meg. Az ismeretlen paraméterek értékére a mintából nyert információ segítségével következtethetünk. Ebből világos, hogy a becslés valójában egy statisztika, amelynek "jósága" definiálandó, s a definíció a gyakorlat követelményeinek kell, hogy megfeleljen.

Legyen tehát $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ egy minta valamilyen $\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta\}$ eloszlásból. *Becslésnek* nevezünk egy tetszőleges T statisztikát. A korábbi szóhasználat ezzel kapcsolatban úgy módosul, hogy a mintát szokás egy ξ *valószínűségi változóra vonatkozó megfigyelésnek* nevezni, ami alatt azt értjük, hogy a mintaelemek ξ-vel megegyező azonos eloszlású, független valószínűségi változók.

Természetesen a becsléssel kapcsolatban vannak bizonyos elvárások, amelyeket heurisztikusan a következő módon foglalhatunk össze. Először is nyilvánvaló követelmény, hogy a $T(\mathbf{X})$ értékeinek a becsülni kívánt ϑ paraméter valódi értéke körül kell ingadozniuk. Másodszor, ennek az ingadozásnak olyannak kell lennie, hogy a $T(\mathbf{X})$ értékei minél kevésbé szóródjanak. Végül, ha a $T(\mathbf{X}) = T_n(\mathbf{X})$ statisztika tetszőleges mintanagyság mellett értelmezett, akkor megkívánjuk, hogy a $T_n(\mathbf{X})$ értékei valamilyen értelemben konvergáljanak a becsülendő paraméter pontos értékéhez, ha $n \to \infty$. Valójában ez a követelmény tehát becslések egy egész sorozatára vonatkozik. Az alábbiakban megadjuk ezen tulajdonságok pontos matematikai definícióját.

17.2. Torzítatlan becslések

Azt mondjuk, hogy a $T(\mathbf{X})$ statisztika a ϑ paraméter *torzítatlan becslése*, ha bármely $\vartheta \in \Theta$ esetén fennáll az

$$M_\vartheta(T(\mathbf{X})) = \vartheta$$

egyenlőség. Itt és a továbbiakban az M_ϑ jelölés arra utal, hogy a $T(\mathbf{X})$ statisztika várható értékét a P_ϑ valószínűség segítségével számoljuk.

Ugyanígy értelmezhető egy ϑ paraméter valamely ψ függvényének torzítatlan becslése; ekkor azt követeljük meg, hogy bármely $\vartheta \in \Theta$ esetén teljesüljön az

$$M_\vartheta(T(\mathbf{X})) = \psi(\vartheta)$$

egyenlőség.

A korábbiakban láttuk, hogy a mintaközép a várható érték torzítatlan becslése. Ebből adódik, hogy λ paraméterű Poisson-eloszlásból származó minta esetén a mintaközép a λ paraméter torzítatlan becslése, hiszen ekkor λ éppen a várható érték. Hasonlóan, λ paraméterű exponenciális eloszlás esetén a mintaközép az $\frac{1}{\lambda}$ torzítatlan becslése. Megjegyezzük, hogy míg egy esemény p valószínűségének a relatív gyakoriság torzítatlan becslése, az $\frac{1}{p}$-re általában nem létezik torzítatlan becslés. Nem gondolhatjuk tehát azt, hogy ha a $T(\mathbf{X})$ a ϑ paraméter torzítatlan becslése, akkor $\psi(T(\mathbf{X}))$ a $\psi(\vartheta)$ torzítatlan becslése lenne.

Ugyancsak láttuk a korábbiakban, hogy az empirikus szórásnégyzet a szórásnégyzetnek nem torzítatlan becslése, de torzítatlan becslését adja a szórásnégyzetnek a korrigált empirikus szórásnégyzet.

17.3. Hatásos becslések

A becslések jóságára vonatkozó második fontos követelmény az, hogy a becslésként használt statisztika értékeinek szóródása minél kisebb legyen. Legyenek a $T_1(\mathbf{X})$ és $T_2(\mathbf{X})$ statisztikák a ϑ paraméter torzítatlan becslései. Akkor mondjuk, hogy a T_1 statisztika *hatásosabb* becslés, mint a T_2 statisztika, ha minden $\vartheta \in \Theta$ esetén teljesül, hogy

$$D^2_\vartheta(T_1) \le D^2_\vartheta(T_2).$$

Ha T a ϑ paraméter olyan torzítatlan becslése, mely minden más torzítatlan becslésnél hatásosabb, akkor azt mondjuk, hogy T a ϑ *hatásos becslése*. Ugyanígy értelmezhető a paraméter valamely függvényének hatásos becslése.

Példaként tekintsünk egy véges várható értékkel és szórással rendelkező valószínűségi változót, melyre vonatkozó minta $(X_1, X_2, \ldots, X_n)$. Könnyű belátni, hogy ha $c_1 + c_2 + \cdots + c_n = 1$ és $c_i \geq 0$ $(i = 1, 2, \ldots, n)$, akkor a $\sum_{i=1}^{n} c_i X_i$ statisztika a várható értéknek torzítatlan becslése. Meg lehet mutatni, hogy az ilyen alakú becslések mindegyikénél hatásosabb a mintaközép, amely a $c_i = \frac{1}{n}$ $(i = 1, 2, \ldots, n)$ választásnak felel meg.

Bár egy adott paraméterre vonatkozó hatásos becslés létezésére általában nincs garancia, az azonban kimutatható, hogy két hatásos becslés szükségképpen 1 valószínűséggel egyenlő.

17.4. Konzisztens becslések

Végül a harmadik, becslésekkel szemben támasztható fontos követelményt fejezi ki a következő fogalom. Tegyük fel, hogy $X_1, X_2, \ldots$ egy valószínűségi változóra vonatkozó megfigyelések sorozata és $T_n = T_n(X_1, X_2, \ldots, X_n)$ a ϑ paraméter becsléseinek egy sorozata. Akkor mondjuk, hogy ez a becsléssorozat a ϑ-ra nézve *konzisztens*, ha bármely $\epsilon > 0$ és bármely $\vartheta \in \Theta$ esetén fennáll

$$\lim_{n\to\infty} P_\vartheta(|T_n - \vartheta| \geq \epsilon) = 0.$$

Ez tehát azt jelenti, hogy a T_n statisztikák sorozata sztochasztikusan konvergál ϑ-hoz minden ϑ esetén. (Itt természetesen a "sztochasztikusan konvergál" kifejezés magában foglalja a ϑ-tól való függést.) Meg lehet mutatni, hogy véges várható értékkel és szórással rendelkező valószínűségi változó esetén a mintaközép a várható értéknek konzisztens becslése. Ez valójában a nagy számok törvényének egy másik megfogalmazása.

Megjegyezzük, hogy ha a szórásnégyzetek sorozata $n \to \infty$ esetén zérushoz tart, akkor torzítatlan becslések sorozata a Csebisev-egyenlőtlenség alapján mindig konzisztens.

Legyenek $X_1, X_2, \ldots$ $\mathcal{N}(m, 1)$ eloszlású valószínűségi változóra vonatkozó megfigyelések. Ha $T_n = X_n$, akkor T_n az m paraméter torzítatlan becslése. Másrészt, tetszőleges $\epsilon > 0$ esetén

$$P_m(|T_n - m| \geq \epsilon) = P_m(|X_n - m| \geq \epsilon) = P_m(m - \epsilon \leq X_n \leq m + \epsilon) = \Phi(\epsilon) - \Phi(-\epsilon),$$

ahol Φ a standard normális eloszlás eloszlásfüggvénye. A fenti kifejezésről könnyű látni, hogy nem tart zérushoz, ha $n \to \infty$, így ez a becsléssorozat nem konzisztens az m paraméterre nézve.

17.5. A maximumlikelihood-becslés

Egy statisztikai paramétert egy adott mintából számos különböző módon lehet becsülni, még akkor is, ha a fentiekben ismertetett tulajdonságok teljesülését is megkívánjuk az alkalmazott becsléstől. A gyakorlat számára nagy jelentőségűek az olyan becslési módszerek, melyeknél a becslést direkt módon tudjuk megkonstruálni a minta alapján. Egy ilyen, gyakran alkalmazott becslési eljárás a *maximumlikelihood*-becslés, mely a maximális valószínűség elvén alapul. Ez azt mondja, hogy a mintaelemek felhasználásával valamely ϑ lehetséges paraméterérték mellett kiszámítható a minta előfordulásának valószínűsége, s a paraméter becslésének azt az értéket tekintjük, amely mellett ez a valószínűség a lehető legnagyobb, maximális. A számítások alapjául az úgynevezett *likelihood függvény* szolgál, melyet a diszkrét, illetve az abszolút folytonos esetben a következőképpen adunk meg:

$$L(\mathbf{x}, \vartheta) = \begin{cases} \prod_{i=1}^{n} f_\vartheta(x_i) & \text{abszolút folytonos eset} \\ \prod_{i=1}^{n} p_\vartheta(x_i) & \text{diszkrét eset,} \end{cases}$$

ahol f_ϑ a (közös) sűrűségfüggvény, illetve $p_\vartheta(x_i) = P_\vartheta(X_i = x_i)$, az úgynevezett *valószínűségfüggvény*. Rögzített minta esetén a likelihood függvény csupán ϑ függvénye, így maximumát analitikus módszerrel lehet meghatározni. Alkalmas feltételek mellett a differenciálszámítás apparátusa használható. Felhasználva a logaritmusfüggvény szigorú monotonitását, valamint a likelihood függvény szorzatként való előállítását, technikai okokból gyakran célszerűbb L helyett $\ln L$ maximumhelyeit meghatározni. Ezeket a maximumhelyeket nevezzük a paraméter *maximumlikelihood-becsléseinek*. Meg lehet mutatni, hogy ha létezik a paraméternek hatásos becslése, akkor ez éppen a maximumlikelihood-becslés.

Bár itt csak egyetlen paraméter becsléséről beszéltünk, természetesen semmi akadálya egyidejűleg több paraméter maximumlikelihood-becslésének, melyhez a többváltozós függvények szélsőértékének számítási módszerei használhatók.

Példaként tekintsük $\mathcal{N}(m, \sigma)$ normális eloszlás esetén az m és σ paraméterek maximumlikelihood-becslését. A likelihood függvény:

$$L(x_1, x_2, \ldots, x_n, m, \sigma) = (2\pi)^{-\frac{n}{2}} \sigma^{-n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m)^2}.$$

Ekkor a

$$\frac{\partial \ln L}{\partial m} = 0, \qquad \frac{\partial \ln L}{\partial \sigma^2} = 0$$

egyenletrendszerből kapjuk a stacionárius helyeket:

$$\hat{m} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

és

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Könnyen belátható, hogy ezek a stacionárius helyek valóban maximumhelyek. Ez azt jelenti, hogy a normális eloszlás esetén a várható érték, illetve a szórásnégyzet maximumlikelihood-becslése a mintaközép, illetve az empirikus szórásnégyzet.

17.6. Konfidenciaintervallumok

Az eddigiekben olyan becslésekkel foglalkoztunk, amikor a valószínűségi változó, illetve az eloszlás valamely ismeretlen paraméterét egyetlen, a mintaelemekből valamilyen módon kiszámított értékkel közelítettük, becsültük. Ezek az úgynevezett *pontbecslések*. A becslések egy másik típusát képezik az olyanok, amikor az ismeretlen paramétert nem egy számértékkel kívánjuk közelíteni, hanem olyan intervallumot keresünk, amely előre megadott, 1-hez közeli valószínűséggel tartalmazza az illető paramétert. Ezek az úgynevezett *intervallumbecslések*. Ha az (u_1, u_2) véletlen helyzetű intervallum $1-\epsilon$ valószínűséggel tartalmazza a ϑ paramétert, akkor azt mondjuk, hogy $1-\epsilon$ valószínűséggel *lefedi* a ϑ paramétert, illetve (u_1, u_2) *konfidenciaintervallum* a ϑ paraméterre $100(1-\epsilon)\%$ *megbízhatósági szinten.* Az intervallum végpontjait *alsó*, illetve *felső konfidenciahatároknak* nevezzük. Az, hogy (u_1, u_2) véletlen helyzetű intervallum azt jelenti: végpontjai maguk is valószínűségi változók. Az $1-\epsilon$ megbízhatóság tehát azt jelenti, hogy elég nagy minta esetén az intervallum átlagosan az esetek $100(1-\epsilon)\%$-ában tartalmazza a paraméter valódi értékét. Tévedésünk mértékét $100\epsilon\%$ fejezi ki.

Konfidenciaintervallumok konstrukciójára különféle módszerek ismeretesek. A következő példában a normális eloszlás várható értékére konstruálunk konfidencia-intervallumot ismert szórás esetén.

Legyen adott egy n elemű minta az $\mathcal{N}(m, \sigma)$ eloszlásból, ahol σ ismert, de az m értéke ismeretlen. Tudjuk, hogy az $\overline{X}$ mintaközép torzítatlan becslése m-nek, és szórása $\frac{\sigma}{\sqrt{n}}$, így $\overline{X}$ egy $\mathcal{N}(m, \frac{\sigma}{\sqrt{n}})$ eloszlású valószínűségi változó. Ezért az

$$u = \frac{\overline{X} - m}{\sigma/\sqrt{n}}$$

valószínűségi változó standard normális eloszlást követ. Adott $0 < \epsilon < 1$ esetén az $\mathcal{N}(0,1)$ eloszlás táblázatából meghatározható az az u_ϵ szám, amelyre

$$2\Phi(u_\epsilon) - 1 = \Phi(u_\epsilon) - \Phi(-u_\epsilon) = P(-u_\epsilon < u < u_\epsilon) = 1 - \epsilon.$$

Ekkor tehát

$$\Phi(u_\epsilon) = 1 - \frac{\epsilon}{2},$$

s ezzel a választással az $(\overline{X} - u_\epsilon \frac{\sigma}{\sqrt{n}}, \overline{X} + u_\epsilon \frac{\sigma}{\sqrt{n}})$ intervallum $1-\epsilon$ valószínűséggel lefedi az eloszlás várható értékét, tehát m-re $100(1-\epsilon)\%$-os megbízhatósági szintű konfidenciaintervallumot kaptunk. A mintaközépnek az m paramétertől való eltérése $100(1-\epsilon)\%$-os szinten az intervallum félhossza, $u_\epsilon \frac{\sigma}{\sqrt{n}}$.

Előfordul, hogy az eltérést kívánjuk korlátozni, s az a kérdés, hogy ehhez (legalább) mekkora mintanagyság szükséges. Ha legfeljebb d a megengedett eltérés, akkor az eddigiek szerint

$$u_\epsilon \frac{\sigma}{\sqrt{n}} \le d,$$

amiből a szükséges mintanagyság

$$n \ge \frac{u_\epsilon^2 \sigma^2}{d^2}.$$

A következőkben a t-eloszlás és a χ^2-eloszlás alkalmazására mutatunk egy-egy példát.

Tegyük fel, hogy ismeretlen várható értékű és szórású normális eloszlású sokaságból vett minta alapján kívánunk a várható érték mintaközéppel való becslésére adott szintű konfidenciaintervallumot megadni. Ha a szórást a fentiekben szerepelt u definíciójában az S^* korrigált empirikus szórással helyettesítjük, akkor a

$$t = \sqrt{n}\frac{\overline{X} - m}{S^*}$$

valószínűségi változót kapjuk, melyről ki lehet mutatni, hogy $n-1$ szabadsági fokú t-eloszlást követ. A t-eloszlás táblázatából adott $0 < \epsilon < 1$ esetén meghatározható az a t_ϵ szám, amelyre

$$P(-t_\epsilon < t < t_\epsilon) = 1 - \epsilon,$$

azaz, t értéke $1-\epsilon$ valószínűséggel esik a $(-t_\epsilon, t_\epsilon)$ intervallumba. Ezért a t definíciója alapján az

$$(\overline{X} - t_\epsilon \frac{S^*}{\sqrt{n}}, \overline{X} + t_\epsilon \frac{S^*}{\sqrt{n}})$$

intervallum $100(1-\epsilon)$%-os szintű konfidenciaintervallum a várható értékre.

Tekintsük a következő példát. Valamely termék minőségi jellemzője ismeretlen várható értékű és szórású normális eloszlást követ. Miután 19 darabot megvizsgáltak, a mintaközépre 2200, a korrigált tapasztalati szórásra pedig 190 adódott. Adjunk konfidenciaintervallumot a várható értékre 99%-os megbízhatósági szinten! A fentiek alapján az intervallum határainak kiszámitásához a

$$t_\epsilon \frac{S^*}{\sqrt{n}}$$

kifejezés értékét kell ismernünk. A szabadsági fokok száma $19 - 1 = 18$, így a 99%-os szintnek megfelelő érték:

$$t_\epsilon = 2.878,$$

amiből a konfidenciaintervallum határai a fenti formula révén megkaphatók, s az intervallum: $(2074.55, 2325.45)$.

A következőkben normális eloszlású sokaság szórásának becslésével foglalkozunk. Legyen adott egy n elemű minta az $\mathcal{N}(m, \sigma)$ eloszlású sokaságból, melynek szórása ismeretlen. Meg lehet mutatni, hogy az $n\frac{S^2}{\sigma^2}$ valószínűségi változó $n-1$ szabadsági fokú χ^2-eloszlású, ahol S^2 a minta empirikus szórásnégyzete. Olyan intervallumot, amelybe ez a változó $1-\epsilon$ valószínűséggel esik, végtelen sokféleképpen lehet választani. Egy lehetséges választás a következő: kimutatható, hogy ha $0 < \epsilon < 1$, akkor

$$P(\chi^2_{1-\frac{\epsilon}{2}} < n\frac{S^2}{\sigma^2} < \chi^2_{\frac{\epsilon}{2}}) = 1 - \epsilon,$$

amiből következik, hogy az

$$\Big(\frac{nS^2}{\chi^2_{\frac{\epsilon}{2}}}, \frac{nS^2}{\chi^2_{\frac{1-\epsilon}{2}}}\Big)$$

intervallum $100(1-\epsilon)$%-os szintű konfidenciaintervallum az ismeretlen σ^2 paraméterre. Ugyanilyen szintű konfidenciaintervallum a szórásra a

$$\Big(\frac{\sqrt{n}S}{\chi_{\frac{\epsilon}{2}}}, \frac{\sqrt{n}S}{\chi_{\frac{1-\epsilon}{2}}}\Big)$$

intervallum.

Adott n és ϵ esetén a $\chi^2_{\frac{\epsilon}{2}}$ és $\chi^2_{\frac{1-\epsilon}{2}}$ értékeket a χ^2 eloszlás táblázatából határozhatjuk meg. Tekintsük a következő példát: egy automata által gyártott alkatrész valamely jellemzőjét 10 esetben megmérve a következő értékek adódtak:

2.18	2.12	2.14	2.15	2.21
2.17	2.14	2.13	2.10	2.14

Ennek alapján az empirikus szórás $S = 0.032$. A σ-ra vonatkozó 95%-os konfidenciaintervallum meghatározása a következőképpen történhet: $\epsilon = 0.05$, a szabadsági fokok száma $10 - 1 = 9$, s a χ^2-eloszlás táblázatából a megfelelő értékek:

$$\chi^2_{\frac{\epsilon}{2}} = 19.023,$$

$$\chi^2_{\frac{1-\epsilon}{2}} = 2.7.$$

Ezek felhasználásával a fenti formulákból a keresett konfidenciaintervallum a következő: $(0.023, 0.062)$.

17.7. Feladatok

17.1. Legyen $X_1, X_2, \ldots, X_n$ a $(0, c)$-ben egyenletes eloszlásból származó minta, X_n^* a maximum és $\overline{X}$ a mintaközép. Mutassuk meg, hogy $\frac{n+1}{n} X_n^*$ torzítatlan becslése c-nek, és hatásosabb, mint $2\overline{X}$!

17.2. Mutassuk meg, hogy ha $(X_k - m)^4$ várható értéke véges $(k = 1, 2, \ldots)$, akkor a korrigált empirikus szórásnégyzet konzisztens becslése a szórásnégyzetnek!

17.3. Tekintsünk egy véges várható értékkel és szórással rendelkező valószínűségi változót, melyre vonatkozó minta $(X_1, X_2, \ldots, X_n)$. Mutassuk meg, hogy ha $c_1 + c_2 + \cdots + c_n = 1$ és $c_i \geq 0$ $(i = 1, 2, \ldots, n)$, akkor a $\sum_{i=1}^{n} c_i X_i$ statisztika a várható értéknek torzítatlan becslése! Mutassuk meg, hogy az ilyen alakú becslések mindegyikénél hatásosabb a mintaközép!

17.4. Az (a, b) intervallumban egyenletes eloszláshoz tartozó minta alapján becsüljük az ismeretlen a és b paraméterek $c = \frac{a+b}{2}$ számtani közepét. Mutassuk meg, hogy bár a mintaközép torzítatlan becslés c-re, de van nála hatásosabb!

17.5. Rendelkezésre áll egy binomiális eloszlásból származó $n = 20$ elemű minta, melyben $k = 13$ bizonyos tulajdonságú elem található. Határozzuk meg a minta alapján az eloszlás p paraméterének maximumlikelihood-becslését!

17.6. Az előző feladat általánosításaként határozzuk meg a binomiális eloszlás p paraméterének maximumlikelihood-becslését!

17.7. Határozzuk meg Poisson-eloszlás esetén a λ paraméter maximumlikelihood-becslését!

17.8. Legyen $(x_1, x_2, \ldots, x_n)$ a $(0, a)$ intervallumban egyenletes eloszlásból vett minta. Mutassuk meg, hogy az a paraméter maximumlikelihood-becslése $\max(x_1, x_2, \ldots, x_n)$, és határozzuk meg e becslés *torzítását*, azaz e becslés várható értékének és a-nak különbségét!

17.9. Valamely mérőeszközzel történő mérési eredmények $\mathcal{N}(m, 0.2)$ normális eloszlást követnek. Miután 25 mérést elvégeztünk, a mérési eredmények átlaga 2.25. Adjunk 95%-os szintű konfidenciaintervallumot az m várható értékre!

17.10. Egy gép előírt hosszúságú darabokat vág le egy acéllemezből, de a tényleges hosszúság ingadozást mutat $\mathcal{N}(m, 2.2)$ normális eloszlás szerint. Adjunk 95%-os szintű konfidenciaintervallumot m-re, ha egy 40 darabból álló minta átlagos hosszúsága 80.5 cm!

17.11. Villanyégők vizsgálatánál egy adott tételből 15 darabnak mérték meg az égési időtartamát, mely közelítőleg normális eloszlású volt. A mintaközépre 1200 óra, a korrigált empirikus szórásnégyzetre 186 óra adódott. Adjunk 99%-os szintű konfidenciaintervallumot a várható értékre!

17.12. Egy markológép által kiemelt homok mennyiségét véletlenszerűen 16-szor lemérték, melynek empirikus szórására 15 kg értéket kaptak. Adjunk meg a szórásra 90%-os szintű konfidenciaintervallumot!

18. Hipotézisvizsgálat, statisztikai próbák

18.1. Statisztikai hipotézisvizsgálat

Általában egy adott statisztikai sokasággal kapcsolatban különböző feltevéseket, *hipotéziseket* fogalmazunk meg. Ezek a hipotézisek a sokaság különböző jellemzőire vonatkozhatnak. A gyakorlati alkalmazások számára nagyon fontosak azok a statisztikai módszerek, melyek segítségével eldönthető, hogy egy adott hipotézis elfogadható-e, illetve elfogadása esetén mekkora az esetleges tévedés valószínűsége. Ilyen kérdésekkel foglalkozik a statisztikai döntések elmélete és a statisztikai hipotézisvizsgálat. A statisztikai döntések meghozatalát általában úgynevezett *statisztikai próbák* előzik meg; ezek alapján fogadunk el vagy utasítunk el egy adott hipotézist. A következőkben összefoglaljuk a hipotézisvizsgálat és a döntéshozatal legfontosabb összetevőit, ezek néhány gyakran alkalmazott módszerét.

Adott statisztikai mező esetén tekintsük a Θ paraméterteret, és bontsuk fel két diszjunkt részhalmazra: $\Theta = \Theta_0 \cup \Theta_1$. Azt a feltevést, hogy valamely ismeretlen ϑ paraméter a Θ_0-hoz tartozik, *nullhipotézisnek* nevezzük, azt pedig, hogy ϑ a Θ_1-nek eleme, *ellenhipotézisnek* vagy *alternatív hipotézisnek*. Ezeket a következő módon jelöljük:

$$H_0 : \vartheta \in \Theta_0$$

$$H_1 : \vartheta \in \Theta_1.$$

18.2. Statisztikai próbák

A hipotézisek elfogadásáról vagy elutasításáról statisztikai próbák segítségével döntünk. Tekintsünk egy T statisztikát, és adjunk meg egy $0 < \epsilon < 1$ számot. A mintatérben adjunk meg egy E *elfogadási tartományt*, melyre teljesül, hogy

$$P(T(X_1, X_2, \ldots, X_n) \in E|H_0) \geq 1 - \epsilon,$$

és egy (E-től diszjunkt) K *kritikus tartományt*, melyre az teljesül, hogy

$$P(T(X_1, X_2, \ldots, X_n) \in K|H_0) \leq \epsilon.$$

Tehát a mintából készített statisztika értékei legalább $1 - \epsilon$ valószínűséggel esnek az elfogadási tartományba, és legfeljebb ϵ valószínűséggel esnek a kritikus tartományba, feltéve, hogy a nullhipotézis fennáll. Ezek után döntésünk természetesen az lesz, hogy ha a statisztika értéke az elfogadási tartományba esik, akkor elfogadjuk a nullhipotézist (vagyis nem utasítjuk el), ha pedig az érték a kritikus tartományba esik, akkor elutasítjuk, és az ellenhipotézist fogadjuk el. Természetesen a döntés lehet helyes is, és hibás is, s célszerű az esetleg elkövetett hiba valószínűségét megbecsülni. Könnyű belátni, hogy kétféle hibát követhetünk el. *Elsőfajú hibát* akkor követünk el, ha elutasítjuk a nullhipotézist, holott az igaz. Ez akkor történik, amikor a statisztika a kritikus tartományba esik, és H_0 igaz, tehát ennek valószínűsége legfeljebb ϵ, amit a próba *szignifikancia szintjének* nevezünk.

Másodfajú hibát akkor követünk el, ha elfogadjuk a hibás H_0 hipotézist. Ez akkor következik be, ha a statisztika az elfogadási tartományba esik, miközben az ellenhipotézis teljesül. Ennek valószínűsége

$$P(T(X_1, X_2, \ldots, X_n) \in E|H_1) = 1 - P(T(X_1, X_2, \ldots, X_n) \in K|H_1).$$

Ezek alapján a hibás döntés valószínűségét az első- és másodfajú hiba valószínűségeinek összege adja. Általánosságban elmondhatjuk, hogy bármelyik típusú hiba valószínűségének csökkentése adott mintanagyság esetén csak a másik típusú hiba rovására történhet. A mintanagyság növelésével azonban esetenként elérhető mindkét típusú hiba valószínűségének, tehát a hibás döntés meghozatala valószínűségének csökkentése.

A következőkben néhány olyan eljárást, *statisztikai próbát* ismertetünk, melyek a gyakorlatban alkamazhatók statisztikai döntések meghozatalakor. A 18.3. és 18.4. pontokban un. *paraméteres próbát* ismertetünk, ugyanis feladatunk valamilyen eloszláscsaládból vett eloszlás ismeretlen paraméterére tett hipotézis vizsgálata. A 18.5. pontban úgynevezett *nem paraméteres* próba szerepel, amelyben egy valószínűségi változó eloszlására tett hipotézisünket vizsgáljuk.

18.3. Az u-próba

Adott egy $(X_1, X_2, \ldots, X_n)$ független minta az ismert σ szórású és ismeretlen m várható értékű $\mathcal{N}(m, \sigma)$ normális eloszlású sokaságból. Feladatunk annak eldöntése, hogy az ismeretlen várható érték megegyezik-e valamely adott m_0 valós számmal. Ennek megfelelően a nullhipotézis és az ellenhipotézis a következőképpen írható fel:

$$H_0 : m = m_0,$$

$$H_1 : m \neq m_0.$$

A döntés meghozatalához a következő próbastatisztikát használjuk:

$$u = \frac{\overline{X} - m_0}{\sigma}\sqrt{n}.$$

Ha helyes a nullhipotézis, akkor az u valószínűségi változó $\mathcal{N}(0, 1)$ standard normális eloszlású. Adott $0 < \epsilon < 1$ számhoz a standard normális eloszlás táblázatából meghatározzuk azt az u_ϵ értéket, amelyre

$$P(|u| > u_\epsilon) = \epsilon.$$

A gyakorlatban általában az $\epsilon = 0.05$, $\epsilon = 0.01$, $\epsilon = 0.001$ értékek használatosak, ezek mellett ugyanis gyakorlatilag biztosnak tekinthető az $(|u| \leq u_\epsilon)$ esemény.

Ezek után, ha a tényleges mintából kiszámított u-ra $|u| > u_\epsilon$ teljesül, tehát a statisztika a kritikus tartományba esik, akkor egy olyan esemény következett be, amelyet gyakorlatilag lehetetlennek minősítettünk, így a kiinduló nullhipotézis helytelen voltára következtetünk, s azt elvetjük. Ekkor azt mondjuk, hogy a tényleges várható érték és a feltételezett m_0 közötti eltérés nem véletlen, hanem *szignifikáns* a $100(1 - \epsilon)$%-os *szignifikanciaszinten*. Ellenkező esetben, tehát ha $|u| < u_\epsilon$, akkor a tényleges várható érték és az m_0 közötti eltérés statisztikailag nem igazolható, így a nullhipotézis elutasítására nincs okunk, azt elfogadjuk. Vegyük észre, hogy az egész eljárás döntő mértékben függ az előre adott ϵ értéktől, hiszen ez határozza meg a szignifikanciaszintet. Ha ez egy konkrét gyakorlati problémával kapcsolatban nincs megadva, akkor általában az $\epsilon = 0.05$ értéket használhatjuk. Mindazonáltal a döntés meghozatalakor a szignifikancia szintjét mindig meg kell említenünk.

Tekintsük a következő problémát: egy csokoládégyár 14 dekagrammos csokoládészeleteket gyárt egy 2 dekagramm szórású gépen. Egy véletlen minta mintaközepe 14.8 dekagramm volt. Feltételezhető-e az eltérés véletlenszerűsége? Esetünkben tehát

$$H_0 : m = m_0 = 14,$$

$$H_1 : m \neq m_0.$$

Az u-próba alkalmazásához válasszuk az általában szokásos $\epsilon = 0.05$ értéket. Ekkor

$$0.05 = P(|u| > u_\epsilon) = 1 - P(|u| \leq u_\epsilon) = 1 - P(-u_\epsilon \leq u \leq u_\epsilon) =$$

$$= 1 - P(u \leq u_\epsilon) + P(u \leq -u_\epsilon) = 2 - 2P(u \leq u_\epsilon) = 2 - 2\Phi(u_\epsilon),$$

tehát

$$\Phi(u_\epsilon) = 1 - \frac{\epsilon}{2} = 0.975.$$

A standard normális eloszlás táblázatából $u_\epsilon = 1.96$ adódik, míg az adatokból $u = 2$, tehát a nullhipotézist 95%-os szignifikanciaszinten el kell vetnünk, mert az eltérés mértéke ezen a szinten szignifikáns.

18.4. A t-próba

A gyakorlati problémák többségében a normális eloszlású sokaságnak a várható értéke és a szórása is ismeretlen, csupán annyit tudunk, hogy $\mathcal{N}(m, \sigma)$ normális eloszlású. Feladatunk ismét annak eldöntése, hogy az ismeretlen várható érték megegyezik-e valamely adott m_0 valós számmal. Ennek megfelelően a nullhipotézis és az ellenhipotézis ismét a következőképpen írható fel:

$$H_0 : m = m_0,$$

$$H_1 : m \neq m_0.$$

A nullhipotézis ellenőrzésére ezúttal a következő próbastatisztikát használjuk:

$$t = \frac{\overline{X} - m_0}{S^*}\sqrt{n}.$$

Ha helyes a nullhipotézis, akkor a t valószínűségi változó $n-1$ szabadsági fokú t-eloszlású. Adott $0 < \epsilon < 1$ számhoz a t-eloszlás táblázatából meghatározzuk azt a t_ϵ értéket, amelyre

$$P(|t| > t_\epsilon) = \epsilon.$$

Az u-próbához hasonlóan a gyakorlatban általában az $\epsilon = 0.05$, $\epsilon = 0.01$, $\epsilon = 0.001$ értékek használatosak. A továbbiakban az u-próbánál alkalmazott eljárást követjük.

Tekintsük most a következő problémát: egy töltőautomata 1000 gramm anyag betöltésére van beállítva. Mintavétel során a következő értékeket kaptuk:

985	987	1003	993	996
991	994	1004	1002	985

A következő hipotézist kell megvizsgálnunk 95%-os szinten: teljesül-e a várható értékre az $m_0 = 1000$ előírás. Tehát

$$H_0 : m = m_0 = 1000,$$

$$H_1 : m \neq m_0.$$

Ha feltételezzük, hogy a minta normális eloszlásból van, akkor t-próbát alkalmazunk a nullhipotézis ellenőrzésére. Az adatok alapján $\overline{X} = 9940$ és ${S^*}^2 = 52.222$, amiből $t = -2.626$ érték adódik. Mivel 10 adatunk van, a 9 szabadsági fokú t-eloszlás táblázatából az $\epsilon = 0.05$ értékhez $t_\epsilon = 2.267$ tartozik, így 95%-os szinten szignifikáns eltérés van, aminek következtében ezen a szinten a nullhipotézist elutasítjuk.

Megjegyezzük, hogy 99%-os szinten a nullhipotézis elfogadhatónak bizonyul. Más kérdés, hogy a mintanagyság túlságosan kicsi ahhoz, hogy ennek alapján 99%-os szinten ”nyugodtan” elfogadhassuk a nullhipotézist.

18.5. A χ^2-próba

Egy valószínűségi változó eloszlására tett hipotézisünket úgy ellenőrizzük, hogy lehetséges értékeit diszjunkt csoportokba osztjuk. Jelölje r a csoportok számát, p_i annak a valószínűségét (a feltételezett eloszlásban), hogy a valószínűségi változó értéke az i-edik csoportba esik, és ν_i az i-edik csoportba eső mintaelemek számát. Legyen

$$\chi^2 = \sum_{i=1}^{r} \frac{(\nu_i - np_i)^2}{np_i}.$$

Ekkor meg lehet mutatni (a Moivre-Laplace-tétel alapján), hogy ez a változó közelítőleg χ^2-eloszlású, $r-1$ szabadsági fokkal. Ennek alapján a próbát a következőképpen kell lefolytatni. Adott $0 < \epsilon < 1$ esetén meghatározzuk azt a χ^2_ϵ számot, amelyre $P(\chi^2 > \chi^2_\epsilon) = \epsilon$. Ezután, ha a minta alapján kapott χ^2 érték nagyobb χ^2_ϵ-nél, akkor a nullhipotézist elvetjük, ellenkező esetben azt mondjuk, hogy nincs ellentmondás a minta és a nullhipotézis között.

Most tekintsük a következő példát. Egy dobókockát 1400-szor feldobunk, és az egyes számok gyakoriságára a következő értékeket kapjuk:

Dobott szám	Gyakoriság
1-es	228
2-es	240
3-as	224
4-es	237
5-ös	235
6-os	236

Szabályosnak tekinthetjük-e a kockát? Mivel szabályos kocka esetén bármely szám dobása egyaránt $\frac{1}{6}$ valószínűségű, így a nullhipotézis

$$H_0 : p_i = \frac{1}{6} \qquad (i = 1, 2, \ldots, 6).$$

A feladatot χ^2-próbával oldjuk meg. A fentiek alapján

$$\chi^2 = \sum_1^6 \frac{(\nu_i - 6p_i)^2}{6p_i} = 0.7857.$$

A szabadsági fokok száma $f = r - 1 = 5$. A χ^2-eloszlás táblázatából az $\epsilon = 0.05$ értéknek megfelelően azt kapjuk, hogy $\chi^2_\epsilon = 11.07$, ami nagyobb, mint a minta alapján számított fenti érték, így a nullhipotézis elutasítására 95%-os szignifikanciaszinten nincs okunk, a kocka szabályosnak tekinthető.

18.6. Feladatok

18.1. Valamely gép által az (1)-es, illetve a (2)-es beállítás mellett gyártott termékek egyik jellemzője $\mathcal{N}(10, 1)$, illetve $\mathcal{N}(10.25, 1)$ eloszlású valószínűségi változó. Egy hosszú gyártási folyamat során elfelejtették feljegyezni, hogy melyik beállítással termelt a gép. Megvizsgálják 100 termékre a szóban forgó méretet. Mi legyen a döntésünk, hogy a helytelen (1)-es döntés valószínűsége ne legyen több 0.02-nél?

18.2. Egy érmét 100-szor feldobtunk, és 63 fej jött ki. Szabályosnak mondható-e az érme?

18.3. Egy betegséget egy bizonyos gyógyszer $p = 0.6$ valószínűséggel gyógyít meg. Valaki egy új gyógyszerről azt állítja, hogy jobb az előzőnél. Ha 20 emberből 15 meggyógyult az új gyógyszer hatására, akkor hogyan döntünk a $p < 0.6$ nullhipotézisről 95%-os szignifikanciaszinten?

18.4. Kételemű mintát veszünk a $(0, c)$-ben ($c \leq 1$) egyenletes eloszlásból. A $c = 1$ nullhipotézisről úgy döntünk, hogy ha $\max(X_1, X_2) \geq 0.8$, akkor elfogadjuk, különben elutasítjuk. Milyen a másodfajú hiba c-től való függése?

18.5. Egy bizonyos betegség esetén 10% a halandóság. 200 hegyilakó közül 29-en haltak meg ebben a betegségben. Jelenti-e ez azt, hogy a hegyilakóknak átlagon felüli a halandósága ebben a betegségben?

18.6. Egy csomagológép által készített bálákat vizsgálunk. 100 bála mérlegelése után azt kaptuk, hogy a bálák átlagos tömege 705 kg. Legyen a bálák tömege normális eloszlású, 50 kg szórással. Van-e szignifikáns eltérés 95%-os megbízhatósági szinten az előírt 700 kg-hoz képest?

19. Regressziók

19.1. Kétváltozós regresszió

A kétváltozós regresszió alapvető problémája úgy fogalmazható, hogy egy η valószínűségi változó értékeit egy másik ξ valószínűségi változó értékeivel akarjuk becsülni. A módszer lényege az, hogy a két valószínűségi változó együttes eloszlásának segítségével próbálunk olyan függvényt keresni, amelybe a ξ konkrét értékeit behelyettesítve az η értékeinek jó közelítését nyerjük. A közelítés "jóságát" természetesen definiálni kellene, s erre több lehetőség kínálkozik. A regressziószámításban a közelítés jóságát általában a négyzetes eltérés várható értékével szokás mérni. Ezzel kapcsolatban bizonyítás nélkül megemlítjük a következő tételt.

19.1.1. Tétel *Legyenek η, ξ valószínűségi változók. Tegyük fel, hogy η-nak létezik várható értéke és a ξ-re vonatkozó feltételes várható értéke. Ekkor az*

$$M[(\eta - h(\xi))^2]$$

várható érték mindazon h függvények körében, amelyekre $h(\xi)$ valószínűségi változó, akkor minimális, ha

$$h(\xi) = M(\eta|\xi).$$

A tétel szerint tehát az η-nak ξ-re vonatkozó regressziója az a függvény, amely a legkisebb négyzetes eltérés értelmében a lehető legjobban közelíti az η-t.

19.2. Lineáris regresszió

A regressziós függvényt általában nem könnyű meghatározni, ezért gyakran nem az összes függvények körében keressük a legjobban közelítőt, hanem valamely speciális függvényosztályban. Ha ez a függvényosztály a lineáris függvények osztálya, akkor a kapott legjobban közelítő függvényt *regressziós egyenesnek* nevezzük. A regressziós egyenest tehát

$$y = ax + b$$

alakban keressük, ismeretlen valós a, b együtthatókkal. Az a és b meghatározásához az

$$M[(\eta - a\xi - b)^2]$$

várható értéket kell minimalizálni. Tételezzük fel a továbbiakban, hogy az η és ξ szórása pozitív. Egyszerűen adódik, hogy a fenti kifejezés akkor minimális, ha

$$a = r\frac{\sigma_2}{\sigma_1},$$

$$b = m_2 - m_1 r\frac{\sigma_2}{\sigma_1},$$

ahol

$$r = \frac{M(\eta\xi) - M(\eta)M(\xi)}{D(\eta)D(\xi)},$$

az η és ξ korrelációs együtthatója, továbbá m_1, m_2, ill. σ_1, σ_2 jelölik a ξ, illetve η várható értékét, illetve szórását. A regressziós egyenes egyenlete tehát

$$y = r\frac{\sigma_2}{\sigma_1}x + m_2 - m_1 r\frac{\sigma_2}{\sigma_1}.$$

A regressziós egyenes iránytangensét *regressziós együtthatónak* nevezzük. Látható, hogy a regressziós egyenes akkor és csak akkor párhuzamos az x tengellyel, ha az η és ξ korrelálatlanok. Ha az egyenes egyenletében x helyére ξ-t írunk, megkapjuk az η közelítését

$$\overline{\eta} = r\frac{\sigma_2}{\sigma_1}\xi + m_2 - m_1 r\frac{\sigma_2}{\sigma_1}$$

alakban. Ehhez a közelítéshez azonban ismernünk kellene a korrelációs együttható, valamint a várható értékek és a szórások pontos értékét. Ezeket általában csak becsülni tudjuk, valamely minta alapján. Ha $((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n))$ n-elemű minta a (ξ, η) vektorváltozóra, akkor a várható értékek becslésére az

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

mintaközepeket, a regressziós együttható becslésére pedig az

$$\overline{a} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

kifejezést használjuk. Ennek az eljárásnak a helyességét a következőképpen támaszthatjuk alá: egy adott minta esetén a megfelelő $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ értékek a síkon n pontot reprezentálnak. Felvethető a kérdés, hogy melyik az az $y = ax + b$ egyenletű egyenes, amely a legkisebb négyzetek elve alapján a legjobban illeszkedik ezekre a pontokra. Más szóval, milyen a és b mellett lesz a

$$\sum_{i=1}^{n}(y_i - ax_i - b)^2$$

kifejezés minimális? A kérdésre válaszként az

$$\overline{a} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

és

$$\overline{b} = \overline{Y} - \overline{X}\overline{a}$$

értékeket kapjuk, összhangban a fentiekkel.

19.3. Regresszió normális eloszlás esetén

A következőkben részletesen megvizsgáljuk azt az esetet, amikor ξ és η együttes eloszlása kétdimenziós normális eloszlás. Az együttes sűrűségfüggvény ekkor

$$h(x, y) =$$

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2} - 2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2} + \frac{(y-m_2)^2}{\sigma_2^2}\right]},$$

ahol m_1, m_2, illetve σ_1, σ_2 jelölik a ξ és η várható értékét, illetve szórását, r pedig a ξ és η korrelációs együtthatója. Ekkor tehát ξ és η $\mathcal{N}(m_1, \sigma_1)$, illetve

$\mathcal{N}(m_2, \sigma_2)$ eloszlású valószínűségi változó. Ha f jelöli a ξ sűrűségfüggvényét, akkor kiszámítható az η változónak a ξ-re vonatkozó feltételes sűrűségfüggvénye:

$$g_\xi(y|x) = \frac{h(x,y)}{f(x)} =$$
$$= \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)\sigma_2^2}\left(y-m_2-r\frac{\sigma_2}{\sigma_1}(x-m_1)\right)^2}.$$

Ha itt x rögzített, akkor ez egy $\mathcal{N}(m_2 + r\frac{\sigma_2}{\sigma_1}(x - m_1), \sigma_2\sqrt{1-r^2}))$ eloszlás sűrűségfüggvénye, amelynek várható értéke

$$M(\eta|\xi = x) = \int_{-\infty}^{\infty} y g_\xi(y|x)dy = m_2 + r\frac{\sigma_2}{\sigma_1}(x - m_1).$$

Ebben az esetben tehát a regressziós függvény egyenes: együttes normális eloszlás esetén a regressziós egyenes nem csupán az egyenesek, hanem az összes függvények osztályában is a legjobb közelítést szolgáltatja.

19.4. Linearizálható regresszió

Két valószínűségi változó esetén gyakran mérési adatok, közvetlen megfigyelések, mintavétel útján igyekszünk következtetést levonni a köztük fennálló esetleges függvénykapcsolat típusára. A megfigyelt értékek közötti feltételezett összefüggéseket úgynevezett *empirikus képlettel* igyekszünk leírni. Ezek a képletek általában bizonyos paramétereket tartalmaznak, melyek értékét a minta alapján akarjuk közelíteni. Természetesen az empirikus képlet nem feltétlenül írja le a két valószínűségi változó között esetleg fennálló pontos függvénykapcsolatot. Ilyenkor az adatoktól való eltérés minimalizására leggyakrabban a legkisebb négyzetek módszerét szokás alkalmazni. A lineáris regresszió esetében az empirikus képlet alapja valójában egy lineáris összefüggés feltételezése a két változó között, amely nem biztos, hogy mindig megalapozott. Lineáris összefüggést olyankor jogos feltételezni, amikor a korrelációs együttható abszolút értéke közel 1. Más esetekben – az adatoktól függően – más függvényosztályokban kell keresni az empirikus képlet típusát. Bizonyos esetekben egyszerű változótranszformációkkal az empirikus képlet *linearizálható.* Például, az $y = ax^b$ empirikus képlet az $u = \ln x$, $v = \ln y$ helyettesítéssel a $v = bu + \ln a$ lineáris összefüggésbe megy át, így ha az eredeti adatokat kettős logaritmikus skálájú koordinátarendszerben ábrázoljuk, akkor ezek egy egyenesen fekszenek. Ha tehát az eredetileg választott empirikus képlet közelítőleg helyesen fejezi ki az adatok közti összefüggést, akkor a logaritmusaik közti összefüggés közelítésére jogosan használjuk a regressziós egyenest. További példák könnyen linearizálható empirikus képletekre a következők: $y = ae^{bx}$, $y = ax^n + b$, $y = \frac{x}{ax+b}$, stb. A megfelelő linearizáló helyettesítések leírását az olvasóra bízzuk.

19.5. Feladatok

19.1. A ξ, η valószínűségi változópárra rendelkezésre áll a következő minta:

x	101.3	103.7	98.6	99.9	97.2	100.1
y	609	626	586	594	579	605

Határozzuk meg a korrelációs együttható becsült értékét, s a két változó között lineáris kapcsolatot feltételezve határozzuk meg a regressziós egyenes egyenletét!

19.2. A ξ, η valószínűségi változókra rendelkezésünkre áll a következő minta:

x	1	2	2.5	3	5	5.6	5.8	6	9	10
y	3	4.5	6.3	7	8	18	16.8	18	21	20

Lineáris kapcsolatot feltételezve határozzuk meg a regressziós egyenest!

19.3. A ξ, η valószínűségi változókra rendelkezésünkre áll a következő minta:

x	1970	1971	1972	1973	1974
y	13	24	32	46	57

Lineáris kapcsolatot feltételezve határozzuk meg a regressziós egyenest!

19.4. A ξ, η valószínűségi változókra rendelkezésünkre áll a következő minta:

x	1	1.5	2	2.5	3	3.5
y	3.1	4.23	5.8	8.26	10.87	14.20

Válasszunk alkalmas empirikus képletet, s paramétereit határozzuk meg a legkisebb négyzetek módszerével!

Irodalom

References

1. Baróti György - Bognár Jánosné - Fejes-Tóth Gábor - Mogyoródi József, *Valószínűségszámítás, Egyetemi jegyzet*, Tankönyvkiadó, Budapest, 1978.
2. Bognár Jánosné - Göndőcs Ferenc - Kászonyi László - Kováts Antal - Michaletzky György - Móri Tamás - Somogyi Árpád - Szeidl László - Székely J. Gábor, *Matematikai statisztika* (Mogyoródi József - Michaletzky György, eds.), Nemzeti Tankönyvkiadó, Budapest, 1995.
3. Denkinger Géza, *Valószínűségszámítás*, Tankönyvkiadó, Budapest, 1969.
4. Denkinger Géza, *Valószínűségszámítás (Példatár)*, Tankönyvkiadó, Budapest, 1970.
5. Feller, W., *Bevezetés a valószínűségszámításba és alkalmazásaiba*, Műszaki Könyvkiadó, Budapest, 1978.
6. Jordan Károly, *Fejezetek a klasszikus valószínűségszámításból*, Akadémiai Kiadó, Budapest, 1956.
7. *Matematika II., Tankönyvpótló jegyzet és feladatgyűjtemény*, Közgazdasági és Jogi Könyvkiadó, Budapest, 1978.
8. Meszéna György - Ziermann Margit, *Valószínűségelmélet és matematikai statisztika*, Közgazdasági és Jogi Könyvkiadó, Budapest, 1981.
9. Obádovics J. Gyula, *Valószínűségszámítás és matematikai statisztika*, Scolar Kiadó, Budapest, 1995.
10. Prékopa András, *Valószínűségelmélet műszaki alkalmazásokkal*, Műszaki Könyvkiadó, Budapest 27, 1962.
11. Rényi Alfréd, *Valószínűségszámítás*, Tankönyvkiadó, Budapest, 1973.
12. Solt György, *Valószínűségszámítás (Példatár)*, Műszaki Könyvkiadó, Budapest, 1969.