walszam2017-jav-180121.doc, ### 2018.01.23., 12:00'

http://math.uni-pannon.hu/~szalkai/walszam2017.pdf

Valószínűségszámítás alapjai szemléletesen

/Kézirat, 2018-01-23. /

dr.Szalkai István

Pannon Egyetem, Veszprém

Matematika Tanszék

http://math.uni-pannon.hu/~szalkai

szalkai@almos.uni-pannon.hu

Bevezetés

Ez csak egy nagyon vázlatos, félkész kézirat, segítség az előadáshoz .
A kisebb egységek végét □ jelzi.
További anyagok találhatók még a honlapomon: http://math.uni-pannon.hu/~szalkai

dr. Szalkai István
Veszprém, 2018.01.23.
szalkai@almos.uni-pannon.hu

Tartalom

1. Események és eseménytér

1.1. Alapfogalmak (Definíciók):

Kísérlet: egy jelenség aktív vagy passzív megfigyelése
Determinisztikus (=meghatározott) kísérlet: mindenképpen csak egy lehetséges kimenetele (végeredménye) lehet.
Sztochasztikus (=véletlen) kísérlet: több kimenetele van
Példák: érme-, kockadobás, két kocka, izzó élettartama, fizikai mennyiségek mérése, vízállások, lottóhúzás, minőségellenőrzés, stb.
Elemi események: A kísérlet lehetséges, tovább már nem bontható kimenetelei. Az összes lehetséges elemi esemény *halmazát* (összességét) **eseménytér**nek nevezzük, mi Ω -val jelöljük. (Más könyvekben eltérő jelölés is lehet). A kísérlet egy elemi eseménye Ω -nak egy eleme: $\omega \in \Omega$ vagy $x \in \Omega$.
Matematikailag Ω egy *tetszőleges*, nem üres halmaz lehet!
Példák: "hatost dobtam", "3V", "2kg", ...
Esemény (általánosan): Ω bármely *részhalmaza*: $A \subseteq \Omega$.
Példák: "páros számot dobtam", "10V és 20V között van", ...
Az elemi eseményeket tehát Ω egyelemű részhalmazainak kell (illik) tekintenünk: $\omega \in \Omega$ helyett inkább $\{\omega\} \subseteq \Omega$ írandó.
Az események összessége (halmaza) nyilvánvalóan Ω **hatványhalmaza** (power set), aminek jól ismert jele $P(\Omega)$.

A kísérlet (vég)**eredménye** tehát mindig Ω egy eleme: $\omega \in \Omega$. Azt mondjuk, hogy egy $A \subseteq \Omega$ esemény egy kísérlet (elvégzése) során **bekövetkezik**, ha $\omega \in A$.

Az alábbi fogalmak szemléletesek, de meg kell gondolnunk matematikai hátterüket.
(Ezeket a fogalmakat a későbbiekben még általánosítjuk.)
Biztos esemény, ami mindig bekövetkezik tehát az egész alaphalmaz Ω .
Lehetetlen esemény, ami nem következhet be, tehát az üres halmaz $\emptyset$.
Egymást kizáró események csak diszjunktak lehetnek, vagyis $A \cap B = \emptyset$.
Az A eseményből **következik** a B esemény, vagy más szóval A **maga után vonja** B -t, ha minden $\omega \in \Omega$ esetén $\omega \in A$ -ból $\omega \in B$ következik. Ez pedig éppen azt jelenti, hogy $A \subseteq B$, vagyis A **részhalmaza** B -nek.

1.2. Műveletek eseményekkel (eseményalgebra)

Mivel az események valójában Ω részhalmazai, ezért a szokásos halmazműveletekről van szó: únió, metszet, komplementer, különbség. A valószínűségszámítás elméletében más elnevezések és jelölések használatosak:

események **összege**	**A+B** =	$A \cup B$ = únió,
események **szorzata**	**A·B** =	$A \cap B$ = metszet,
események **különbsége**	**A-B** =	$A \setminus B$ = különbség,
A esemény **tagadása**	$\overline{A}$ =	$\overline{A}$ = komplementer.

Mi legtöbbször a hagyományos, halmaz- elnevezéseket használjuk.

1.3. A műveletek tulajdonságai

A fentiek alapján a halmazműveletek (Boole-algebrák) jól ismert tulajdonságairól van szó. Tanácsoljuk az Olvasónak az egyenlőségeket a valószínűségszámítás nyelvére lefordítani.

Boole-algebrák axiómái

$$\mathfrak{B} = (X, \cup, \cap, \overline{}, \emptyset, I)$$

kommutativitás	$A \cup B = B \cup A$	(BA1)
	$A \cap B = B \cap A$	(BA2)
asszociativitás	$A \cup (B \cup C) = (A \cup B) \cup C$	(BA3)
	$A \cap (B \cap C) = (A \cap B) \cap C$	(BA4)
disztributivitás	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	(BA5)
	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	(BA6)
elnyelési tulajdonságok	$A \cup (A \cap B) = A$	(BA7)
	$A \cap (A \cup B) = A$	(BA8)
$\emptyset$ és I tulajdonságai	$A \cup \overline{A} = I$	(BA9)
	$A \cap \overline{A} = \emptyset$	(BA10)
	$A \cup \emptyset = A$	(BA11)
	$A \cap \emptyset = \emptyset$	(BA12)
	$A \cup I = I$	(BA13)
	$A \cap I = A$	(BA14)

A fenti axiómák következményei például az alábbiak:

De Morgan-azonosságok: $\overline{A \cup B} = \overline{A} \cap \overline{B}$ és $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

2. A relatív gyakoriság és a valószínűség

2.0. Ha egy kísérletet a gyakorlatban n -szer elvégzünk, és egy A esemény k -szor következett be (gyakorisága k), akkor a k/n hányadost **relatív gyakoriság**nak hívjuk.

Tapasztalatok alapján hihető, hogy ha n vagy, akkor a k/n hányados egy bizonyos, az A eseményre jellemző p szám körül ingadozik. (Ezt az összefüggést a *Nagy számok törvényei* fejezetben Bernoulli tétele igazolja.) Ezt a p számot hívjuk az A esemény **valószínűségének**, és **P(A)** -val jelöljük.

2.1. Definíció: A valószínűség **Kolmogorov-féle axiómái:**

P (egy) valószínűség az Ω eseménytéren, ha

(o) P : $\mathcal{P}(\Omega) \to \mathbf{R}$, vagyis $A \subseteq \Omega$ esetén $P(A) \in \mathbf{R}$ valós szám,

(i) $0 \le P(A) \le 1$,

(ii) $P(\emptyset) = 0$, $P(\Omega) = 1$,

(iii) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ ha az A_i -k egymást *páronként kizárják*, vagyis ha $A_i \cap A_j = \emptyset$. □

2.2. A fenti axiómák **következményei:**

1. tetszőleges $A,B \subseteq \Omega$ eseményekre $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,
2. ha A és B kizárják egymást, akkor $P(A \cup B) = P(A) + P(B)$ (additivitás),
3. $P(\overline{A}) = 1 - P(A)$,

4. tetszőleges A,B$\subseteq\Omega$ eseményekre $P(A \setminus B) = P(A) - P(A \cap B)$,
5. ha $B \subset A$, akkor $P(A \setminus B) = P(A) - P(B)$,
6. ha $B \subset A$, akkor $P(B) \le P(A)$ (monotonitás),
7. tetszőleges A,B,C $\subseteq\Omega$ eseményekre
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$
(logikai szitaformula) ,
8. ... □

2.3. Vegyük észre, hogy *P* fenti tulajdonságai nagyon hasonlítanak a síkbeli halmazok *területé*-nek tulajdonságaival (ami nem meglepő, mert mindkettő *additív halmazfüggvény*). Tehát tanuláskor nyugodtan olvashatunk *P*(A) helyett T_A -t.

Az 1.1.Definíció *általánosításaként* mondjuk az alábbiakat:

2.4. Definíció: Tetszőleges A,B$\subseteq\Omega$ eseményekre:
A **biztos esemény**, ha P(A)=1 ,
A **lehetetlen esemény**, ha P(A)=0 ,
A és B **egymást kizáró események**, ha P(A$\cap$B)=0 . □

2.5. Megjegyzés: Általánosan az "*esély*" és "*valószínűség*" szavak egymás szinonímái, de egyes statisztikai művek az "*esély*" fogalom alatt <u>mást</u> (*p*/*q*) értenek.

3. A valószínűség kiszámítása

Most csak a két legegyszerűbb módszert ismertetjük, hiszen az egész anyag célja végig a valószínűség kiszámítása !

3.1. Kombinatorikus (klasszikus) valószínűségi mező
<u>Amennyiben</u> Ω véges elemszámú halmaz, <u>és</u> minden $\omega\in\Omega$ elemi eseménynek *ugyanakkora* a valószínűsége (pl. szabályos kockával dobás), <u>akkor</u>
egyrészt $P(\{\omega\}) = \frac{1}{n}$ minden elemi esemény esetén,
másrészt tetszőleges A$\subseteq\Omega$ eseményre $P(A) := \frac{|A|}{|\Omega|} = \frac{kedvez\H{o}}{\ddot{o}sszes}$. □

3.2. Geometriai valószínűség:
<u>Amennyiben</u> a Ω halmaz a számegyenes, a sík vagy a tér valamely olyan részhalmazával reprezentálható (modellezhető, megjeleníthető), <u>amely</u> esetén minden $\omega\in\Omega$ elemi eseménynek *ugyanakkora* a valószínűsége, <u>akkor</u> tetszőleges A$\subseteq\Omega$ eseményre
$P(A) = \frac{\mu(A)}{\mu(\Omega)}$ (arányos a halmaz hosszával/területével/térfogatával) ,
ahol értelemszerűen μ(A) és $\mu(\Omega)$ a halmazok hosszát, területét vagy térfogatát jelöli. □

3.3. Például: *buszra várok* (de csak véletlenszerűen mentem ki a megállóba),
célbadobás (de csak véletlenszerűen célzok, nem vagyok profi, <u>és</u> minden lövésemmel legalább a céltáblát eltalálom).
Érdemes még a "*Randevú a könyvtárban*" -típusú feladatokat us tanulmányozni.
A "*célbadobás*" -nál kihangsúlyozott feltételeket minden esetben alaposan ellenőrizni kell !

4. Feltételes valószínűség, események függetlensége

Ha már tudjuk, hogy valami bekövetkezett (egy B esemény), akkor ez mennyire befolyásolhatja az általunk vizsgált A esemény bekövetkezését? A relatív gyakoriság elemzésével kapjuk a következő képletet:

4.1. Definíció: Amennyiben a B esemény nem lehetetlen, vagyis P(B)>0 , akkor A -nak B bekövetkezése utáni valószínűsége $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$.

Tehát P(A|B) az A esemény valószínűsége, **feltéve**, hogy B már bekövetkezett, ezért hívjuk **feltételes valószínűség**nek. □

4.2. Megjegyzések: A *feltételes valószínűség* teljesíti a valószínűség axiómáit, ha a B feltétel rögzített. Ez azt jelenti, hogy a 2.Fejezet képletei mind igazak maradnak, ha $P(...)$ helyett mindenütt $P(...|B)$ -t írunk.

Természetesen merül fel a kérdés: B hogyan befolyásolja A -t (erősíti, gyengíti vagy nem befolyásolja). Ezt alább, a *Függetlenség* alfejezetben vizsgáljuk.

A 4.1.Definíció képletének átszorzása után kapjuk a következő egyszerű, de fontos összefüggést:

4.3. Szorzás Tétel: $P(A \cap B) = P(A \mid B) \cdot P(B)$. □

4.4. Definíció: A $B_1, B_2, B_3, ..., B_n \subseteq \Omega$ események **teljes eseményrendszert** alkotnak, ha egymást **páronként kizárják**, és uniójuk a biztos esemény, vagyis $B_i \cap B_j = \emptyset$, és

$$B_1 \cup B_2 \cup ... \cup B_n = \Omega \text{ .} \quad \square$$

(Általánosabban: $P(B_i \cap B_j)=0$ és $P(B_1 \cup B_2 \cup ... \cup B_n)=1$.)

A fenti tulajdonságú halmazrendszereket általában *partíciónak* (*felosztásnak*) nevezünk, lásd az alábbi ábrát.

4.5. Teljes valószínűség tétele:

Ha $\{ B_1, B_2, B_3, ..., B_n \}$ teljes eseményrendszer és $P(B_i)>0$ minden i-re, akkor bármely A eseményre

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$$

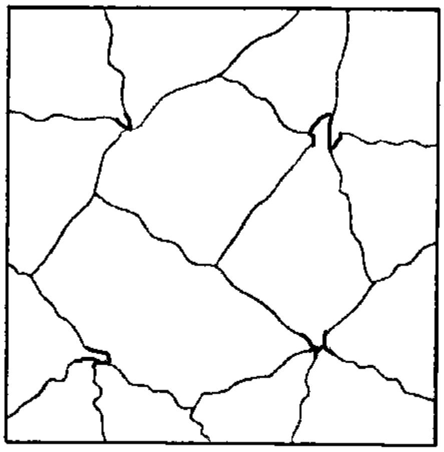
Bizonyítás: A Szorzás Tétel alapján a fenti képlet így írható:

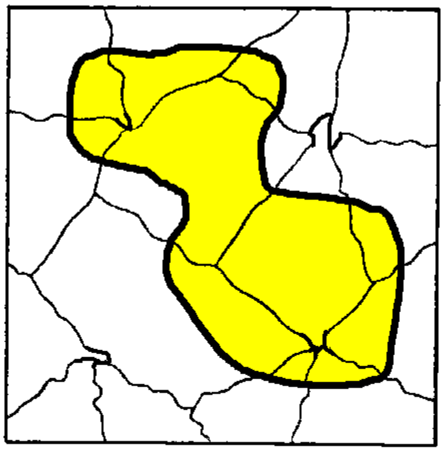
$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$

ami nyilván igaz, hiszen

$$A = \bigcup_{i=1}^{n} \left(A \cap B_i \right) \text{ .} \quad \square$$

A gondolatmenetet szemléltetik az alábbi ábrák (P helyett gondoljunk ismét a területre):





Teljes eseményrendszer (partíció) **Teljes valószínűség**

4.6. Példa: Egy gyárban 3 műszakban gyártanak termékeket. Az I. műszak az összes termék 40%-át, a második az összes termék 35%-át, a harmadik műszak az összes termék 25%-át gyártja. Az első műszak által gyártott termékek mindegyike 0.05, a második műszak által gyártott termékek mindegyike 0.06, a harmadik műszak által gyártott termékek mindegyike 0.07 valószínűséggel selejtes. Kiválasztva egy akármilyen terméket, mennyi az esélye, hogy selejtes?
Megoldás: Legyenek B_1,B_2, B_3 rendre azon események, hogy a kiválsztott terméket melyik műszak gyártotta, és jelölje S azt, amikor a termék selejes. A feltételek szerint $P(B_1)=0.4$, $P(B_2)=0.35$, $P(B_3)=0.25$, (ellenőrzés: $P(B_1)+P(B_2)+P(B_3)= 0.4+0.35+0.25=1$). Továbbá $P(S|B_1)=0.05$, $P(S|B_2)=0.06$, $P(S|B_1)=0.07$.
A teljes valószínűség tétele szerint

$$P(S) = P(S|B_1)\cdot P(B_1) + P(S|B_2)\cdot P(B_2) + P(S|B_3)\cdot P(B_3) =$$
$$= 0.05*0.4 + 0.06*0.35 + 0.07*0.25 = 0.0585 .$$

Megfordítás: Ha a kiválasztott termék selejtes, mennyi az esélye, hogy az első (vagy a második, vagy a harmadik) műszak gyártotta? Kit szidjunk (legnagyobb eséllyel)? Mert ugye a harmadik műszak termékei között van a legtöbb selejt, de ugyanekkor a harmadik műszak állítja elő a legkevesebb terméket ... . Az alábbi tétel erre a kérdésre válaszol:

4.7. Bayes tétele (Megfordítási tétel): Ha $B_1, B_2, B_3,...., B_n$ teljes eseményrendszer és $P(B_i)>0$ minden *i*-re, akkor bármely A eseményre (ha $P(A)>0$)

$$P(B_j \mid A) = \frac{P(A \mid B_j)\cdot P(B_j)}{P(A)} \quad . \quad \square$$

A tétel könnyen következik a Szorzás Tételből.

A 4.6.Példa folytatása:

$P(B_1|S) = P(S|B_1)\cdot P(B_1)/P(S) = 0.05*0.4 /0.0585 \approx 0.341\,880$,
$P(B_2|S) = P(S|B_2)\cdot P(B_2)/P(S) = 0.06*0.35/0.0585 \approx 0.358\,974$,
$P(B_3|S) = P(S|B_3)\cdot P(B_3)/P(S) = 0.07*0.25/0.0585 \approx 0.299\,145$,

tehát (kis eltéréssel) a selejtes darabok legnagyobb része a második műszaktól származik.
(Ellenőrzés: $P(B_1|S)+P(B_2|S)+P(B_3|S)=1$.)

Események függetlensége

Felmerül a kérdés, hogy a már bekövetkezett B mennyire befolyásolja A -t ? Nyilván három eset lehetséges:

$P(A|B) < P(A)$, vagyis B gyengíti A -t ,
$P(A|B) > P(A)$, vagyis B erősíti A -t ,
$P(A|B) = P(A)$, vagyis B nem befolyásolja A -t .

4.8. Speciális esetek: Az 1.Fejezetben megismert speciális $B \subseteq A$ ("B -ből *következik* A") esetben a 4.1.Definíció képlete szerint

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$, vagyis ekkor B -ből valóban következik A .

Ha pedig kizárják egymást (a 2.4.Definíció szerint), akkor

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{P(B)} = 0$$, vagyis B valóban kizárja A -t.

4.9. Nyilvánvaló követelmény, hogy az A és B események akkor lehetnek *függetlenek*, ha egyik sem befolyásolja a másikat, vagyis

$$P(A|B) = P(A) \quad \text{és} \quad P(B|A) = P(B) .$$

Kis számolás után (a Szorzás tétel segítségével) kapjuk, hogy a fenti két követelmény ekvivalens (azonos értékű) a következő, egyetlen összefüggéssel:

4.10. Definíció: Az *A* és *B* események **egymástól függetlenek,** ha $P(A \cap B) = P(A) \cdot P(B)$. □

Hangsúlyozzuk, hogy a fenti egyenlőség nem érvényes tetszőleges eseményekre !!!

4.11. Állítás: Ha $0 < P(A) < 1$, $0 < P(B) < 1$ és A és B függetlenek egymástól, akkor $\overline{A}$ és B is, A és $\overline{B}$ is, valamint $\overline{A}$ és $\overline{B}$ is függetlenek egymástól. □

5. Valószínűségi változók és jellemzőik

A legtöbb kísérletnél nem csak "esemény"-t észlelünk (piros, sikerült, stb.), hanem mérünk is (valamit). Általában többször megmérve "ugyanazt" a jelenséget (pl. testtömeg), mindig más eredményt kapunk, véletlenszerűen váltakozva.

5.1. Definíció: Az olyan ξ függvényeket, amelyek elemi eseményekhez rendelnek valós számokat, azaz $\xi : \Omega \to R$, **valószínűségi változó**knak (v.v.) nevezzük. □
(Könnyen megjegyezhető: "val.vált." = "a kísérlet számszerű végeredménye".)

A valószínűségi változók alábbi két típusának megkülönböztetése lényeges az anyag további részében: minden definíciónál, tételnél lényeges, hogy az milyen típusú val.vált.-ra érvényes!

5.2. Definíció: A ξ valószínűségi változó **diszkrét** (elkülönült), ha csak véges vagy megszámlálhatóan (felsorolhatóan) sok lehetséges értéke van, tehát értékkészlete $Im(\xi) = \{x_1, x_2, \ldots, x_n, \ldots\}$ alakban írható.
A ξ valószínűségi változó **folytonos**, ha értékkészlete tartalmaz egy intervallumot: $Im(\xi) \supseteq (a,b)$. □

5.3. Definíció: A ξ diszkrét valószínűségi változó **eloszlásán** értjük a lehetséges értékeinek halmazát a hozzájuk tartozó valószínűségekkel: $\{x_1, x_2, \ldots, x_n, \ldots\}$ és $\{p_1, p_2, \ldots, p_n, \ldots\}$ ahol $p_i := P(\xi = x_i)$. □

Nyilván ξ és η -nak akkor van **azonos eloszlása,** ha $\{x_1, x_2, \ldots, x_n, \ldots\}$ és $\{p_1, p_2, \ldots, p_n, \ldots\}$ ugyanazok.

5.4. Állítás: Egy $\{p_1, p_2, \ldots, p_n, \ldots\}$ számsorozat akkor és csak akkor egy diszkrét valószínűségi változó eloszlása, ha az alábbi **axiómák**at teljesíti:
(i) $0 \le p_i \le 1$,
(ii) $p_1 + p_2 + \ldots + p_n + \ldots = 1$. □

5.5. Definíció: Egy ξ tetszőleges (akár diszkrét, akár folytonos) valószínűségi változó **eloszlásfüggvénye** $F: R \to R$, ahol $F(x) := P(\xi < x)$. □

5.6. Tétel: Az eloszlásfüggvény alaptulajdonságai (axiómák):

1) $0 \le F(x) \le 1$,
2) $F(x)$ monoton nő,
3) $F(x)$ balról folytonos ("teli karika" a jobb végén),
4) $\lim_{x\to\infty} F(x) = 1$ és $\lim_{x\to-\infty} F(x) = 0$. □

Mire való az *eloszlásfüggvény*? Segítségével az alábbi kérdésekre tudunk gyors választ adni:

5.7. Tétel:

$P(\xi < a) = F(a)$

$P(\xi \ge a) = 1 - F(a)$

$P(\xi \le a) = F(a) + P(\xi = a)$

$P(\xi > a) = 1 - F(a) - P(\xi = a)$

$P(a \le \xi < b) = F(b) - F(a)$

$P(a \le \xi \le b) = F(b) - F(a) + P(\xi = b)$

$P(a < \xi < b) = F(b) - F(a) - P(\xi = a)$

$P(a < \xi \le b) = F(b) - F(a) + P(\xi = b) - P(\xi = a)$

$P(\xi = a) = \lim_{x\to a+} F(x) - F(a)$ □

Az 5.2.Definíció helyett matematikailag az alábbi követelmény pontosabb:

5.8. Definíció: ξ **folytonos**, ha létezik egy olyan f : R→R folytonos függvény, amelyre (legfeljebb véges sok pont kivételével)

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

Ekkor f(x) -et ξ **sűrűségfüggvényé**nek nevezzük. □

A ξ és η v.v. **azonos eloszlásúak**, ha sűrűségfüggvényeikre $f_\xi(x)=f_\eta(x)$ véges sok $x \in$R pont kivételével. □

A "*sűrűségfüggvény*" elnevezést az 5.15. megjegyzésben világítjuk meg.

5.9. Állítás: A sűrűségfüggvény alaptulajdonságai (**axiómái**):

0) f : R→R   1) $f(x) \ge 0$ 2) $\int_{-\infty}^{\infty} f(x) = 1$. □

5.10. Állítás: ha az $f: R \to R$ (legfeljebb véges sok pont kivételével) folytonos függvény rendelkezik a fenti tulajdonságokkal, akkor van olyan ξ valószínűségi változó, aminek f a sűrűségfüggvénye. □

5.11. Állítás: *(i)* Ha ξ folytonos, akkor eloszlásfüggvénye ($F(x)$) minden pontban folytonos.
(ii) Ha $F(x)$ folytonos és véges sok pont kivételével folytonosan differenciálható, akkor van ξ folytonos eloszlású valószínűségi változó, amelynek $F(x)$ eloszlásfüggvénye.
(iii) Ahol F deriválható, ott $F'(x) = f(x)$. □

5.12. Állítás: Ha $F(x)$ folytonos x -ben, akkor $P(\xi = x) = 0$. □

Mire való az $f(x)$ *sűrűségfüggvény*? Segítségével az alábbi kérdésekre tudunk gyors választ adni:

5.13. Állítás: ha ξ folytonos eloszlású, akkor

$$P(a \le \xi < b) = P(a \le \xi \le b) = P(a < \xi \le b) = P(a < \xi < b) = F(b) - F(a) \ . \ \square$$

Most összefoglaljuk a fejezet legfontosabb képleteit:

5.14. Tétel: "Tipikus" kérdések (és a válaszok)

$P(\xi<b) \quad = \int_{-\infty}^{b} f(x)dx \; = \; F(b)$

$P(a\leq\xi) \quad = \int_{a}^{\infty} f(x)dx \; = \; 1\text{-}F(a) \; = \; 1\text{-} P(\xi<a)$

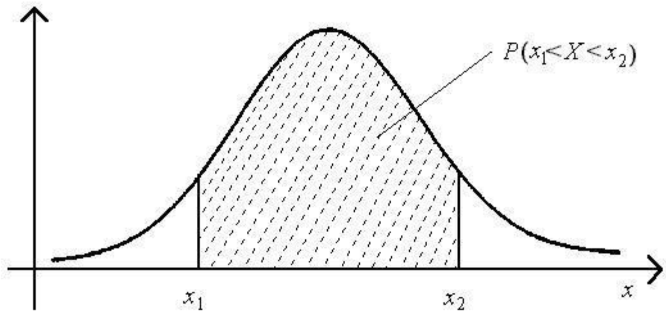
$P(a\leq\xi<b) \quad = \int_{a}^{b} f(x)dx \; = \; F(b)\text{-}F(a)$ (Newton-Leibniz szabály)

$P(\xi=b) \quad = \; 0$ (ha ξ folytonos v.v.)

$P(\xi\approx c) = P(|\xi\text{-}c|<\varepsilon) = P(c\text{-}\varepsilon<\xi<c+\varepsilon) = F(c+\varepsilon)\text{-}F(c\text{-}\varepsilon)$. □

5.15. Miért nevezzük $f(x)$ -et sűrűségfüggvénynek?

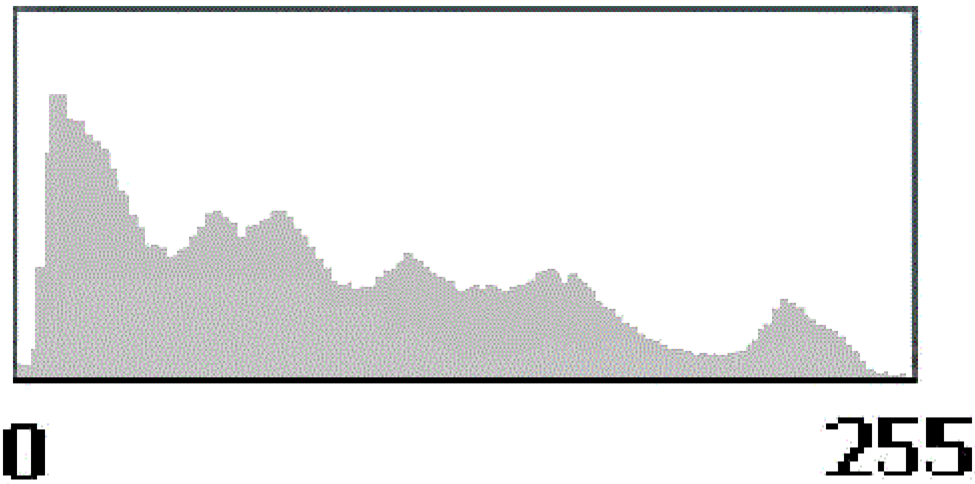
A Newton-Leibniz szabály alapján:



$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx.$$

vagy másképpen: $\lim_{\Delta a\to 0} \frac{P(a\leq\xi<a+\Delta a)}{\Delta a} = f(a)$. □

Hisztogram (oszlopdiagram) és sűrűségfüggvény kapcsolata:



Valószínűségi változók függetlensége

5.16. Definíció: A ξ és η valószínűségi **változókat független**nek nevezzük, ha bármely $x,y \in$R valós számok esetén

$$P(\xi < x \quad \text{és} \quad \eta < y) = P(\xi < x) \cdot P(\eta < y) \ . \qquad \square$$

5.17. Állítás: Ha ξ és η *diszkrét* v.v., és lehetséges értékeik Im(ξ)={ $x_1, x_2, \ldots$.} illetve Im(η)= { $y_1, y_2 \ldots$. }, akkor : ξ és η pontosan akkor függetlenek, ha

$$P(\xi = x_i \ , \eta = y_j) \ = \ P(\xi = x_i) \cdot P(\eta = y_j) \ . \ \square$$

5.18. Tétel: Ha : ξ és η folytonos eloszlásúak, sűrűségfüggvényük $f(x)$ illetve $g(y)$, akkor: ξ és η pontosan akkor függetlenek, ha $\dfrac{\partial^2 P(\xi < x \ , \eta < y)}{\partial x \, \partial y} = f(x) \cdot g(y)$. □

6. Várható érték és szórás

Egyetlen mérés helyett általában többször szoktunk mérni, és a méréseket utána átlagoljuk, ismétlődő értékek esetén súlyozzuk az értékeket.

6.1. Definíció: Ha ξ *diszkrét* v.v., lehetséges értékei { $x_1, \ldots$.} és eloszlása {p_1,...} , akkor ξ **várható értéke** vagy **átlaga**

- ha ξ-nek véges sok lehetséges értéke van, akkor

$$M(\xi) = E(\xi) = \sum_{i=1}^{n} x_i p_i \ ,$$

- ha ξ-nek végtelen sok lehetséges értéke van, akkor

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

feltéve, hogy $\sum_{i=1}^{\infty} |x_i| p_i < \infty$. □

6.2. Definíció: Ha ξ *folytonos* v.v., sűrűségfüggvénye *f*(*x*) , akkor ξ **várható értéke** vagy **átlaga**

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

amennyiben az $\int_{-\infty}^{\infty} |x| \cdot f(x) dx$ improprius integrál konvergens. □

6.3. Megjegyzések: (i) E(ξ) a régebben használt angol "*expected value*" (várt érték) kezdőbetűje, újabban inkább az M(ξ) jelölést használjuk, ami az angol "*mean*" (átlag!) szót jelöli.
(ii) Úgy tapasztaljuk hogy a ξ mennyiség többszöri mérésekor kapott értékek az M(ξ) átlag körül ingadoznak. Ezt a **Nagy számok gyenge törvénye (Csebisev-alak)** tétel igazolja a *Nagy számok* Fejezetben.

6.4. Tétel: A várható érték tulajdonságai:

1. Ha $\xi = c$, akkor $M(\xi) = c$ ("beragadt a mérőműszer") .
2. Ha létezik $M(\xi)$, akkor $M(a\xi + b) = aM(\xi) + b$ (pl. ^{o}C helyett ^{o}F) .

3. Ha $M(\xi)$ és $M(\eta)$ létezik, akkor $M(\xi+\eta)=M(\xi)+M(\eta)$.
4. Ha $a\leq\xi\leq b$, akkor $a\leq M(\xi)\leq b$.
5. Ha $0\leq\xi$, akkor $0\leq M(\xi)$.
6. Ha ξ és η függetlenek, akkor $M(\xi\eta)=M(\xi)M(\eta)$.
7. Ha $\xi_1,\xi_2,...\xi_n$ független azonos eloszlású valószínűségi változók, akkor

$$M(\sum_{i=1}^{n}\xi_i)=nM(\xi_1) \text{ és } M\left(\frac{\sum_{i=1}^{n}\xi_i}{n}\right)=M(\xi_1)$$

8. Ha $g:H\subset R\to R$ folytonos függvény, akkor diszkrét eloszlás esetén

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

folytonos eloszlás esetén $M(g(\xi))=\int_{-\infty}^{\infty}g(x)f(x)dx$.

9. Speciálisan: $g(x)=x^2$ esetén $M(\xi^2)=\sum_{i=1}^{\infty}x_i^2p_i$ illetve $M(\xi^2)=\int_{-\infty}^{\infty}x^2f(x)dx$. □

A szórás

Többször tapasztalhattuk (pl. iskolai dolgozatok osztályzatainál), hogy az átlagtól való szóródás néha kicsi, néha nagy. Ezeket az eltéréseket nyilván átlagolnunk kell.
Ha ξ-nek van várható értéke, akkor $M(\xi-M(\xi))=0$. $M(|\xi-M(\xi)|)$ sem jó, nem könnyű számolni.
De $(\xi-M(\xi))^2$ a kis eltéréseket még kisebbé, a nagyokat még nagyobbá teszi, ezért a $M((\xi-M(\xi))^2)$ mennyiséget használjuk az ingadozás mérésére.

6.6. Definíció: A ξ v.v. **szórásnégyzete** $D^2(\xi)=M\left((\xi-M(\xi))^2\right)$ (amennyiben ez a kifejezés véges), **szórása** $D(\xi)=\sqrt{M\left((\xi-M(\xi))^2\right)}$. □

6.7. Állítás: A fenti, gyök alatti mennyiség nemnegativitása biztosított a várható érték 6.4. Tétel 5) tulajdonsága miatt, tehát gyököt lehet vonni. □

6.8. Állítás: Ha $M(\xi^2)$ véges, akkor létezik a szórásnégyzet és $D^2(\xi)=M(\xi^2)-M^2(\xi)$.
Bizonyítás: A 6.4.Tétel pontjait alkalmazva

$$D^2(\xi)=M\left((\xi-M(\xi))^2\right)=M\left(\xi^2-2\xi\cdot M(\xi)+M^2(\xi)\right)=M(\xi^2)-2M(\xi)M(\xi)+M^2(\xi)=$$
$$=M(\xi^2)-M^2(\xi)$$ □

6.9. Következmény: A várható érték (6.4.Tétel 9.) tulajdonsága miatt a szórásnégyzetre a következő számolási képleteink vannak:

diszkrét valószínűségi változó esetén: $D^2(\xi)=\sum_{i=1}^{\infty}x_i^2p_i-\left(\sum_{i=1}^{\infty}x_ip_i\right)^2$,

folytonos valószínűségi változó esetén: $D^2(\xi)=\int_{-\infty}^{\infty}x^2f(x)dx-\left(\int_{-\infty}^{\infty}xf(x)dx\right)^2$. □

6.10. Tétel: A szórásnégyzet és a szórás tulajdonságai:

1. $\xi = c$ akkor és csak akkor, ha $D^2(\xi) = D(\xi) = 0$. ("Beragadt a mérőműszer".)
2. $D^2(a\xi + b) = a^2 D^2(\xi)$ és $D(a\xi + b) = |a| D(\xi)$.
3. CSAK Ha ξ és η *függetlenek,* akkor

$$D^2(\xi+\eta) = D^2(\xi) + D^2(\eta) \quad \text{és} \quad D(\xi+\eta) = \sqrt{D^2(\xi) + D^2(\eta)} \ .$$

4. Ha $\xi_1, \xi_2, ... \xi_n$ független, azonos eloszlású valószínűségi változók, akkor

$$D^2(\sum_{i=1}^{n} \xi_i) = n \cdot D^2(\xi_1) \quad , \quad D^2\left(\sum_{i=1}^{n} \xi_i \Big/ n\right) = \frac{D^2(\xi_1)}{n} \quad ,$$

$$D(\sum_{i=1}^{n} \xi_i) = \sqrt{n} D(\xi_1) \quad , \quad D\left(\sum_{i=1}^{n} \xi_i \Big/ n\right) = \frac{D(\xi_1)}{\sqrt{n}} \ . \qquad \square$$

6.11. Definíció: *Diszkrét* valószínűségi változó **módusza** a legvalószínűbb x_k érték (vagyis amelyre p_k=P(x_k) a legnagyobb).
Folytonos valószínűségi változó **módusza** a sűrűségfüggvény lokális maximumhelye. □

6.12. Definíció: *Diszkrét* valószínűségi változó **mediánja** a következő $m \in$R valós szám:
- ha az F(x) függvény sehol sem veszi fel az ½ értéket, akkor m legyen a legkisebb olyan $x \in$R amelyre F(x)> ½ .
- ha van olyan $x \in$R valós szám, amelyre F(x)= ½ , akkor ezen x számok egy intervallumot alkotnak, és legyen m az intervallum közepe.

Folytonos valószínűségi változó **mediánja** az F(x)= ½ egyenlet megoldása, ha pedig több ilyen szám van, akkor m az intervallum közepe. □
A *medián* tehát lényegében azon x érték, aki a nagyság szerinti sorban *középen áll*.

7. Nevezetes diszkrét eloszlású valószínűségi változók

Az alábbikban Im(ξ)={x_1,...,x_n} vagy Im(ξ)={x_1,...,x_n,...} és használjuk a p_k := P(ξ=x_k) rövidítést ahol $k \in$N tetszőleges természetes szám.

Diszkrét egyenletes eloszlású valószínűségi változók

7.1. Definíció: ξ **diszkrét egyenletes v.v.**, ha Im(ξ)={x_1,...,x_n} egy tetszőleges véges halmaz és ξ mindegyik x_i értéket ugyanakkora valószínűséggel veszi fel, azaz p_i =P(ξ=x_i) = 1/n minden i=1,...,n esetén. □

7.2. Például szabályos dobókockával dobunk, a fizetendő összeg utolsó számjegye, véletlenül mondunk egy számot 1 és n között, a dobozban levő n db (különböző) tombolajegyek számai {x_1,...,x_n}, szabályos sokszög alapú egyenes hasáb (majdnem henger) alakú ceruza oldalaira különböző számokat írunk és a ceruzát elgurítjuk, stb.

7.3. Állítás: $M(\xi) = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{x_1 + x_2 + ... + x_n}{n}$ és $D(\xi) = \sqrt{M(\xi^2) - M^2(\xi)} = \sqrt{\dfrac{\sum_{i=1}^{n} x_i^{\,2}}{n} - \left(\dfrac{\sum_{i=1}^{n} x_i}{n}\right)^2}$

(Ez lényegében csak a definíció felírása, tehát ellenőrzése nem nehéz házi feladat. □

Hipergeometrikus eloszlású valószínűségi változók

7.4. Definíció: Legyenek $N,S,n \in \mathrm{N}$ rögzített természetes számok, $N \geq 2$, $0 \leq S \leq N$, $1 \leq n \leq S$, $0 \leq N - S \leq n$. Ekkor ξ hipergeometrikus eloszlású v.v. az S,N,n *paraméterekkel*, ha ξ lehetséges értékei $\mathrm{Im}(\xi) = \{0,1,\ldots,n\}$ és

$$p_k = P(\xi = k) = \frac{\binom{S}{k}\binom{N-S}{n-k}}{\binom{N}{n}} .$$ □

ahol $\binom{\ldots}{\ldots}$ természetesen binomiális együtthatókat jelöl.

7.5. Állítás: Az alábbi típusú kísérletek, melyeket **visszatevés nélküli mintavételek**nek nevezünk, hipergeometrikus eloszlásúak. Tegyük fel, hogy egy halmazban N db elem van, amelyek között S db másmilyen, mint a többi (például selejt). Visszatevés nélkül, akár egy marokkal kiveszünk közülük n -et. Legyen ξ a kivett elemek között a másmilyenek (S-beliek) száma. Ekkor ξ hipergeometrikus v.v.

Az állítás belátása nem nehéz, házi feladat. □

7.6. Példák:

7.7. Tétel: $M(\xi) = n \cdot \frac{S}{N}$ $(= np)$ és $D(\xi) = \sqrt{np(1-p)\left(1 - \frac{n-1}{N-1}\right)}$ ahol $p = \frac{S}{N}$.

(A $\sqrt{1 - \frac{n-1}{N-1}}$ szorzótényezőt **korrekciós tényező**nek hívjuk.)

7.8. Tétel: Ha n és k állandó, $S \to \infty$, $N \to \infty$, $\frac{S}{N} \to p = \textit{állandó}$, akkor

$$\frac{\binom{S}{k}\binom{N-S}{n-k}}{\binom{N}{n}} \to \binom{n}{k} p^k (1-p)^{n-k}$$ □

7.9. Magyarázat: A fenti tétel szerint a hipergeometriai eloszlású v.v. hoz tartozó valószínűségek határértékben a binomiális eloszláshoz (lásd köv. fejezet) tartozó valószínűségekhez tartanak. Ez összhangban azzal a tapasztalati ténnyel és elméleti meggondolással, hogy nagy elemű sokaság (alaphalmaz) esetén majdnem mindegy, hogy visszatevéssel vagy visszatevés nélkül választunk ki elemeket.

Binomiális vagy Bernoulli eloszlású valószínűségi változók

7.10. Definíció ξ **binomiális** vagy **Bernoulli** eloszlású v.v. az n,p *paraméterekkel*, ha $n \in \mathrm{N}$ tetszőleges nemnulla természetes szám és $0<p<1$ tetszőleges szám, továbbá ha ξ lehetséges értékei (!)
$\mathrm{Im}(\xi) = \{x_0, x_1, \ldots, x_n\} = \{0,1,2\ldots,n\}$ és *eloszlása*: tetszőleges k=0,1,2…,n esetén

$$p_k = \mathrm{P}(\xi = \mathrm{k}) = \binom{\mathrm{n}}{\mathrm{k}} p^k (1-p)^{n-k}$$

ahol $\binom{\mathrm{n}}{\mathrm{k}}$ a binomiális együttható. □

7.11. Tétel: A következő típusú kísérletek, melyeket **visszatevéses mintavételek**nek nevezünk, binomiális eloszlásúak. Rögzítsünk egy kísérletet (Ω) és egy $A\subseteq\Omega$ eseményt, rögzítsünk egy tetszőleges $n\in N$, $n\neq 0$ tetszőleges természetes számot, és legyen p=P(A). Most végezzük el a kísérletet n -szer egymás után, teljesen azonos körülmények között, egymástól függetlenül, és jelölje (számolja) ξ azt, hogy az n kísérlet alatt hányszor következett be az A esemény. Ekkor ξ binomiális eloszlású v.v. az n,p paraméterekkel. Ennek belátása nem nehéz, házi feladat. □

Hangsúlyozzuk, hogy a fenti tételben említett feltételek ("teljesen azonos körülmények között, egymástól függetlenül") nem mindig teljesülnek maradéktalanul a gyakorlatban, az ilyen kísérletsorozatok tehát nem teljesen binomiális eloszlású v.v. -k.

7.12. Példák:

a) 10-szer gurítunk egy kockával, ξ a dobott hatosok száma, ekkor ξ binomiális eloszlású az n=10 és p=1/6 paraméterekkel. Azonos feladat: egyszerre gurítunk 10 szabályos kockát, amik nem akadályozzák (befolyásolják) egymást.

b) 7-szer feldobunk egy szabályos érmét, ξ a dobott fejek száma. ξ binomiális eloszlású az n=7, p=0.5 paraméterekkel.

c) *Visszatevéses mintavételek:* visszatevéssel (=egyesével) választunk rögzített N elem közül n-szer. Az összes elem között S db selejtes van. Legyen ξ a kivett elemek közt a selejtes elemek száma. Ekkor ξ binomiális eloszlású az n és a p=S/N paraméterekkel.

d) *Visszatevéssel* választunk a magyar kártyából 5-ször, ξ a kivett lapok közt a pirosak száma. Ekkor ξ binomiális eloszlású az n=5 és p=8/32 paraméterekkel.

7.13. Állítás: Ha ξ binomiális v.v., akkor $M(\xi)=np$ és $D(\xi)=\sqrt{np(1-p)}$. □

7.14. Állítás: Ha ξ binomiális v.v., akkor ξ legvalószínűbb értéke (*modusza*) :

$m = [(n+1)p]$ akkor, ha $(n+1)p$ nem egész szám,

m = (n+1)p és m = (n+1)p-1, akkor, ha $(n+1)p$ egész szám,

ahol [x] az x valós szám **egész részét** (lefelé kerekítés vagy csonkítás) jelenti. □

7.15. Tétel: Ha ξ_1 binomiális eloszlású val. változó az n_1 és p paraméterekkel, ξ_2 binomiális eloszlású val. változó az n_2 és (ugyanazon!) p paraméterekkel, valamint ξ_1 és ξ_2 függetlenek, akkor $\xi_1+\xi_2$ is binomiális eloszlású valószínűségi változó az $n=n_1+n_2$ és p paraméterekkel. □

7.16. Tétel: Ha $n\to\infty$ és $p\to 0$ úgy, hogy $np=\lambda=$ állandó, akkor

$$\binom{n}{k}p^k(1-p)^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda} .$$ □

A legutolsó eredmény, többek között, segítségünkre van p_k kiszámolásában. Ugyanis nagy n , k és kis p esetén a binomiális együttható nagyon nagy míg p^k nagyon kicsi, és ezek szorzata nagy pontatlanságot okoz a számológépeken. A Poisson-eloszlások alapja a fenti 7.16.Tétel.

Poisson eloszlású valószínűségi változók

7.17. Definíció: A ξ v.v. **Poisson** eloszlású a λ >0 valós *paraméterrel*, ha lehetséges értékei Im(ξ) = N = {0,1,2,…,n,...} az *összes* természetes szám, és eloszlása ($e\approx$ 2,71828 az Euler-féle szám)

$$p_k=P(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda} .$$ □

7.18. Alkalmazás: A Poisson eloszlások egyik alkalmazása, mint az előző fejezetben említettük, a binomiális eloszlások közelítése nagy n és kis p esetén. Erre egy mintapéldát a 10. fejezet végén mutatunk.

A gyakorlatban a következő típusú kísérleteknek van Poisson eloszlása: *rögzítünk* egy "fizikai halmazt", ami lehet térbeli (vagy sík- vagy egyenesbeli) halmaz vagy időintervallum, ezt a fizikai halmazt a továbbiakban egységesen "**V térfogatnak**" hívjuk. Tegyük még fel, hogy ezen a halmazon belül sok, egymástól független jelenség léphet fel, melyek egyszerre történő megjelenése négyzetesen csökken. Ekkor bebizonyított tétel, hogy ξ, a fellépett jelenségek száma, Poisson v.v.

Ebből a megfogalmazásból ugye érezhető a Poisson és binomiális eloszlások hasonlósága, amit a 7.16. Tétel igazol. □

7.19. Példák: "*egységnyi*" területen/hosszon/térfogatban keletkezett anyaghibák száma (textil/cső/agyag), adott időintervallumban a hívások / ügyfelek száma, sajtóhibák száma egy könyvben.

7.20. Tétel: Ha ξ Poisson v.v. $\lambda>0$ paraméterrel, akkor $M(\xi)=\lambda$ és $D(\xi)=\sqrt{\lambda}$. □

7.21. Tétel: ξ legvalószínűbb értéke (*modusz*):

$[\lambda]$ ha λ nem egész szám ,

λ és $\lambda-1$ ha λ egész szám . □

7.22. Tétel: Ha ξ_1 Poisson eloszlású val. változó a λ_1 paraméterrel, ξ_2 Poisson eloszlású val. változó a λ_2 paraméterrel, valamint ξ_1 és ξ_2 *függetlenek*, akkor $\xi_1+\xi_2$ is Poisson eloszlású valószínűségi változó az $\lambda=\lambda_1+\lambda_2$ paraméterrel.

Következmény: Ha ξ Poisson eloszlású a V térfogaton λ paraméterrel és $t\in R^+$ tetszőleges pozitív szám, akkor ξ leszűkítése/kiterjesztése a V/t térfogatra szintén Poisson eloszlású λ/t paraméterrel.
Ezt a jelenséget röviden úgy hívjuk, hogy a Poisson eloszlások **korlátlanul oszthatók.**

Geometriai eloszlású valószínűségi változók

7.23. Definíció: ξ **geometriai eloszlású** v.v. a p $(0<p<1)$ paraméterrel, ha ξ lehetséges értékei $\mathrm{Im}(\xi)=N=\{1,2,\dots,n,\dots\}$ az összes pozitív természetes szám, és ξ eloszlása:

$$p_k=P(\xi=k)=p\cdot(1-p)^{k-1} \ . \qquad \square$$

Megjegyzés: szokás a $q:=1-p$ rövidítést használni, ekkor p_k képlete: $p_k=p\cdot q^{k-1}$. Ezért hívják ξ -t mértani (geometriai) eloszlásnak, hiszen ekkor a p_1 , p_2 , ... egy mértani sorozat.

7.24. Tétel: A gyakorlatban a következő kísérletek mértani eloszlású v.v.-k. Egymástól függetlenül végezzük ugyanazt a kísérletet, és legyen A egy rögzített esemény, amelynek valószínűségét jelölje $p:=P(A)$ és legyen $0<p<1$. Ha a kísérletet addig ismételgetjük, teljesen azonos körülmények között, egymástól függetlenül, amíg az A be nem következik, és ξ azt a legkisebb természetes számot jelöli, ahányadik kísérletnél A legelőször bekövetkezett ("*addig, amíg*"), akkor ξ eloszlása geometriai, a p paraméterrel. Ennek belátása nem nehéz, házi feladat.

7.25. Példák: addig jár a korsó a kútra, amíg el nem törik,
addig járunk vizsgázni, amíg végre sikerül a vizsga,
addig ütünk a dióra kalapáccsal amíg meg nem reped,
addig lövünk az ellenségre amíg el nem találjuk,
addig ugrunk árkot, amíg végre száraz ruhával megússzuk,
addig vonulnak el a háremhölgyek Szindbád előtt, amíg Szindbád nem választ, ...

Megjegyezzük, hogy a valódi, gyakorlati életben nem mindegyik fenti példában illetve nem mindig teljesülnek maradéktalanul a feltételek ("teljesen azonos körülmények között, egymástól függetlenül"). Gondoljunk csak a vizsgákra vagy a diókra: a következő vizsgára már jobban készülünk, vagy éppen idegesebbek vagyunk, jobban izguluk vagy puskázunk, no és ki hallott már a 18 645 -dik vizsgáról ?

Ide kapcsolódik még a 8.14. Tétel is. □

7.26. Tétel: Ha ξ geometriai eloszlású v.v. p pararméterrel, akkor

$$M(\xi)=\frac{1}{p} \quad \text{és} \quad D(\xi)=\frac{\sqrt{1-p}}{p} \ . \qquad \square$$

Gyakorlati feladatoknál gyakori a "*hány kísérlet lesz elég*?" kérdés:

7.27. Állítás: $P(\xi \le k) = 1-(1-p)^k$.

Bizonyítás: (ezt a számolást feladatnak is kitűzhetnénk, tehát kivételesen hasznos lesz átolvasnunk):

$P(\xi\le k) = P(\xi=1) + P(\xi=2) + ... + P(\xi=k) = p + p\cdot q^1 + p\cdot q^2 + ... + p\cdot q^{k-1} = p\cdot\frac{q^k-1}{q-1} = p\cdot\frac{q^k-1}{-p} = 1-q^k$. □

7.28. Tétel: Ha ξ geometriai eloszlású valószínűségi változó, akkor

$$P(\xi > m+n \mid \xi > m) = P(\xi > n) \ ,$$

vagyis, az eddigi kísérletek nem befolyásolják a továbbiakat (" ξ **nem fiatalodó**").

Bizonyítás: $P(\xi>k)=\sum_{i=k+1}^{\infty}P(\xi=i)=\sum_{i=k+1}^{\infty}p(1-p)^{i-1}=p(1-p)^k\frac{1}{1-(1-p)}=(1-p)^k$. $m+n>m$ miatt

$P(\xi>m+n\cap\xi>m)=P(\xi>m+n)=(1-p)^{m+n}$, így $P(\xi>m+n\mid\xi>m)=\frac{P(\xi>m+n)}{P(\xi>m)}=(1-p)^n$. □

8. Nevezetes folytonos eloszlású valószínűségi változók

Folytonos egyenletes eloszlású valószínűségi változók

8.1. Definíció Tetszőleges rögzített $a,b\in$R valós számok, $a<b$ esetén ξ **folytonos egyenletes eloszlású** v.v. az *a,b paraméterekkel*, ha sűrűségfüggvénye $f(x)=\begin{cases} c & ha \quad a\le x\le b \\ 0 & különben \end{cases}$. □

Ebben a fejezetben *nagyon egyszerűen* bizonyítható (kiszámolható) állítások vannak, a számolások elvégzése mindegyik esetben hasznos házi feladat.

8.2. Állítás: c$=\frac{1}{b-a}$ és $F(x)=\begin{cases} 0 \quad ha \quad x\le a \\ \frac{x-a}{b-a} \quad ha \quad a\le x\le b \\ 1 \quad ha \quad x\ge b \end{cases}$. □

8.3. Állítás: $M(\xi)=\frac{a+b}{2}$ és $D(\xi)=\frac{b-a}{\sqrt{12}}$. □

8.4. Tétel: A gyakorlatban a következő kísérletek *folytonos egyenletes eloszlású* v.v.-k: véletlenszerűen, minden befolyástól mentesen, "*egyenletesen*" választunk az [*a,b*] intervallumban egy *x* valós számot. □

A fenti és az alábbi állítások is igazolják a folytonos egyenletes eloszlások és a geometriai valószínűség (3.fejezet második része) azonosságát!

8.5. Példák: henger alakú ceruzát elgurítva mely pontján áll meg,
papírszalagot hol vágunk ketté, pálcát hol törünk ketté,
buszmegállóba megyünk ki "csak úgy", a buszok egyenletesen 15 percenként jönnek, stb.

8.6. Állítás: $P(c<\xi<d)=\dfrac{d-c}{b-a}$ arányos a [*c,d*] intervallum hosszával.

Exponenciális eloszlású valószínűségi változók

8.7. Definíció Tetszőleges rögzített $\lambda\in R^+$ pozitív valós szám esetén ξ **exponenciális eloszlású** v.v. a $\lambda>0$ paraméterrel, ha sűrűségfüggvénye

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & ha \quad x\geq 0\\ 0 \quad különben & \end{cases}.$$ □

Megjegyzés: A gyakorlatban a következő kísérletek *exponenciális eloszlású* v.v.-k: gépek, élettelen tárgyak (pl. izzók, gumiabroncsok, stb.) élettartama, azaz a tönkremenésig eltelt idő. Ez nem egy elméleti tétel, hanem gyakorlati tapasztalat, pontosabban statisztikai módszerekkel ("*illeszkedésvizsgálat*") ellenőrzött tény. □

8.8. Tétel: Ha ξ exponenciális eloszlású v.v. a λ paraméterrel, akkor $F(x)=\begin{cases}0 & ha \quad x<0\\ 1-e^{-\lambda x} & ha \quad x\geq 0\end{cases}.$ □

8.9. Tétel: $M(\xi)=D(\xi)=\dfrac{1}{\lambda}$. □

8.10. Tétel: Ha ξ exponenciális eloszlású val. változó, és $x\geq 0$, $y\geq 0$ tetszőleges számok, akkor

$$P(\xi\geq x+y|\xi\geq y)=P(\xi\geq x) \ .$$ (" ξ *örökifjú, nem öregedő*") □

A fenti egyenlőség kiszámolása nem nehéz, javasolt házi feladat. Érdemes még összehasonlítani a 7.28. Tétellel is. A következő tétel kiemeli az örökifjú tulajdonság és az exponenciális eloszlás kapcsolatát.

8.11. Tétel: Ha ξ folytonos eloszlású, F(0)=0, F(*x*)<1 , F minden nemnegatív *x* -re deriválható, $\lim\limits_{0+} F'(x)=\lambda>0$ és ξ örökifjú, akkor ξ exponenciális eloszlású v.v. □

8.12. Tétel: Az exponenciális eloszlás és a *Poisson* eloszlás kapcsolata:
Legyenek $\xi_1,\xi_2,....$ független exponenciális eloszlású v.v. azonos λ paraméterrel, legyen T>0 rögzített, és legyen η a T "időpontig tönkrement/kicserélt alkatrészek" száma, azaz legyen Im(η)=N, és legyen:

$\eta=0$ ha $\xi_1>T$, $\eta=1$ ha $\xi_1\leq T$ de $\xi_1+\xi_2>T$,

általában

$$\eta=k \quad \text{ha} \quad \sum_{i=1}^{k}\xi_i\leq T \quad \text{de} \quad \sum_{i=1}^{k+1}\xi_i>T \ .$$

Ekkor η Poisson eloszlású v.v. a $\lambda\cdot T$ paraméterrel. □

8.13. Példa: Egy bizonyos típusú izzó élettartama exponenciális eloszlású valószínűségi változó (ξ) 1000 óra várható értékkel (M(ξ)). Ha egy izzó tönkremegy, azonnal kicseréljük egy ugyanolyan típusú másik izzóra, aminek az élettartama független az előző izzó élettartamától. Mennyi a valószínűsége annak, hogy 2500 óra alatt 2 izzócsere szükséges? (Azaz P(η≥2)=? .)

8.14. Tétel: az exponenciális eloszlás és a *geometriai* eloszlás kapcsolata:
Ha ξ *exponenciális eloszlás*ú v.v. a λ paraméterrel, akkor $\eta=[\xi]+1$ geometriai eloszlású v.v. a $p=1-e^{-\lambda}$ paraméterrel.
Bizonyítás: η értékei lehetnek 1,2,3,... , és

$$P(\eta=k)=P(k-1\le\xi<k)=F(k+1)-F(k)=1-e^{-\lambda k}-(1-e^{-\lambda(k-1)})=e^{-\lambda(k-1)}(1-e^{-\lambda}) \ . \qquad \square$$

Magyarázat: η jelentése: hány (egész) óráig működik ξ , meddig kell még várnom, hogy tönkremenjen, vagyis "egy-egy kockadobás" = "1 óráig működtetem" . □

8.15. Példa: Egy telefontársaság folytonos alapon számláz, 20 Ft/perc díjjal. Egy másik társaság perc alapon számláz 15 Ft/ perc illetve percdíjjal. Melyik számlázási módot érdemes választani, ha egy hívás hossza exponenciális eloszlású valószínűségi változó 2 perc várható értékkel?
Megoldás: Legyen ξ a hívás hossza, θ_1 a folytonos alapon számlázott díj, θ_2 a perc alapon számlázott díj. Ekkor

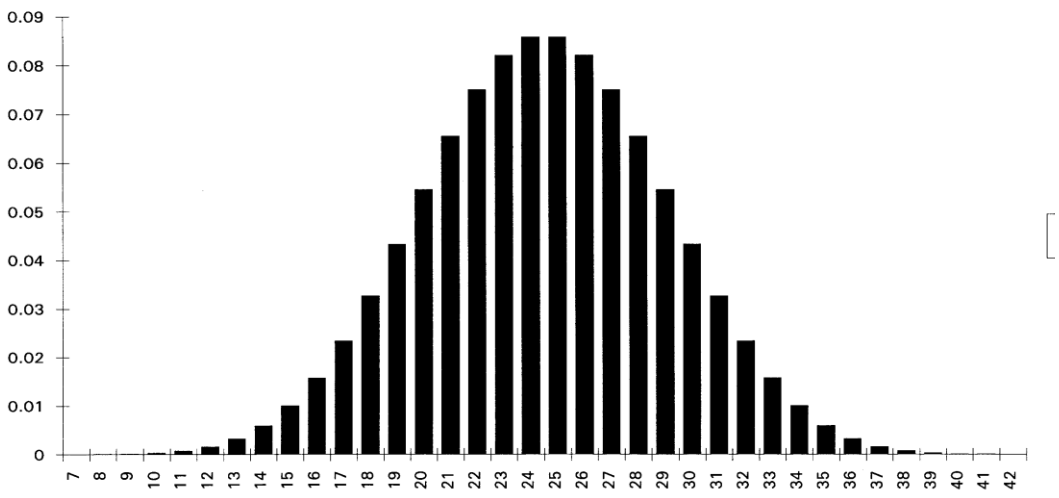
$$M(\theta_1)=20M(\xi)=40 \text{ Ft} \quad \text{és} \quad M(\theta_2)=15M([\xi]+1)=15\frac{1}{1-e^{-0.5}}=38.1 \text{ Ft.}$$

Tehát, ha 17 Ft a percdíj, akkor már 40 Ft felé kerül a várható érték. □

9. Normális eloszlású valószínűségi változók

9.1. Bevezetés: Nagyon sok fizikai és egyéb mennyiség nagyon sok egyednél történt mérésének tanulmányozása után jutott Gauss arra a következtetésre, hogy a mennyiségek hisztogramjait (oszlopdiagramjait) jól közelítik az e^{-x^2} függvény lineáris ("vízszintes és függőleges") transzformációi.

Általában olyan mennyiségekről van szó, amelyek nagyon sok, apró +/- hatás összegeződéseként jönnek létre (testmagasság, tömeg, térfogat, feszültség, stb.) Ezen feltételezés szemléltetésére sok példát láthatunk: Galton deszka, több kocka dobásainak összege (lásd a honlapomon és az alábbi rajzon), stb. Ezt a jelenséget a 10.5. "Központi Határeloszlás Tétel" és a Statisztika tárgy "Illeszkedésvizsgálat" módszere igazolja.



Hét kocka összegének eloszlása

http://math.uni-pannon.hu/~szalkai/7kocka.gif

9.2. Definíció: ξ **standard normális eloszlású v.v.**, ha sűrűségfüggvénye $f(x)=\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}$ ($x\in$R). □

9.3. Tétel: $M(\xi)=0$ és $D(\xi)=1$. □

9.4. Megjegyzések: Érdemes alaposan tanulmányoznunk az e^{-x^2} függvény és lineáris transzformáltjainak képleteit és grafikonjait!

Liouville tétele szerint az e^{-x^2} függvény primitív függvényét nem lehet képlettel felírni. Ezért $F(x)$ -re képletet nem tudunk felírni, értékeit külön-külön kiszámolva táblázatba rendezték a matematikusok, mi pedig az értékeket a táblázatból keressük ki. Egy kisebb táblázat a jegyzet végén található.

9.5. Jelölés: $\Phi(x):=F(x)=\int_{-\infty}^{x}f(t)dt$ tetszőleges $x\in$R esetén. □

$\Phi(x)$ nyilván a **standard normális eloszlású v.v. eloszlásfüggvénye.**

9.6. Állítás: $\Phi(-x)=1-\Phi(x)$ tetszőleges $x\in$R esetén. □

Következmény: $-\xi$ is standard normális eloszlású v.v.

Bizonyítás: $F(x)=P(-\xi<x)=P(-x<\xi)=1-\Phi(-x)=1-(1-\Phi(x))=\Phi(x)$. □

9.7. Definíció: Legyen ξ standard normális eloszlású v.v. és legyenek $m,\sigma\in$R tetszőleges valós számok, σ>0 . Ekkor az $\eta=\sigma\cdot\xi+m$ valószínűségi változót m,σ ***paraméterű*** **normális eloszlású v.v.** -nak nevezzük és erre a $\eta\sim N(m,\sigma)$ *jelölést* használjuk ! □

Az alábbi tétel ekvivalens definíciót jelent a normális eloszlások részére:

9.8. Tétel: η sűrűségfüggvénye $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-m)^2}{2\sigma^2}}$ minden valós x -re. □

9.9. Állítás: $M(\eta)=m$ és $D(\eta)=\sigma$. □

9.10. Tétel: $\eta\sim N(m,\sigma)$ esetén $F_\eta(x)=F_{m,\sigma}(x)=\Phi\left(\frac{x-m}{\sigma}\right)$. □

9.11. Tétel: "*k-szor szigma szabály*" **:**

Ha $\xi\sim N(m,\sigma)$ és $k\in$R$^+$ tetszőleges pozitív szám, akkor $P(m-k\sigma<\xi<m+k\sigma)=2\Phi(k)-1$. □

Jelentése: Annak a valószínűsége, hogy ξ értéke a várható értékének *k·szórás*a sugarú környezetébe esik, pontosan 2·Φ(*k*)-1 .

Az egyenlőség könnyen kiszámolható a 9.10. és 9.6. összefüggések alapján, tanulságos házi feladat. □

A tétel speciális eseteit érdemes külön is kiszámolnunk:

9.12. Tétel: Speciális esetek:

$k=1 \quad\Rightarrow\quad P(m-\sigma<\xi<m+\sigma)=0.68$,

$k=2 \quad\Rightarrow\quad P(m-2\sigma<\xi<m+2\sigma)=0.95$,

$k=3 \quad\Rightarrow\quad P(m-3\sigma<\xi<m+3\sigma)=0.997$. □

9.13. Tétel: Ha $\xi\sim N(m,\sigma)$ és $\eta=a\xi+b$, $a\neq 0$, akkor $\eta\sim N(am+b\,,|a|\sigma)$. □

9.14. Tétel: Ha $\xi\sim N(m_1,\sigma_1)$ és $\eta\sim N(m_2,\sigma_2)$ független valószínűségi változók, akkor

$$\xi+\eta\sim N(m_1+m_2,\sqrt{\sigma_1^2+\sigma_2^2}).$$ □

Megjegyzés: Az általános 6.4. és 6.10. Tételekből már tudjuk, hogy $\xi+\eta$ várható értéke m_1+m_2 és szórása $\sqrt{\sigma_1^2+\sigma_2^2}$, a Tétel újdonsága az, hogy: "normális eloszlások összege is normális eloszlás" ! □

Az alábbi, gyakorlatban is fontos összefüggés már könnyen következik az előző Tételből:

9.15. Tétel: Ha $\xi_1, \xi_2, \ldots. \xi_n \sim N(m, \sigma)$ független, azonos eloszlású normális v.v.-k, akkor

$$\sum_{i=1}^{n} \xi_i \sim N(nm, \sigma\sqrt{n}) \quad \text{és} \quad \frac{\sum_{i=1}^{n} \xi_i}{n} \sim N\left(m, \frac{\sigma}{\sqrt{n}}\right) .$$ □

Megjegyzés: Mint már többször említettük, a gyakorlatban egy mérés helyett többször mérünk és átlagolunk, ezáltal a szórás csökken, a pontosság nő. Jó tudnunk azt is, hogy az öszeg és az átlag is *normális eloszlású* v.v. □

9.16. Példa: egy felnőtt tömege normális eloszlású valószínűségi változó 75 kg várható értékkel és 15 kg szórással , egy iskolás gyerek tömege normális eloszlású 35 kg várható értékkel és 6 kg szórással. Ha a két személy tömegét független valószínűségi változónak tekintjük, akkor

a) mekkora valószínűséggel lesz egy felnőtt tömege nagyobb, mint egy gyerek tömege,
b) mennyi a valószínűsége annak, hogy az össztömegük 80 és 140 kg közé esik?

9.17. Példa: Egy liftet 8 felnőtt személyre méreteznek. A beszállók tömegét független normális eloszlású valószínűségi változónak tekintjük 75 kg várható értékkel és 15 kg szórással. Mennyi legyen a lift teherbíró képessége, ha azt szeretnénk, hogy 4 személy beszállása esetén 0.99 valószínűséggel ne gyulladjon ki a túlterheltséget jelző lámpa?

Megoldás: Tehát $\xi_i \sim N(75,15)$, $i=1,\ldots,8$, $m=75$, $\sigma=15$, és legyen $\eta := \sum_{i=1}^{8} \xi_i$.

Ekkor, a 9.15. Tétel szerint $\eta \sim N(8\cdot m, \sqrt{8}\cdot\sigma) = N(8\cdot 75, \sqrt{8}\cdot 15) = N(600{,}42.43)$.

Olyan $x \in \mathrm{R}$ valós számot kell keresnünk, amelyre $P\left(\sum_{i=1}^{8} \xi_i < x\right) = P(\eta<x) = 0.99$.

Mivel $P(\eta<x) = F_\eta(x) = \Phi\left(\frac{x-600}{42.43}\right) = 0.99$ és Φ táblázatából tudjuk, hogy $\Phi(2.32)=0.99$, tehát

$\frac{x-600}{42.43} = 2.32$ ahonnan $x=698.5 \sim 700$kg . □

A következő összefüggésekre a Matematikai Statisztika tárgyban lesz majd szükségünk.

9.18. Tétel: Ha $\xi \sim N(0,1)$ akkor $\eta = \xi^2$ eloszlás- és sűrűségfüggvénye

$$F_\eta(x) = \begin{cases} 0 & ha \quad x \le 0 \\ 2\Phi(\sqrt{x}) - 1 & ha \quad x > 0 \end{cases} \quad \text{és} \quad f_\eta(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}} & ha\ x > 0 \\ 0 & különben \end{cases} ,$$

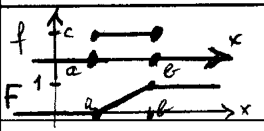
valamint várható értéke $M(\eta)=1$. □

9.19. Tétel: Ha $\xi \sim N(0,1)$ akkor $\eta = e^{\xi}$ (ún. "**lognormális eloszlás**") eloszlás- és sűrűségfüggvénye

$$F_\eta(x) = \begin{cases} 0 & ha \quad x \le 0 \\ \Phi(\ln x) & ha \quad x > 0 \end{cases} \quad \text{és} \quad f_\eta(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{(\ln x)^2}{2}} & ha\ x > 0 \\ 0 & különben \end{cases} ,$$

valamint várható értéke $M(\eta) = \sqrt{e}$. □

A következő táblázat nagy méretben megtalálható a honlapomon:
http://math.uni-pannon.hu/~szalkai/Eloszlasok(pdf)+kezjav+.gif

	Név	Paraméter	Definíció	$M=\mu$	σ^2	Alkalmazás
	Nevezetes eloszlások					
diszkrét eloszlások	Indikátorváltozó	p	$p_1 = p,\ p_0 = 1-p$	p	$p(1-p)$	Egy p valószínűségű esemény bekövetkezik-e?
	Binomiális eloszlás	n, p	$p_k = \binom{n}{k} p^k (1-p)^{n-k}\ (k=0,1,\ldots,n)$	np	$np(1-p)$	Visszatevéses mintavétel. Egy p valószínűségű esemény n független ismétlés során hányszor következik be?
	Geometriai eloszlás	p	$p_k = (1-p)^{k-1} \cdot p$	$\frac{1}{p}$	$\frac{(1-p)}{p^2}$	
	Hipergeometrikus eloszlás	N, K, n	$p_k = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}\ (k=0,1,\ldots,n)$	$\frac{Kn}{N}$ $= n \cdot p$	$\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}$ $= n \cdot p \cdot q \left(\frac{N-n}{N-1}\right)$	Visszatevés nélküli mintavétel. *Megjegyzés:* A $p = \frac{K}{N}$ jelöléssel μ és σ egyszerűbb alakra írható.
	Poisson-eloszlás	$\lambda (>0)$	$p_k = \frac{\lambda^k}{k!} e^{-\lambda}\ (k=0,1,\ldots)$	λ	λ	Egységnyi idő alatt megfigyelt események száma. Egységnyi területre eső hibák száma.
folytonos eloszlások	Standard normális eloszlás	—	$f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$	0	1	Normális eloszlás standardizáltja.
	Normális eloszlás	$\mu,\ \sigma(>0)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2	Populáció egyedeinek méretei, tömegei.
	Egyenletes eloszlás	a, b	 $c = \frac{1}{b-a}$ $F(x) = \frac{x-a}{b-a}$	$\frac{a+b}{2}$	$\left(\frac{b-a}{\sqrt{12}}\right)^2$	Gyártási folyamatban fellépő méretingadozások. Fizikai mennyiségek.
	Exponenciális eloszlás	$\lambda(>0)$	$f(x) = \lambda e^{-\lambda x}$ ha $x>0$; $F(x) = 1 - e^{-\lambda x}$ ha $x>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	Élettartam, várakozási idő. Időintervallumok a Poisson-eloszláshoz kapcsolódóan.

10. Nagy számok törvényei

Az alábbi egyenlőtlenségek általános közelítéseket adnak bármely η valószínűségi változó értékeinek eloszlásáról. "Természetesen" a közelítések pontosságát a mérések számának növelésével javíthatjuk.

10.1. Tétel: Markov-egyenlőtlenség

Ha a ξ *valószínűségi változónak létezik* $M(\xi)$ *várható értéke, akkor* $\forall a\in R^+$ *esetén*

$$\boxed{P(\xi \ge a) \;\le\; \frac{M(\xi)}{a}}\;.$$

Magyarázat: A tétel szerint a mérés eredménye (ξ) egyre nagyobb értéket egyre kisebb valószínűséggel vesz fel. Pontosabban: ξ bármilyen nagy lehet ($\xi\ge a$), de ennek az eseménynek a valószínűsége legfeljebb M/a, amely korlát $a\to\infty$ esetén 0 -hoz tart.

10.2. Tétel: Csebisev-egyenlőtlenség

Ha a ξ *valószínűségi változónak létezik* $M(\xi)$ *várható értéke és* $D(\xi)$ *szórása, akkor* $\forall k\in R^+$ *és* $\forall \varepsilon\in R^+$ *esetén*

$$\boxed{P\big(\,|\xi - M(\xi)| \ge kD(\xi)\,\big) \;\le\; \frac{1}{k^2}}, \quad \textit{vagy másként} \quad \boxed{P\big(\,|\xi - M(\xi)| \ge \varepsilon\,\big) \;\le\; \frac{D^2(\xi)}{\varepsilon^2}}\;,$$

ill. az esemény tagadása:

$$\boxed{P\big(\,|\xi - M(\xi)| \le \varepsilon\,\big) \;\ge 1 - \frac{D^2(\xi)}{\varepsilon^2}}\;.$$

Magyarázat: Azt várjuk, hogy a mérés (ξ) értékei az átlag **(**$M=M(\xi)$**)** körül ingadoznak, ezt igazolja a Csebisev-egyenlőtlenség: a mérés és az átlag *eltérése* (vagyis $|\xi-M|$) lehet ugyan *nagy*, vagyis *lehetséges*, hogy $|\xi-M| \ge \varepsilon$, de ennek a valószínűsége legfeljebb D^2/ε^2. Ez a korlát pedig 0 -hoz tart amennyiben $k\to\infty$. Természetesen ez a korlát is kisebb ha a D szórás is kicsi.

10.3. Tétel: Bernoulli-féle nagy számok törvénye

Legyen A *egy tetszőlege esemény amelyre* $P(A)=p$. *Végezzünk* n *(független) kísérletet és jelölje* ξ_n *az A esemény gyakoriságát (hányszor "sikerült"* A*), ekkor a* relatív gyakoriság = ξ_n/n. *Ekkor tetszőleges* $\varepsilon>0$ *és* $n\in N$ *esetén* ($q=1-p$):

$$\boxed{P\left(\left|\frac{\xi_n}{n} - p\right| \ge \varepsilon\right) \le \frac{pq}{\varepsilon^2 n}}, \quad \textit{vagyis} \quad \boxed{P\left(\left|\frac{\xi_n}{n} - p\right| < \varepsilon\right) \ge 1 - \frac{pq}{\varepsilon^2 n}}.$$

Átfogalmazva: *tetszőleges* $\varepsilon>0$ *és* $\delta>0$ *esetén* $\exists n_0$ *hogy minden* $n>n_0$ *esetén*

$$\boxed{P\left(\left|\frac{\xi_n}{n} - p\right| < \varepsilon\right) \ge 1 - \delta}, \quad \textit{vagy másként} \quad \boxed{P\left(\left|\frac{\xi_n}{n} - p\right| \ge \varepsilon\right) \le \delta}\;.$$

(ahol $\delta=\dfrac{pq}{\varepsilon^2 n}\to 0$ ha $n\to\infty$.)

Magyarázat: A tétel szerint az *elméleti* valószínűség (p) és a *tapasztalati* (gör: empirikus) relatív gyakoriság (azaz ξ_n/n) *eltérése* (vagyis $|\,\xi_n/n - p\,|$) **kicsi**, sőt tetszőleges ε számnál kisebb lehet - legalábbis majdnem 100% valószínűséggel igaz a mondat előző fele.

Másképpen fogalmazva (jobboldali képlet): az eltérés lehet nagy, de csak kis valószínűséggel. (Ez nem jelenti azt, hogy a relatív gyakoriság konvergál a valószínűséghez, hanem csak azt,hogy a nagy eltérés valószínűsége kicsi!) Ha p nem ismert, akkor a $p(1-p)\le 1/4$ becslés alapján $\delta\le 1/4\varepsilon^2 n$ is írható.

10.4. Tétel: A nagy számok gyenge törvénye (Csebisev-alak)

Legyenek ξ_1 , ξ_2 , ... , ξ_n *független, azonos eloszlású valószínűségi változók, amelyeknek létezik (azonos) várható értékük és szórásuk*: $m:=M(\xi_i)$ *és* $\sigma:=D(\xi_i)$, *és legyen* $S_n:= \xi_1+\xi_2+\ldots+\xi_n$. *Ekkor minden* $\varepsilon\in R^+$ *esetén*

$$\boxed{P\left(\left|\frac{S_n}{n}-m\right|\le\varepsilon\right)\ge 1-\frac{\sigma^2}{n\varepsilon^2}}\ ,\quad \textit{vagy másként}\quad \boxed{P\left(\left|\frac{S_n}{n}-m\right|\ge\varepsilon\right)\le\frac{\sigma^2}{n\varepsilon^2}}\ .$$

Magyarázat: ξ_1 , ξ_2 , ... , ξ_n ugyanazon mennyiség többszöri független mérését jelölik (azonos eloszlásúak és függetlenek), így S_n/n éppen ezen mérések (tapasztalati) **átlaga**. A tétel pedig azt mondja ki, hogy a *tapasztalati* átlag (S_n/n) és az *elméleti* átlag (m) eltérése (azaz $|S_n/n-m|$) mekkora lehet. Például tetszőleges rögzített ε korlát esetén az első képletben szereplő hibatag $\sigma^2/n\varepsilon^2 \to 0$ midőn $n\to\infty$, vagyis: a kísérletek számának növelésével ($n\to\infty$) a tapasztalati és az elméleti átlag eltérés ε -nál kisebb (ill. ezen állítás 1-hez közeli valószínűséggel igaz).

E tételnek speciális esete a Bernoulli-féle törvény, ugyanis a 10.3.Tételben $m=p$ és $\sigma^2=pq$.

10.5. Tétel: Központi (=centrális) határeloszlás tétel (nagy számok erős törvénye)

Legyenek ξ_1 , ξ_2 , … , ξ_n , ... *független, azonos eloszlású valószínűségi változók, amelyeknek létezik (azonos) várható értékük és szórásuk:* $M(\xi_i)=m$ *és* $D(\xi_i)=\sigma$ *(i=1,2…n,...). Ekkor* **összegük** *(pontosabban: standardizált átlaguk) közelítőleg normális eloszlású: a*

$$\varsigma_n=\frac{\xi_1+...+\xi_n-n\cdot m}{\sigma\sqrt{n}}\quad \textit{jelöléssel}\quad \lim_{n\to\infty}P(\varsigma_n<y)=\Phi(y)\ .$$

Magyarázat: Mivel mindegyik ξ_1 , ξ_2 , … , ξ_n ... mérés átlagosan m, így összegük $+\infty$ -be tart: $\lim_{n\to\infty}(\xi_1+\xi_2+\ldots+\xi_n)=+\infty$, ezért kell az összeg ζ_n (standardizált) változatát tekintenünk. Mivel a $P(\zeta_n\le y)$ kifejezés éppen az ζ_n valószínűségi változó eloszlásfüggvénye és Φ a (standard) Normális eloszlás eloszlásfüggvénye, ezért a Tétel utolsó képlete valóban azt igazolja, hogy sok azonos (apró) hatás összeg valóban a Normális eloszláshoz tart.

10.6. Moivre-Laplace tétel:

Tetszőleges $0\le p\le 1$ *és* $u,v\in\mathbb{R}$ *valós számokra*

$$\lim_{n\to\infty}\left(\sum_{u\le k\le v}\binom{n}{k}p^kq^{n-k}\right)=\Phi(v^*)-\Phi(u^*)$$

vagy egyszerűbben

$$\sum_{u\le k\le v}\binom{n}{k}p^kq^{n-k}\approx\Phi(v^*)-\Phi(u^*)$$

ahol

$$u^*=\frac{u-np}{\sqrt{npq}}=\frac{u-m}{\sigma}\quad\text{és}\quad v^*=\frac{v-np}{\sqrt{npq}}=\frac{v-m}{\sigma}\ .$$

Magyarázat: A tétel szerint nagyon sok azonos kísérletet elvégezve az, hogy a sikeres kísérletek száma u és v közé esik, normális eloszlással közelíthető.

Más szavakkal: a <u>binomiális eloszlást</u> nagy n esetén Normális eloszlással közelíthetjük.

Ötlet: A második (közelítő) képlet <u>hasonlít</u> a már "megszokott"

$$P(u<\xi<v)=F(v)-F(u)=\Phi(v^*)-\Phi(u^*)=\Phi\left(\frac{v-m}{\sigma}\right)-\Phi\left(\frac{u-m}{\sigma}\right)$$

formulához! A tételt szokták a következő alakban is írni:

$$\lim_{n\to\infty}\left(\sum_{u^*\le\frac{k-np}{\sqrt{npq}}\le v^*}\binom{n}{k}p^kq^{n-k}\right)=\Phi(v^*)-\Phi(u^*)$$

A Moivre-Laplace tétel a 10.5. Tétel speciális esete

Megjegyzés: Mint tanultuk: nagy n esetén a <u>hipergeometrikus</u> eloszlás közelíthető a <u>binomiális</u> eloszlással, ami pedig a <u>Poisson</u> eloszlással is közelíthető, és végül, a Poisson eloszlás is közelíthető a <u>normális</u> eloszlással.

10.7. Mintafeladat: N=1000 állat van egy farmon. Egy járvány esetén *P*(meggyógyul)=0.7 . Mekkora eséllyel lesz a meggyógyult állatok száma $\frac{N}{3}$ és $\frac{2N}{3}$ között?

Megoldás:

1. rész: A feladat valójában binomiális (Bernoulli) eloszlású:

N=1000 , p=0,7 , q=1-p=0,3 ,

ξ=1000 állat közül mennyi gyógyul meg .

Binomiális eloszlás:

$P(\xi=k)=\binom{n}{k}p^k(1-p)^{n-k}$ (k=0;1;…;n)

$$P(334 \le \xi \le 666) = \sum_{k=334}^{k=666} \binom{1000}{k} 0{,}7^k 0{,}3^{1000-k} \quad .$$

2. rész: Mivel N nagy, ezért a számolási nehézségek miatt közelíthetünk Poisson eloszlással.

Poisson-eloszlás:

$P(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda}$ (*k*=0;1;…)

$\lambda = np = 1000 \cdot 0{,}7 = 700$

$$P(334 \le \xi \le 666) = \sum_{k=334}^{k=666} \frac{700^k}{k!} e^{-700} \quad .$$

3. rész: Mivel N nagyon nagy, ezért a még mindig fennálló számolási nehézségek miatt a Moivre-Laplace tétellel kell közelítenünk és számolnunk.

Moivre - Laplace tételt használata

$\sum_{u \le k \le v} \binom{n}{k} p^k q^{n-k} \approx \Phi(v^*) - \Phi(u^*)$, ahol: $u^* = \frac{u-np}{\sqrt{npq}} = \frac{u-m}{\sigma}$ és $v^* = \frac{v-np}{\sqrt{npq}} = \frac{v-m}{\sigma}$,

M=m=np= 1000*0,7= 700 , D=σ=$\sqrt{npq} = \sqrt{1000 x 0{,}7 x 0{,}3} = \sqrt{210} \approx 14{,}49$,

$P(334 \le \xi \le 666) = \sum_{k=334}^{k=666} \binom{1000}{k} 0{,}7^k 0{,}3^{1000-k} \approx \Phi\left(\frac{666-700}{\sqrt{210}}\right) - \Phi\left(\frac{334-700}{\sqrt{210}}\right) =$

$= \Phi(-2{,}35) - \Phi(-25{,}26) = 1 - \Phi(2{,}35) - 0 = 1 - 0{,}99065 =$ **0,00935 .** □

Függelék

Valószínűségszámítás - Matematika szótár

Valószínűségszámítás	Matematika
Eseménytér (=kísérlet összes lehetséges kimenetele)	$H \neq \emptyset$ (tetszőleges) alap**halmaz**,
kísérlet **végeredménye**	$x \in H$ tetszőleges **elem**,
esemény (=kísérlet aktuális kimenetele)	$A \subseteq H$ (tetszőleges) rész**halmaz**,
elemi esemény	$\{x\} \subseteq H$ egyelemű részhalmaz ("singleton"),
A esemény bekövetkezik	$x \in A$
A esemény nem következik be	$x \notin A$
biztos esemény	$H \subseteq H$ (az alaphalmaz), ill. ha $P(A)=1$,
lehetetlen esemény	$\emptyset \subseteq H$ (üres halmaz), ill. ha $P(A)=0$,
ellentett esemény (esemény tagadása)	$\overline{A}$ (komplementer halmaz),
események összege: A+B ("vagy")	$A \cup B$ (únió),
szorzata: A·B ("és")	$A \cap B$ (metszet),
különbsége: A-B	$A \setminus B$ (különbség),
kizáró események	$A \cap B = \emptyset$ (diszjunkt halmazok), ill. ha $P(A \cap B)=0$,
A -ból következik B (=A maga után vonja B-t)	$A \subseteq B$ (A részhalmaza B -nek),
valószínűség P(A)	$\mathbf{P}:\mathcal{P}(H) \to R$ tetsz. függvény ($\mathcal{P}(H)$=H hatványhalmaza) a Kolmogorov axiómákkal, (pl. terület),
független események A,B	$P(A \cap B)=P(A) \cdot P(B)$,
teljes eseményrendszer	H egy partíciója,
valószínűségi **v**áltozó (mérés számszerű végeredménye)	$\xi : H \to R$ tetszőleges függvény,
diszkrét v.v.	$Im(\xi) = \{x_1, x_2, ..., x_n, ...\}$ felsorolható,
- " - eloszlása	$\{ p_1 , p_2 , ... , p_n , ... \}$ ahol $p_i=P(\xi=x_i)$ ha $i \in N$
folytonos v.v.	van $[a,b] \subseteq Im(\xi)$ intervallum,
eloszlásfüggvény	$F:R \to R$ tetszőleges függvény az axiómákkal, vagy: $F(t) = P(\xi<t)$ vagy: f primitív függvénye: $F(t)=\int_{-\infty}^{t} f(x)dx$
sűrűségfüggvény	$f:R \to R$ tetszőleges függvény az axiómákkal, vagy: F deriváltfüggvénye: $f(x)=F'(x)$
$P(a \leq \xi < b) = \int_a^b f(x)dx = F(b)-F(a)$	Newton-Leibniz szabály
várható érték: $M(\xi)$	átlag (=számtani közép, **mean** (angolul))
szórás: $D(\xi)$	"szóródás" (=**dispersion** (latin))

A standard normális eloszlású változó eloszlásfüggvényének (Φ) táblázata

x	Φ(x)
0,00	0,5000
0,01	0,5040
0,02	0,5080
0,03	0,5120
0,04	0,5160
0,05	0,5199
0,06	0,5239
0,07	0,5279
0,08	0,5319
0,09	0,5359
0,10	0,5398
0,11	0,5438
0,12	0,5478
0,13	0,5517
0,14	0,5557
0,15	0,5596
0,16	0,5636
0,17	0,5675
0,18	0,5714
0,19	0,5753
0,20	0,5793
0,21	0,5832
0,22	0,5871
0,23	0,5910
0,24	0,5948
0,25	0,5987
0,26	0,6026
0,27	0,6064
0,28	0,6103
0,29	0,6141
0,30	0,6179
0,31	0,6217
0,32	0,6255
0,33	0,6293

z	Φ(x)
0,34	0,6331
0,35	0,6368
0,36	0,6406
0,37	0,6443
0,38	0,6480
0,39	0,6517
0,40	0,6554
0,41	0,6591
0,42	0,6628
0,43	0,6664
0,44	0,6700
0,45	0,6736
0,46	0,6772
0,47	0,6808
0,48	0,6844
0,49	0,6879
0,50	0,6915
0,51	0,6950
0,52	0,6985
0,53	0,7019
0,54	0,7054
0,55	0,7088
0,56	0,7123
0,57	0,7157
0,58	0,7190
0,59	0,7224
0,60	0,7257
0,61	0,7291
0,62	0,7324
0,63	0,7357
0,64	0,7389
0,65	0,7422
0,66	0,7454
0,67	0,7486

x	Φ(x)
0,68	0,7517
0,69	0,7549
0,70	0,7580
0,71	0,7611
0,72	0,7642
0,73	0,7673
0,74	0,7704
0,75	0,7734
0,76	0,7764
0,77	0,7794
0,78	0,7823
0,79	0,7852
0,80	0,7881
0,81	0,7910
0,82	0,7939
0,83	0,7967
0,84	0,7995
0,85	0,8023
0,86	0,8051
0,87	0,8078
0,88	0,8106
0,89	0,8133
0,90	0,8159
0,91	0,8186
0,92	0,8212
0,93	0,8238
0,94	0,8264
0,95	0,8289
0,96	0,8315
0,97	0,8340
0,98	0,8365
0,99	0,8389
1,00	0,8413
1,01	0,8438

x	Φ(x)
1,02	0,8461
1,03	0,8485
1,04	0,8508
1,05	0,8531
1,06	0,8554
1,07	0,8577
1,08	0,8599
1,09	0,8621
1,10	0,8643
1,11	0,8665
1,12	0,8686
1,13	0,8708
1,14	0,8729
1,15	0,8749
1,16	0,8770
1,17	0,8790
1,18	0,8810
1,19	0,8830
1,20	0,8849
1,21	0,8869
1,22	0,8888
1,23	0,8907
1,24	0,8925
1,25	0,8944
1,26	0,8962
1,27	0,8980
1,28	0,8997
1,29	0,9015
1,30	0,9032
1,31	0,9049
1,32	0,9066
1,33	0,9082
1,34	0,9099
1,35	0,9115

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
1,36	0,9131	1,70	0,9554	2,08	0,9812	2,76	0,9971
1,37	0,9147	1,71	0,9564	2,10	0,9821	2,78	0,9973
1,38	0,9162	1,72	0,9573	2,12	0,9830	2,80	0,9974
1,39	0,9177	1,73	0,9582	2,14	0,9838	2,82	0,9976
1,40	0,9192	1,74	0,9591	2,16	0,9846	2,84	0,9977
1,41	0,9207	1,75	0,9599	2,18	0,9854	2,86	0,9979
1,42	0,9222	1,76	0,9608	2,20	0,9861	2,88	0,9980
1,43	0,9236	1,77	0,9616	2,22	0,9868	2,90	0,9981
1,44	0,9251	1,78	0,9625	2,24	0,9875	2,92	0,9982
1,45	0,9265	1,79	0,9633	2,26	0,9881	2,94	0,9984
1,46	0,9279	1,80	0,9641	2,28	0,9887	2,96	0,9985
1,47	0,9292	1,81	0,9649	2,30	0,9893	2,98	0,9986
1,48	0,9306	1,82	0,9656	2,32	0,9898	3,00	0,9987
1,49	0,9319	1,83	0,9664	2,34	0,9904	3,05	0,9989
1,50	0,9332	1,84	0,9671	2,36	0,9909	3,10	0,9990
1,51	0,9345	1,85	0,9678	2,38	0,9913	3,15	0,9992
1,52	0,9357	1,86	0,9686	2,40	0,9918	3,20	0,9993
1,53	0,9370	1,87	0,9693	2,42	0,9922	3,25	0,9994
1,54	0,9382	1,88	0,9699	2,44	0,9927	3,30	0,9995
1,55	0,9394	1,89	0,9706	2,46	0,9931	3,35	0,9996
1,56	0,9406	1,90	0,9713	2,48	0,9934	3,40	0,9997
1,57	0,9418	1,91	0,9719	2,50	0,9938	3,45	0,9997
1,58	0,9429	1,92	0,9726	2,52	0,9941	3,50	0,9998
1,59	0,9441	1,93	0,9732	2,54	0,9945	3,55	0,9998
1,60	0,9452	1,94	0,9738	2,56	0,9948	3,60	0,9998
1,61	0,9463	1,95	0,9744	2,58	0,9951	3,65	0,9999
1,62	0,9474	1,96	0,9750	2,60	0,9953	3,70	0,9999
1,63	0,9484	1,97	0,9756	2,62	0,9956	3,75	0,9999
1,64	0,9495	1,98	0,9761	2,64	0,9959	3,80	0,9999
1,65	0,9505	1,99	0,9767	2,66	0,9961		
1,66	0,9515	2,00	0,9772	2,68	0,9963		
1,67	0,9525	2,02	0,9783	2,70	0,9965		
1,68	0,9535	2,04	0,9793	2,72	0,9967		
1,69	0,9545	2,06	0,9803	2,74	0,9969		

Irodalom

- http://math.uni-pannon.hu/~szalkai

- Solt György: Valószínűségszámítás (példatár)