1. VÉLETLEN ESEMÉNYEK

A valószínűségszámítás a véletlen jelenségek törvényszerűségének vizsgálatával foglalkozik. **Véletlen jelenség**en olyan jelenséget értünk, amelynek lejátszódását nagyon nagy számú külső körülmény határozza meg. A jelenség megfigyelésekor ezen nagy számú befolyásoló tényező közül csak a leglényegesebbeket tudjuk figyelembe venni. Ezek viszont még nem határozzák meg egyértelműen a jelenség lejátszódását. Ennek következtében a jelenség ismételt megfigyelésekor azt tapasztaljuk, hogy noha ezen lényeges befolyásoló tényezők mindig azonosak a jelenség lejátszódása nem mindig azonos, hanem véletlenszerű.

Valamely véletlen jelenség megfigyelését **véletlen kísérlet**nek (röviden kísérletnek) nevezzük függetlenül attól, hogy az előidézésében aktívan közreműködtünk-e, vagy csak passzív megfigyelők voltunk.

1.1 ELEMI ESEMÉNY, ÖSSZETETT ESEMÉNY

Egy véletlen kísérlettel kapcsolatosan mindig pontosan definiálnunk kell, hogy mit tekintünk a kísérlet különböző kimeneteleinek.

1.1 Definíció. A véletlen kísérlet lehetséges kimeneteleit **elemi esemény**eknek nevezzük. Az elemi események jelölésére az $e_1, e_2, \ldots$ szimbólumokat fogjuk használni.

1.2. Definíció.A véletlen kísérlettel kapcsolatos elemi események halmazát **eseménytér**nek nevezzük, és a továbbiakban *H*-val jelöljük.

A fenti két definíció megértéséhez olvassa el az 1.2. példát!

Valamely kísérlettel kapcsolatosan az elemi eseményeknél összetettebb történéseket is megfogalmazhatunk.

1.3. Definíció. Az eseménytér részhalmazait (véletlen) **esemény**eknek nevezzük. Az eseményeket nyomtatott nagy betűkkel fogjuk jelölni. Ha *A* valamely a vizsgált kísérlettel kapcsolatos esemény, akkor $A \subseteq H$.

Legyen $H = (e_1, e_2, \ldots)$ és $A = (e_{i_1}, e_{i_2}, \ldots, e_{i_k})$. Akkor mondjuk, hogy a kísérlet végrehajtásakor az ***A* esemény bekövetkezik**, ha az $e_{i_1}, e_{i_2}, \ldots, e_{i_k}$ elemi események valamelyike bekövetkezik. A kísérlet $e_{i_1}, e_{i_2}, \ldots, e_{i_k}$ kimeneteleit (az *A* esemény szempontjából) **kedvező esetek**nek , *k*-t pedig a kedvező esetek számának nevezzük.

Mivel $H \subseteq H$, ezért *H* is esemény, amit **biztos esemény**nek nevezzük, mivel ez a kísérlet bármely kimenetele esetén bekövetkezik.

Legyen $\varnothing$ az üres halmaz. $\varnothing \subseteq H$ miatt $\varnothing$ is esemény, mégpedig olyan, amelyik sohasem következhet be, ezért **lehetetlen esemény**nek nevezzük.

Hasonlóan, mivel az egyes elemi események az eseménytér egyelemű részhalmazai, az elemi események maguk is események (az adott kísérlettel kapcsolatosan a lehető legegyszerűbbek). Az események értelmezéséből adódóan az események halmazokkal reprezentálhatók.

1.1. Példa. Tekintsük azt a kísérletet, hogy feldobunk egy szabályos kockát. A kísérlet eredményét úgy tudjuk megadni, hogy megmondjuk, hogy hányat dobtunk. A rövidség kedvéért megegyezhetünk abban, hogy a továbbiakban pl. 1-gyel fogjuk jelölni azt a történést, hogy 1-et dobtunk stb. Nyilván a kísérlettel kapcsolatosan 6 különböző kimenetel, azaz 6 elemi esemény létezik és így

H=(1,2,3,4,5,6).

Ezzel a kísérlettel kapcsolatos események ennek a 6 elemű halmaznak a részhalmazai:

a) A=(2,4,6)="páros számot dobunk",

b) B=(1,2,3,4,5)="nem dobunk 6-ot",

c) C=(5,6)="4-nél nagyobbat dobunk", stb.

1.2. Példa. A kísérlet most abban áll, hogy kétszer egymás után feldobunk egy szabályos kockát. Most a kísérletnek egy lehetséges kimenetele, hogy elsõre 6-ot, a másodikra pedig 5-öt dobunk. Ettõl különbözõ kimenetele a kísérletnek, ha ugyanezeket a számokat dobjuk, de fordított sorrendben. A kísérlet kimenetelét most egyrendezett számpárral tudjuk leírni, ha megállapodunk abban, hogy az elsõ szám jelenti az elsõ dobás eredményét, a második pedig a második dobásét.

A különböző elemi események tehát:

$$\begin{array}{cccccc}
\{1,1\} & \{1,2\} & \{1,3\} & \{1,4\} & \{1,5\} & \{1,6\} \\
\{2,1\} & \{2,2\} & \{2,3\} & \{2,4\} & \{2,5\} & \{2,6\} \\
\{3,1\} & \{3,2\} & \{3,3\} & \{3,4\} & \{3,5\} & \{3,6\} \\
\{4,1\} & \{4,2\} & \{4,3\} & \{4,4\} & \{4,5\} & \{4,6\} \\
\{5,1\} & \{5,2\} & \{5,3\} & \{5,4\} & \{5,5\} & \{5,6\} \\
\{6,1\} & \{6,2\} & \{6,3\} & \{6,4\} & \{6,5\} & \{6,6\}
\end{array}$$

A H eseménytér pedig nem más, mint a felsorolt 36 elemi esemény halmaza. A kísérlettel kapcsolatos események ennek a 36 elemű halmaznak a részhalmazai:

a) A=$(\{1,1\},\{2,2\},\{3,3\},\{4,4\},\{5,5\},\{6,6\})$="egyforma számokat dobunk"

b) B=$(\{1,1\},\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{1,6\})$="elsőre 1-et dobunk",

c) C=$(\{1,1\},\{1,2\},\{2,1\},\{2,2\})$=" mindkét alkalommal 3-nál kisebbet dobunk"

1.3. Példa. Egyszerre feldobunk két különbözõ színû, például piros és fekete kockát. Vegyük észre, hogy a kísérlet lehetséges kimeneteleit most is rendezett számpárokkal tudjuk megadni, ha megállapodunk abban, hogy például mindig az elsõ helyre írjuk azt, hogy a piros kockával hányat dobtunk, és a másodikra helyre pedig azt, hogy hányat dobtunk a feketével. Tehát a $\{3,2\}$ azt jelenti, hogy a piros kockával 3-at, a feketével pedig 2-t dobtunk, a $\{4,4\}$ pedig azt, hogy mindkét kockával 4-et dobtunk, stb. A H eseménytér most is felfogható úgy, mint a fenti rendezett számpárok 36 elemû halmaza, és

$$A=(\{1,4\},\{2,4\},\{3,4\},\{5,4\},\{6,4\})$$

most azt az eseményt jelenti, hogy a fekete kockával 4-et dobunk.

1.4. Példa. Két egyforma (megkülönböztethetetlen) kockát egyszerre feldobunk. Mik lesznek most az elemi események? Lehetséges például, hogy 2 kettest dobunk, vagy hogy egy 1-et és egy 2-t dobunk stb. Az utóbbi esetben ennél a kísérletnél annak nincs értelme, hogy melyik kockával dobtunk 2-t, hiszen a két kockát nem tudjuk megkülönböztetni. Az elemi eseményeket most is számpárokkal tudjuk megadni, csak most ezek nem lesznek rendezettek. Például a fenti két lehetséges kimenetelt a (2,2), illetve az (1,2) számpárok fogják jelenteni. Tehát az elemi események a következõk:

$$\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
 & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
 & & (3,3) & (3,4) & (3,5) & (3,6) \\
 & & & (4,4) & (4,5) & (4,6) \\
 & & & & (5,5) & (5,6) \\
 & & & & & (6,6)
\end{array}$$

A H eseménytér most a fenti $6+5+4+3+2+1=\dfrac{6\cdot 7}{2}=21$ elemi eseményből áll.

Gyakorlásként gondoljuk meg, hogy a H melyik részhalmazai jelentik a következő eseményeket:

a) Mindkét dobott szám kisebb 5-nél.

b) A dobott számok összege páros.

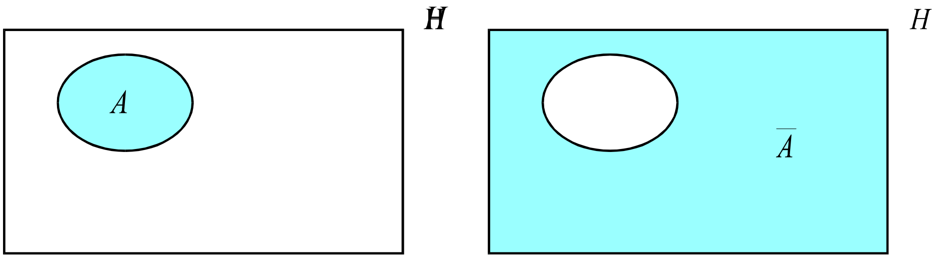
c) Egyik kockával sem dobunk 6-ot.

1.2. MŰVELETEK ESEMÉNYEKKEL

Legyen *H* valamely véletlen kísérlettel kapcsolatos eseménytér. A következőkben az eseményekről feltesszük, hogy ezzel a kísérlettel kapcsolatosak, tehát ugyanannak a *H* eseménytérnek a részhalmazai.

1.4. Definíció. Az *A* esemény komplementerén értjük azt az $\overline{A}$ eseményt, amely akkor és csakis akkor következik be, ha az A esemény nem következik be.

Az $\overline{A}$ esemény mindazokat az elemi eseményeket tartalmazza, amelyek az A-nak nem elemei, vagyis az $\overline{A}$ eseményt úgy kapjuk meg, hogy képezzük A-nak a H-ra vonatkozó komplementer(kiegészítő) halmazát.(1.1. ábra)



1.1. ÁBRA

Nyilvánvaló, hogy $\overline{H} = \varnothing$, és $\overline{\varnothing} = H$.

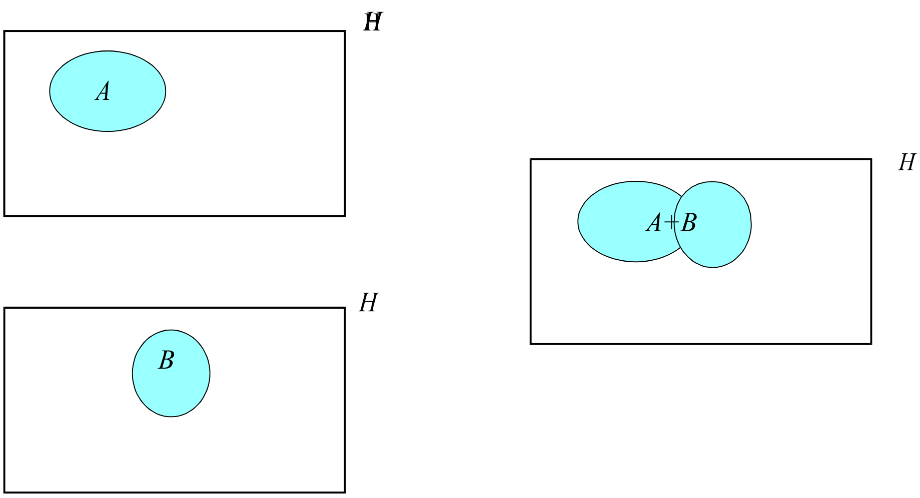
1.5. Definíció. Az A és B események $A+B$ összegén értjük azt az eseményt, amelyik akkor és csakis akkor következik be, ha A és B közül legalább az egyik bekövetkezik.

Az $A+B$ eseményt azok az elemi események alkotják, amelyek az A és B események közül legalább az egyikben benne vannak, azaz az eseménytérnek az $A+B$ eseményt jelentő részhalmazát úgy kapom meg, hogy képezem az A és B halmazok egyesítését. (1.2. ábra).

Ezután kettőnél több, akár végtelen sok esemény összegét is értelmezni tudjuk:

$$\sum_i A_i$$

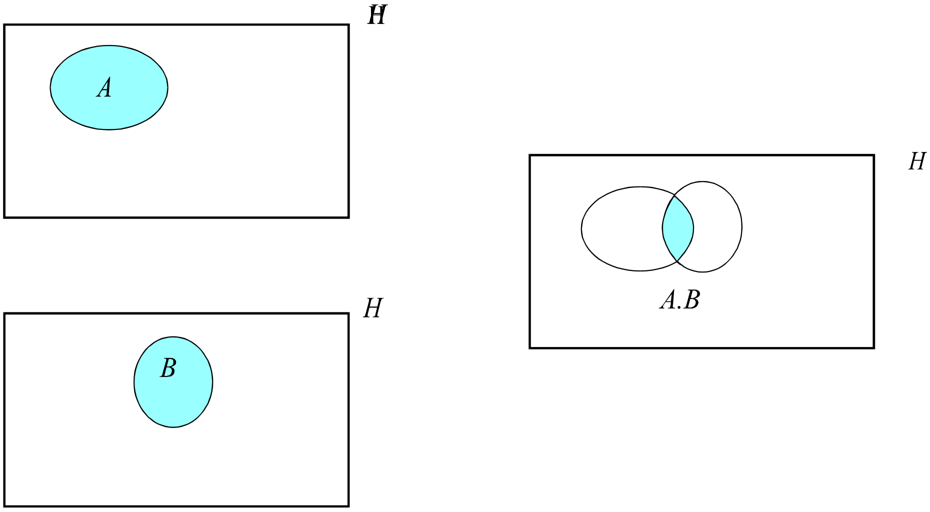
jelenti azt az eseményt, hogy az $A_1, A_2, \ldots$ események közül legalább egy bekövetkezik.



1.2. ÁBRA

1.6. Definícó. Az A és B események $A \cdot B$ szorzatán értjük azt az eseményt, amelyik akkor és csakis akkor következik be, ha az A és B események egyidejűleg bekövetkeznek. (A és B).

Az $A \cdot B$ eseményt azok az elemi események alkotják, amelyek az A és B események mindegyikben benne vannak, azaz az eseménytérnek az $A \cdot B$ eseményt jelentő részhalmazát úgy kapom meg, hogy képezem az A és B halmazok metszetét. (1.3. ábra).



1.3. ÁBRA

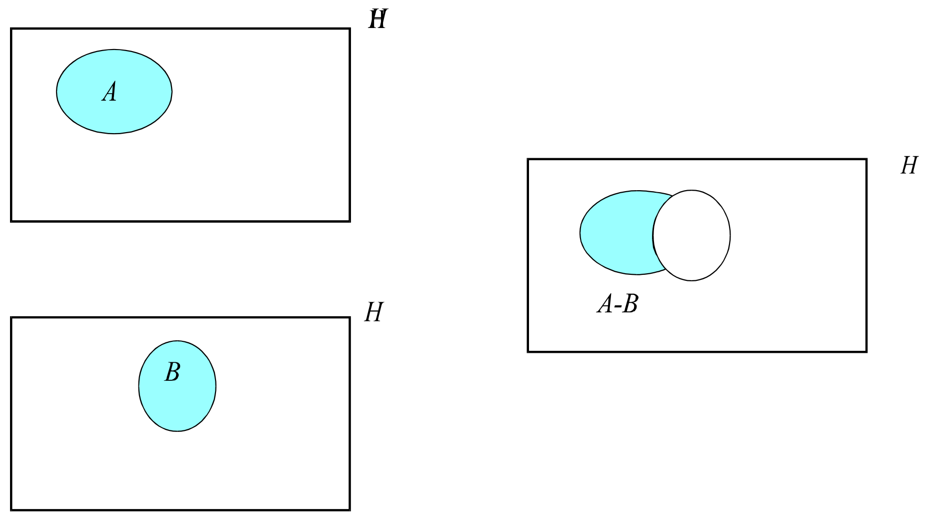
Kettőnél több esemény szorzata

$$\prod_i A_i$$

azt az eseményt jelenti, hogy az A_1, A_2, K események mindegyike bekövetkezik.

1.7. Definíció. Az A és B események $A - B$ különbségén értjük azt az eseményt, amelyik akkor és csakis akkor következik be, ha az A bekövetkezik de a B nem.

Az $A - B$ eseményt azok az elemi események alkotják, amelyek az A eseményben benne vannak, de nincsenek benne a B eseményben., Azaz az eseménytérnek az $A - B$ eseményt jelentő részhalmazát úgy kapom meg, hogy képezem az A és B halmazok különbségét.. (1.4. ábra).



1.4. ÁBRA

1.5. Példa. Az 1.2. Példában szereplõ kísérlettel kapcsolatban tekintsük a következõ eseményeket:

A="elsőre páros számot dobunk"

B="másodikra páros számot dobunk".

Ekkor

$\overline{A}$ ="elsőre páratlan számot dobunk"

A+*B*=" van páros a dobott számok között",

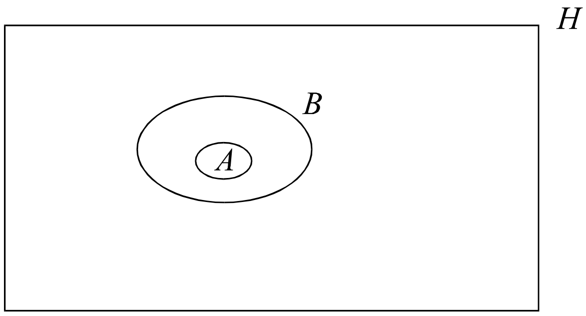
$A \cdot B$="mindkét dobott szám páros"

A-*B*="elsőre páratlan számot dobunk és a másodikra páratlant"=$A \cdot \overline{B}$.

1.3 AZ ESEMÉNYMŰVELETEK TULAJDONSÁGAI

1.8. Definíció. Az *A* esemény maga után vonja a *B* -t, ha az A bekövetkezésekor egyúttal a *B* is bekövetkezik.

Ez azt jelenti, hogy mindazok az elemi események, amelyek elemei *A*-nak elemei a *B*-nek is. Vagyis az *A* részhalmaza a *B*-nek. (1.5. ábra) Ezért, ha az *A* maga után vonja *B*-t, akkor ezt úgy jelöljük, hogy $A \subseteq B$.



1.5. ÁBRA

1.6. Példa. A kockadobás kísérleténél *A* jelentse, hogy 2-t dobunk, *B* pedig azt, hogy páros számot:

$$A = (2)$$
$$B = (2,4,6).$$

Ha kettőt dobunk, akkor egyúttal az is bekövetkezett, hogy páros számot dobtunk, tehát $A \subseteq B$.

1.9. Definíció.Az A és B események egyenlők (A=B), ha $A \subseteq B$ és $B \subseteq A$.

Két esemény tehát akkor egyenlő, ha a két esemény ugyanazokat az elemi eseményeket tartalmazza, ami azt jelenti, hogy ha *A* bekövetkezik, akkor *B* is, és ha *A* nem következik be, akkor *B* sem.

1.10. Definíció. Az A és B események egymást kizárják, ha a kettő egyszerre nem következhet be, azaz, ha

$$A \cdot B = \emptyset.$$

1.11. Definíció. Az $A_1, \dots, A_n$ események egymást páronként kizárják, ha közülük bármely kettő kizárja egymást, azaz

$$A_i . A_j = \emptyset, \text{ ahol } 1 \le i, j \le n, \;\; i \ne j.$$

1.7. Példa. A kockadobás kísérletével kapcsolatban tekintsük a következő eseményeket:

A="1-et vagy 2-t dobunk"=(1,2),

B="3-at vagy 4-et dobunk"=(3,4),

C="5-öt vagy 6-ot dobunk"=(5,6).

Ezek egymást páronként kizáró események, hiszen

$$A \cdot B = \emptyset, \quad A \cdot C = \emptyset, \quad B \cdot C = \emptyset.$$

Ebből következik, hogy ez a három esemény nem következhet be egyszerre:

$$A \cdot B \cdot C = \emptyset.$$

Fordítva viszont az nem igaz, hogy ha három esemény egyszerre nem következhet be, akkor azok egymást páronként kizárják. Az előbbi kísérlettel kapcsolatban vegyük a következő három eseményt:

E="1-et, vagy 2-öt dobunk"=(1,2),

F="1-et, vagy 3-at dobunk"=(1,3),

G="4-et, vagy 5-öt dobunk"=(4,5).

Ez a három esemény is olyan, hogy egyszerre nem következhetnek be, $E \cdot F \cdot G = \emptyset$, de nem igaz, hogy egymást páronként kizárják, mivel

$$E \cdot F = (1,2) \cdot (1,3) = (1) \neq \emptyset.$$

Ezek után rátérünk az eseményműveletek tulajdonságainak tárgyalására. Akiknek az alábbi tulajdonságok és bizonyításaik először túlságosan formálisnak tűnnek, azok az eseményeket mindig úgy képzeljék el, mint egy alaphalmaz (pl. egy téglalap) egymáshoz viszonyítva általános helyzetű részhalmazait, amint ez az 1.2. -1.6. ábrákon is látható.

1. Az események összegének és szorzatának definíciójából azonnal adódik, hogy tetszőleges *A* esemény esetén

$$A + A = A, \quad A \cdot A = A.$$

2. Az események összeadása és szorzása kommutatív és asszociatív művelet, továbbá a szorzás az összeadásra nézve disztributív:

$$A + B = B + A,$$
$$A.B = B.A,$$
$$(A + B) + C = A + (B + C),$$
$$(A.B) \cdot C = A. \cdot (B \cdot C),$$
$$(A + B). \cdot C = A \cdot C + B \cdot c$$

3.. Tetszőleges *A* és *B* eseményekkel érvényesek a következők:

$$a)\overline{\overline{A}} = A,$$
$$b)A + \emptyset = A, A.\emptyset = \emptyset,$$
$$c)A + H = H, A.H = A,$$
$$d)A + \overline{A} = H, A.\overline{A} = \emptyset,$$

.

4. $\overline{A+B} = \overline{A}.\overline{B}, \overline{A.B} = \overline{A} + \overline{B}$

5. Ha $A \subseteq B$, akkor

$$A + B = B, \text{ és } A \cdot B = A.$$

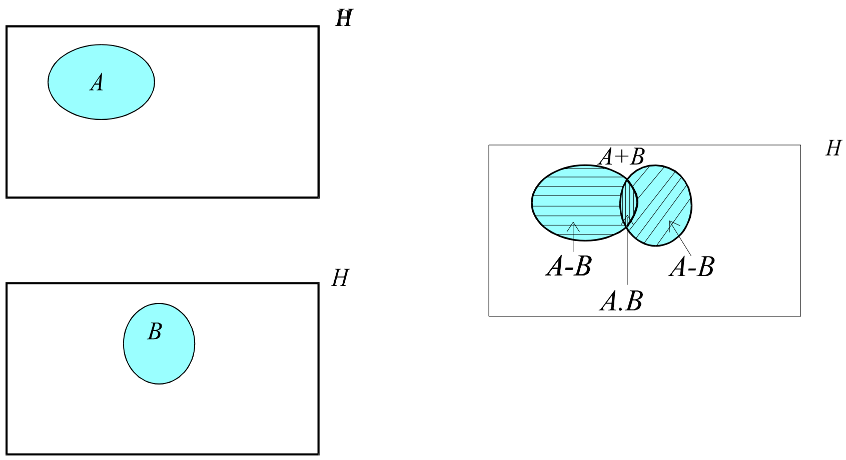
6. Az események különbségével kapcsolatban érvényesek a következő azonosságok:

$a) A - B = A.\overline{B} = A - A.B,$
$b) (A - B) + B = A + B,$
$c)$ ha $B \subseteq A,$ akkor $(A - B) + B = A.$

7. Tetszőleges A és B esemény esetén az összegük felírható egymást páronként kizáró események összegeként:

$$A + B = (A - B) + A \cdot B + (B - A).$$

Ezt a felbontást az alábbi1.6. ábrán szemléltetjük.



1.6. ÁBRA

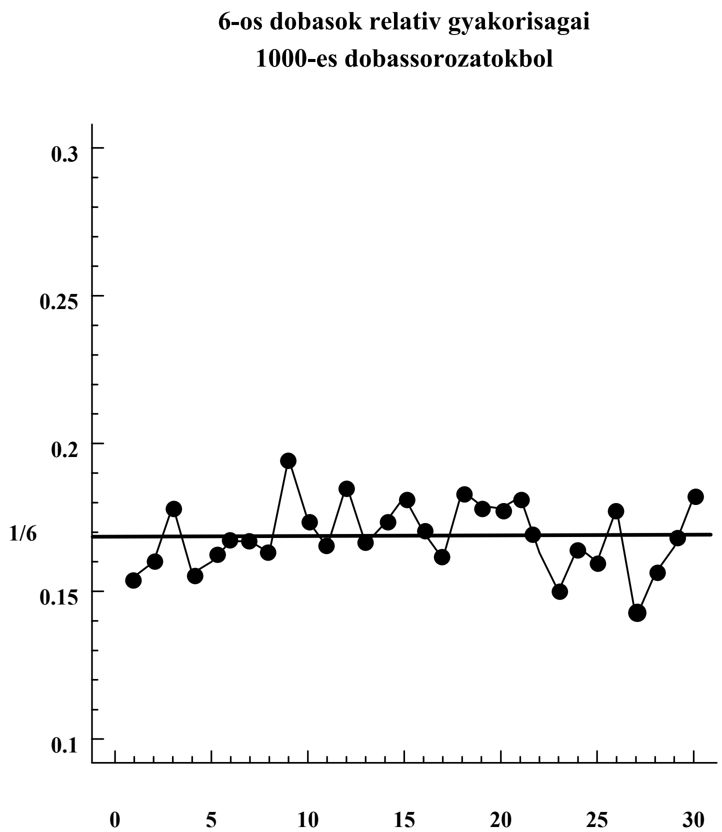
2. VALÓSZÍNŰSÉGSZÁMÍTÁSI ALAPOK

2.1 VALÓSZÍNŰSÉGEK

Legyen A valamely kísérlettel kapcsolatos esemény. Hajtsuk végre n-szer a kísérletet egymástól függetlenül, és mindegyik végrehajtáskor csak az érdekel bennünket, hogy az A esemény bekövetkezett vagy sem. Ezt az A-ra vonatkozó n hosszúságú, független kísérletsorozatnak (röviden kísérletsorozatnak) nevezzük.

2.1. Definíció.Valamely n hosszúságú kísérletsorozat folyamán az A esemény bekövetkezéseinek n_A számát az esemény **gyakoriságának**, az n_A / n hányadost pedig az esemény relatív gyakoriságának nevezzük a kísérletsorozat folyamán.

Tekintsük a kockadobás kísérletét, és A legyen az az esemény, hogy hatost dobunk. 1000-es dobássorozatokat ismételtünk 30-szor, és minden sorozatból kiszámítottuk a 6-os dobás relatív gyakoriságát. Ezeket a relatív gyakoriságokat ábrázoltuk a 2.1. ábrán



2.1.. ÁBRA

Azt tapasztaljuk, hogy a relatív gyakoriságok különböznek egymástól, de bizonyos törvényszerűséget felfedezhetünk, nevezetesen azt, hogy a kapott relatív gyakoriságok az 1/6 körül ingadoznak . Azt is megfigyelték, hogy az ingadozások mértéke annál kisebb, minél hosszabb dobássorozatokból számítjuk a relatív gyakoriságokat.

Az emberiségnek a véletlen eseményekre vonatkozó több évszázados tapasztalata azt mutatja, hogy minden véletlen eseményhez tartozik egyetlen olyan valós szám, amely körül a kísérletsorozatokból számított relatív gyakoriságok ingadoznak. Ezt a tényt szokás a *nagy számok empirikus törvényének* nevezni. Ennek alapján adjuk meg a valószínűség szemléletes -nem egzakt matematikai- definícióját.

2.2. Definíció. Minden A eseményhez létezik egy olyan valós szám, amely körül az A esemény ismételt kísérletsorozatokból számított relatív gyakoriságai ingadoznak. Ezt a számot nevezzük az **esemény valószínűségének**.

Az A esemény valószínűségét $P(A)$-val jelöljük.

2.2 A VALÓSZÍNŰSÉG AXIÓMÁI

A valószínűség axiómái a valószínűség olyan alapvető tulajdonságai, amelyeket a tapasztalatok alapján igaznak tekintünk, ezeket más, elemibb tételekből nem tudjuk levezetni. Az axiómák a relatív gyakoriságok alábbi, bizonyítható tulajdonságait tükrözik.

a) Tetszőleges A esemény n hosszúságú kísérletsorozatból számított n_A / n relatív gyakoriságára érvényes, hogy

$$0 \leq \frac{n_A}{n} \leq 1,$$

mivel $0 \leq n_A \leq n$.

b) Mivel a H biztos esemény abszolút gyakorisága bármely n hosszúságú kísérletsorozat folyamán n, így a relatív gyakorisága

$$\frac{n_H}{n} = \frac{n}{n} = 1.$$

c) Legyenek $A_1, A_2, \dots, A_m$ egymást páronként kizáró események. Ekkor könnyen beláthatjuk, hogy az $A_1 + A_2 + \cdots + A_m$ esemény abszolút gyakorisága egyenlő az egyes események abszolút gyakoriságainak összegével. (Az $A_1 + A_2 + \cdots + A_m$ esemény akkor és csakis akkor következik be ha az $A_1, A_2, \dots, A_m$ események közül legalább az egyik bekövetkezik. De ezek egymást páronként kizáró események, tehát egyszerre közülük csak egy következhet be, Azaz az $A_1 + A_2 + \cdots + A_m$

bekövetkezéseinek a száma= A_1 bekövetkezéseinek a száma+ A_2 bekövetkezéseinek a száma+...+ A_m bekövetkezéseinek a száma.)

Így

$$\frac{n_{A_1+\cdots A_m}}{n}=\frac{n_{A_1}+\cdots n_{A_m}}{n}=\frac{n_{A_1}}{n}+\cdots\frac{n_{A_m}}{n},$$

vagyis egymást páronként kizáró események összegének relatív gyakorisága egyenlő az egyes események relatív gyakoriságainak összegével.

A valószínűség axiómái:

I. Tetszőleges A eseményhez tartozik egy $P(A)$ szám, amit az esemény valószínűségének nevezünk, és erre fennáll, hogy

$$P(A)\geq 0.$$

II. A biztos esemény valószínűsége 1,

$$P(H)=1.$$

III. Ha a véges, vagy megszámlálhatóan végtelen sok $A_1, A_2, \ldots$ események egymást páronként kizárják, akkor

$$P\left(\sum_i A_i\right)=\sum_i P(A_i).$$

A III. axiómát szokás a valószínűség additivitásának nevezni.

Ezeket az axiómákat a valószínűség **Kolmogorov-féle axiómáinak** nevezzük.

2.3 A VÉLETLEN KÍSÉRLET MATEMATIKAI MODELLJE

Valamely véletlen kísérletet egy {H, $\mathcal{A}$, P} hármassal modellezhetünk, ahol

a) H az illető kísérlettel kapcsolatos elemi események halmaza, amit eseménytérnek nevezünk.

b) $\mathcal{A}$ az események halmaza..

c) P az $\mathcal{A}$-n értelmezett függvény, amit valószínűségnek nevezünk, és amelyre teljesülnek a Kolmogorov-féle axiómák:

I. tetszőleges $A\in\mathcal{A}$ esetén $P(A)\geq 0$,

II. $P(H)=1$,

III. Ha a véges, vagy megszámlálhatóan végtelen sok $A_1, A_2, \ldots$ események egymást páronként kizárják, akkor

$$P\left(\sum_i A_i\right)=\sum_i P(A_i).$$

2.4 Klasszikus valószínűségi kísérlet

2.3. Definíció. Ha a kísérletnek véges sok lehetséges kimenetele van és ezek mind egyformán valószínűek, akkor a kísérletet **klasszikus valószínűségi kísérletnek** nevezzük.

2.1. Tétel. Legyen $H=\{e_1,\ldots,e_N\}$ és $P(e_1)=\ldots=P(e_N)=p$. Ekkor tetszőleges, k elemi eseményt tartalmazó $A=\{e_{i_1},\ldots,e_{i_k}\}\subseteq H$ esemény esetén

$$P(A)=\frac{k}{N}.$$

A fenti eredményt sokszor úgy mondjuk hogy "a kedvező esetek száma osztva az összes esetek számával", mivel az elemi események között k olyan van, amelyek esetén az A esemény bekövetkezik (ezek kedvezők az A szempontjából). N viszont a kísérlet összes lehetséges kimeneteleinek a száma. Ilyen esetekben szokás kombinatorikus valószínűségről beszélni, mivel az összes, illetve a kedvező esetek számának meghatározása sok esetben kombinatorikus meggondolásokat igényel. A szükséges kombinatorikai alapelemeket az 1.Függelékben foglaltuk össze.

2.1. Példa. Háromszor egymás után feldobunk egy kockát. Mi annak a valószínűsége, hogy az első és az utolsó dobás is 6-os lesz?

A kísérlet összes lehetséges kimeneteleinek a száma

$$N=6\cdot 6\cdot 6=6^3.$$

(Gondoljuk végig, hogy mivel reprezentálhatók az elemi események!)

A kedvező esetek számának meghatározásához vegyük figyelembe, hogy a vizsgált esemény bekövetkezését jelenti minden olyan rendezett számhármas, amelynek első és harmadik eleme 6-os, a második helyen pedig az 1, ..., 6 számok bármelyike szerepelhet:

$$\{6,\quad \text{bármi},\quad 6\}$$

Az ilyen sorozatok száma 6, tehát $k=6$, azaz

$$P = \frac{k}{N} = \frac{6}{6^3} = \frac{1}{6^2}.$$

2.2. Példa.A 32 lapos magyar kártyából véletlenszerűen kihúzunk egy lapot úgy, hogy bármely lap választása egyformán valószínű. Mi annak a valószínűsége, hogy a kihúzott lap piros?

Ha A jelenti a kérdéses eseményt, akkor

$$P = \frac{k}{N},$$

ahol most N=32. Kérdés, hogy ezek között hány olyan van, amelyik az A bekövetkezését jelenti. Ha a 8 piros lap bármelyikét húzzuk ki, akkor az A esemény bekövetkezik, tehát k=8 és

$$P(A) = \frac{8}{32} = \frac{1}{4}.$$

2.3. Példa. A 32 lapos magyar kártyából egyszerre kihúzunk 3 lapot úgy, hogy bármely 3 lap kiválasztása egyformán valószínű. Mi a valószínűsége annak, hogy egy piros és két zöld lesz a kiválasztottak között?

Legyen A a vizsgált esemény. A kísérlet kapcsolatos elemi események száma

$$N = \binom{32}{3}.$$

Az A esemény bekövetkezését jelenti minden olyan kiválasztás, amikor 1-et választunk a 8 piros közül, és 2-t a 8 zöld lap közül:

$$k = \binom{8}{1}\binom{8}{2},$$

vagyis

$$P(A) = \frac{\binom{8}{1}\binom{8}{2}}{\binom{32}{3}} = \frac{\frac{8!}{1!\cdot 7!}\cdot\frac{8!}{2!\cdot 6!}}{\frac{32!}{3!\cdot 29!}} = \frac{8\cdot\frac{7\cdot 8}{2}}{\frac{30\cdot 31\cdot 32}{2\cdot 3}} = \frac{7}{155}.$$

Megjegyzés. Hogy egyszerre húzzuk ki a három lapot azzal egyenértékű, hogy egymás után, visszatevés nélkül választunk ki három lapot, és a kiválasztás sorrendjére nem vagyunk tekintettel, más szóval nem tekintjük különböző kiválasztásnak azokat az eseteket, amikor ugyanazokat a lapokat választjuk ki, de különböző sorrendben. Ha így definiáljuk az elemi eseményeket, akkor ezzel a kísérlettel kapcsolatban nincs értelme pl. annak a z eseménynek, hogy a 2. kiválasztott lap piros.

2.4. Példa. Az előző feladatban említett kísérletet most úgy hajtjuk végre, hogy visszatevéssel választunk ki 3 lapot úgy, hogy bármely 3-as sorozat kiválasztásának ugyanaz legyen a valószínűsége. (Most különböző kimeneteleknek tekintjük, ha ugyanazokat a lapokat választjuk ki, de különböző sorrendben.) Határozzuk meg, hogy mi lesz most az előző példában vizsgált esemény valószínűsége.

Mivel a választások visszatevéssel történnek, mind a három lap kiválasztásánál adódhat a 32 lap bármelyike, tehát

$$N = 32 \cdot 32 \cdot 32 = 32^3.$$

Vegyük számba a kedvező eseteket!

Piros lap lehetősége	Zöld lapok lehetőségei		Ilyen esetek száma
1. lap	2. lap	3. lap	$8 \cdot 8 \cdot 8 = 8^3$
2. lap	1. lap	3. lap	$8 \cdot 8 \cdot 8 = 8^3$
3. lap	1. lap	2. lap	$8 \cdot 8 \cdot 8 = 8^3$

Tehát, ha a kedvező eseteket először úgy osztályozzuk, hogy a piros lap hányadik helyen fordul elő, akkor $\binom{3}{1}$ osztályt kapunk, és ezek mindegyikébe 8^3 elemi esemény esik, tehát

$$k = \binom{3}{1} 8^3,$$

és így

$$P(A) = \frac{\binom{3}{1} 8^3}{32^3} = 3 \cdot \left(\frac{1}{4}\right)^3.$$

2.5 A valószínűség tulajdonságai

Ebben a bekezdésben olyan tételeket veszünk sorra, amelyek lehetővé teszik, hogy már ismert valószínűségű események segítségével más események valószínűségét meghatározhassuk.

1. Ha az A esemény $P(A)$ valószínűsége ismert, akkor

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

Következmény.

$$P(\varnothing) = P(\overline{H}) = 1 - P(H) = 0.$$

2.4. Definíció. Az $A_1,\ldots,A_n$ események teljes eseményrendszert alkotnak, ha összegük a biztos esemény és egymást páronként kizárják, azaz ha

$$A_1+\ldots+A_n=H,$$

és

$$A_i \,.\, A_j=\varnothing \quad i\neq j.$$

2. Ha az $A_1,\ldots,A_n$ események teljes eseményrendszert alkotnak, akkor

$$P(A_1)+\ldots+P(A_n)=1.$$

3. Ha $B\subseteq A$, akkor $P(A-B)=P(A)-P(B)$.

4. Ha $B\subseteq A$,akkor $P(B)\leq P(A)$.

5. Tetszőleges *A* és *B* eseményekre igaz, hogy

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

2.6 FELTÉTELES VALÓSZÍNŰSÉG, ESEMÉNYEK FÜGGETLENSÉGE

2.6.1. A feltételes valószínűség

2.5. Definíció. Legyen adott két esemény *A* és *B*, továbbá tegyük fel, hogy $P(B)>0$. Az *A* esemény *B*-re vonatkozó feltételes valószínűsége

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

Szemléletes jelentése: ha tudom, hogy a *B* esemény bekövetkezik, akkor mi annak a valószínűsége, hogy az *A* esemény is bekövetkezik. Legyen $H=\{e_1,\ldots,e_N\}$, ahol az elemi események egyformán valószínűek, és $B=\{e_{i_1},\ldots,e_{i_{k_B}}\}\subseteq H$. Ilyen esetekben nem tudom , hogy konkrétan melyik elemi esemény következik be, de az az információ, hogy a *B* esemény bekövetkezik csökkenti a bizonytalanságom mértékét, mert azt tudom, hogy biztosan az $e_{i_1},\ldots,e_{i_{k_B}}$ elemi események valamelyike következik be. Ezen elemi események között általában vannak olyanok, amelyek egyúttal az *A* esemény bekövetkezését is jelentik, és vannak olyanok, amelyek nem vonják maguk után az *A* bekövetkezését. A *B* esemény bekövetkezte, mint a kísérlet kimenetelére vonatkozó információ azzal ekvivalens, mintha egy olyan kísérletről lenne szó, amelynek esetén $e_{i_1},\ldots,e_{i_{k_B}}$ az összes lehetséges, egyformán valószínű kimenetele (a lehetséges esetek száma k_B). Tegyük fel, hogy közöttük $k_{AB}\,(\le k_B)$ olyan elemi esemény van, amelyik egyúttal az A bekövetkezését is jelent. Egy ilyen kísérletben (amikor B a biztos esemény) a A esemény bekövetkezésének valószínűsége

$$\frac{k_{AB}}{k_B}.$$

Ez indokolja azt, hogy az eredeti kísérlettel kapcsolatosan, amikor $H=\{e_1,\ldots,e_N\}$ a

$$\frac{k_{AB}}{k_B}=\frac{\frac{k_{AB}}{N}}{\frac{k_B}{N}}=\frac{P(A.B)}{P(B)}$$

hányadossal értelmezzük a feltételes valószínűséget.

2.5. Példa. Egymás után kétszer feldobunk egy kockát. Feltéve, hogy a dobott számok összege páros, mi a valószínűsége annak, hogy az első szám kisebb mint 4?

Legyen A az az esemény, hogy az első szám kisebb mint 4, B pedig az, hogy a dobott számok összege páros.

A $P(A|B)$ feltételes valószínűség a kérdés.

a) A kísérlet összes lehetséges kimeneteleinek szám $N=6^2$.

A B esemény szempontjából kedvezők azok az esetek, amikor mindkét szám páratlan, vagy amikor mindkettő páros, tehát

$$k_B=3^2+3^2=18,$$

vagyis

$$P(B)=\frac{18}{36}.$$

Az $A\cdot B$ esemény azt jelenti, hogy a dobott számok összege páros és az első szám kisebb, mint 4. Tehát

első szám	második szám	lehetőségek száma
1	1	
	3	
	5	3
2	2	
	4	
	6	3
3	1	
	3	
	5	3

$$k_{AB} = 3 \cdot 3 = 9,$$

$$P(A \cdot B) = \frac{9}{36}.$$

A feltételes valószínűség definíciója szerint

$$P(A|B) = \frac{\frac{9}{36}}{\frac{18}{36}} = \frac{9}{18} = \frac{1}{2}.$$

b) A kérdéses feltételes valószínűséget a következő meggondolással is meghatározhatjuk.

Ha tudom, hogy a B esemény bekövetkezett, akkor ez azt jelenti, hogy a kísérlet 36 lehetséges kimenetele közül annak a 18 esetnek valamelyike következett be, amikor mindkét szám páros, vagy mindkettő páratlan. , tehát a B esemény bekövetkezése esetén 18 az összes esetek száma, és ezek mindegyike egyformán valószínűek

Ezek között 9 olyan van, amely egyúttal az A esemény bekövetkezését is jelenti, tehát a kedvező esetek száma 9, és így

$$P(A|B) = \frac{9}{18} = \frac{1}{2}.$$

2.6.2. A feltételes valószínűség tulajdonságai

1. $0 \le P(A|B) \le 1$.

2. $P(H|B) = 1, P(B|B) = 1$.

3. Ha A_1 és A_2 egymást kizáró események, akkor

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

4. Ha $B \subseteq A$, akkor $P(A|B) = 1$.

5. Ha $A \subseteq B$, akkor $P(A|B) = \dfrac{P(A)}{P(B)}$.

6. Ha A és B egymást kizárják, akkor $P(A|B) = 0$.

2.6.3. Események függetlensége

2.6. Definíció. Az A és B események függetlenek, ha

$$P(A \cdot B) = P(A)P(B)$$

Következmények

1) Ha $P(A) = 0$, akkor A bármely B eseménytől független. $A \cdot B \subseteq A$ miatt $P(A \cdot B) = 0$.

2) Ha A és B függetlenek és $P(A)$>0, akkor $P(B|A) = P(B)$.

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

3) Az előzőből adódik, hogy ha az A és B pozitív valószínűségű események függetlenek, akkor

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

2.7. Tétel. Ha az A és B események függetlenek, akkor az A és $\overline{B}$, B és $\overline{A}$, valamint az $\overline{A}$ és $\overline{B}$ események is függetlenek.

2.7. Definíció. Az $A_1, \ldots, A_n$ események **páronként függetlenek**, ha közülük bármelyik kettő független , azaz

$$P(A_i \cdot A_k) = P(A_i)P(A_k) \quad (1 \le i < k \le n).$$

2.8. Definíció. Az $A_1, \ldots, A_n$ események **teljesen függetlenek** (röviden függetlenek), ha közülük akárhányat kiválasztva azok szorzatának valószínűsége egyenlő a valószínűségeik szorzatával;

$$P(A_i \cdot A_k) = P(A_i)P(A_k) \quad (1 \le i < k \le n).$$

$$P(A_i \cdot A_j \cdot A_k) = P(A_i)P(A_j)P(A_k) \quad (1 \le i < j < k \le n),$$

$$\vdots$$

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

A definíciók alapján nyilvánvaló, hogy ha az $A_1, \ldots, A_n$ események teljesen függetlenek, akkor páronként is függetlenek. Fordítva viszont nem igaz, a páronkénti függetlenségből nem következik a teljes függetlenség.

Bebizonyítható a 2.7. Tétel következő általánosítása is.

2.8. Tétel. Ha az $A_1, \ldots, A_n$ események függetlenek, akkor közülük akárhányat a komplementerére cserélünk, akkor az így kapott események is függetlenek lesznek.

2.9. Tétel. Ha az $A_1, \ldots, A_n$ események függetlenek és azonos *p* valószínűségűek, akkor annak a valószínűsége, hogy közülük *k* esemény következik be

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

2.6.4. Független kísérletek

2.9. Definíció. Két kísérletet függetlennek nevezünk, ha az egyik kísérlet kimenetele nem befolyásolja a másik kísérlettel kapcsolatos történéseket.

Ha van két független kísérlet, és A az egyikkel, B pedig a másikkal kapcsolatos esemény, akkor $A \cdot B$ jelenti azt, hogy az egyik kísérletben az A esemény, a másik kísérletben pedig a B esemény következik be. Ekkor érvényes, hogy

$$P(A \cdot B) = P(A)P(B).$$

Több kísérlet függetlenségét hasonló módon értelmezzük.

2.10. Definíció. Független kísérletsorozatról beszélünk, ha egy kísérletet *n*-szer úgy ismétlünk meg, hogy a soron következő végrehajtáskor semmi módon nem befolyásol bennünket, hogy az előzőekben már mit tapasztaltunk. Nyilván az így ismételt kísérletek függetlenek.

2.10. Tétel. Tekintsünk egy kísérletet, és egy vele kapcsolatos A eseményt. Hajtsuk végre a kísérletet *n*-szer, egymástól függetlenül, és mindegyik ismétlésnél csak az érdekel bennünket, hogy A esemény bekövetkezik-e. Annak a valószínűsége, hogy az *n* hosszúságú független kísérletsorozat folyamán az A esemény *k*-szor következik be

$$\binom{n}{k} p^k (1-p)^{n-k},$$

ahol *p* az A esemény valószínűsége.

2.6. Példa. Tízszer feldobunk egy kockát. Mi a valószínűsége annak, hogy háromszor dobunk párosat?

Tekintsük azt a kísérletet, hogy feldobunk egy kockát. Jelölje A azt az eseményt, hogy páros számot dobunk.

$$p = P(A) = \frac{3}{6} = \frac{1}{2}.$$

A feladatban arról van szó, hogy az A esemény bekövetkezésére vonatkozóan végzünk 10 hosszúságú kísérletsorozatot, és arra vagyunk kíváncsiak, hogy mi annak a valószínűsége, hogy ennek folyamán az A esemény háromszor következik be. A 2.9. Tételből azonnal adódik, hogy a keresett valószínűség

$$\binom{10}{3}\left(\frac{1}{2}\right)^3\left(1-\frac{1}{2}\right)^{10-3} = \frac{120}{1024} = \frac{15}{128}.$$

2.7 A VALÓSZÍNŰSÉGEK SZORZÁSI SZABÁLYA

Az előzőkben láttuk, hogy független események együttes bekövetkezésének valószínűségét úgy számíthatjuk ki, hogy az egyes események valószínűségeit összeszorozzuk. A következő tételben azt mutatjuk meg, hogy miként lehet az együttes bekövetkezés valószínűségét meghatározni akkor, ha az események nem függetlenek.

2.11. Tétel. Tetszőleges A és B események esetén

$$P(A.B) = \begin{cases} 0 & \text{ha } P(B) = 0 \\ P(A|B)P(B) & \text{ha } P(B) \neq 0. \end{cases}$$

A tétel könnyen általánosítható több tényező szorzatára is. Tekintsünk tetszőleges $A_1, \ldots, A_n$ eseményeket. Ha közöttük van olyan, amelyiknek a valószínűsége nulla, akkor az előbbi meggondolásokhoz hasonlóan adódik, hogy a szorzatuk is nulla valószínűségű lesz.

2.12. Tétel. Ha $A_1, \ldots, A_n$ tetszőleges események és $P(A_i) \neq 0, i = 1, \ldots, n,$ akkor

$$P(A_1.A_2 \ldots A_n) = P(A_1|A_2.A_3 \ldots A_n)P(A_2|A_3 \ldots A_n)\ldots P(A_{n-1}|A_n)P(A_n).$$

Például n=3 esetén ez a következőt jelenti:

$$P(A_1.A_2.A_3) = P(A_1|A_2.A_3)P(A_2|A_3)P(A_3).$$

2.7. Példa. Két dobozban golyók vannak. Az I.-ben 4 fehér és két fekete, a II.-ban 8 fehér és két fekete. Tekintsük a következő kísérletet. Feldobunk egy érmét, ha fejet dobunk, akkor az I. dobozból választunk véletlenszerűen egy golyót, ha írást dobunk, akkor pedig a II. dobozból. Mi a valószínűsége annak, hogy az I. dobozból fehér golyót választunk?

A: fehér golyót választunk,
B: az I. dobozból választunk.

A szorzat esemény valószínűségére vagyunk kíváncsiak. Két kísérletet hajtunk végre egymás után. Az első kísérlet az érme feldobása. Lényegében ezzel kapcsolatos a B esemény, mert B akkor és csakis akkor következik be, ha fejet dobunk. A Második kísérlet pedig az, hogy választunk egy golyót, és ezzel kapcsolatos a A esemény. Ez a két kísérlet azonban nem tekinthető egymástól függetlennek, mert az első kísérlet eredménytől függ az, hogy melyik dobozból választjuk a golyót. Tehát a szorzat esemény valószínűségének meghatározására nem alkalmazhatjuk a független kísérletekkel kapcsolatban elmondottakat.

A szorzási szabály szerint

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

Mivel B akkor és csakis akkor következik be, ha fejet dobunk, a valószínűsége

$$P(B) = \frac{1}{2}.$$

Ha a B esemény bekövetkezett, akkor a második kísérlet abban áll, hogy az I. dobozból, 6 golyó(4 fehér+2 fekete) közül választunk.

$$P(A|B) = \frac{4}{6}.$$

Végül

$$P(A \cdot B) = \frac{4}{6} \cdot \frac{1}{2} = \frac{1}{3}.$$

2.8 TELJES VALÓSZÍNŰSÉG TÉTELE, BAYES -ÉTEL

2.13. Tétel. Legyen $A_1, \ldots, A_n$ teljes eseményrendszer és *B* egy tetszőleges esemény. Ekkor

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n).$$

Ezt a tételt nevezik a **teljes valószínűség tételének**. Arra nyújt lehetőséget, hogy egy esemény valószínűségét meghatározzuk, ha ismerjük az eseménynek egy teljes eseményrendszer elemeire vonatkozó feltételes valószínűségeit, és a teljes eseményrendszer elemeinek valószínűségeit.

2.8. Példa. A 2.7. példában szereplő kísérlettel kapcsolatban határozzuk meg annak a valószínűségét, hogy fekete golyót választunk.

B: fekete golyót választunk.

A_1: az I. dobozból választunk,
A_2: a II. dobozból választunk.

A_1, A_2 teljes esemény rendszer, és

$$P(A_1) = P(A_2) = \frac{1}{2},$$

továbbá a B eseménynek a teljes eseményrendszer elemeire vonatkozó feltételes valószínűségei

$$P(B|A_1) = \frac{2}{6}, \text{ és } P(B|A_2) = \frac{2}{10}.$$

A teljes valószínűség tétele szerint

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = \frac{2}{6} \cdot \frac{1}{2} + \frac{2}{10} \cdot \frac{1}{2} = \frac{4}{15}.$$

Gyakorlati alkalmazásoknál sokszor előfordul az is, hogy egy teljes eseményrendszerrel kapcsolatban nem valamely másik A esemény valószínűsége érdekel bennünket, hanem az, hogy bizonyos kísérleti eredmények birtokában a teljes eseményrendszer elemeinek valószínűségeit hogyan lehetne kiszámítani.. Ilyen problémák esetén alkalmazható az alábbi, u.n. **Bayes-tétel**:

2.14. Tétel. Ha $A_1, \ldots, A_n$ teljes eseményrendszer, akkor bármely pozitív valószínűségű B eseményre igazak a

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n)} \qquad (i = 1, \ldots, n)$$

összefüggések.

2.9. Példa. A 2.7. példában szereplő kísérletet elvégeztük, és fekete golyót húztunk. Mi a valószínűsége annak, hogy ez az I. dobozban lévő golyó?

B: fekete golyót választunk.

A_1: az I. dobozból választunk,
A_2: a II. dobozból választunk.

A_1, A_2 teljes esemény rendszer, és

$$P(A_1) = P(A_2) = \frac{1}{2},$$

Most a $P(A_1|B)$ feltételes valószínűség a kérdés.

A Bayes-tétel szerint

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}$$

A 2.8. példa eredményeit felhasználva

$$P(A_1|B) = \frac{\frac{2}{6} \cdot \frac{1}{2}}{\frac{2}{6} \cdot \frac{1}{2} + \frac{2}{10} \cdot \frac{1}{2}} = \frac{5}{8}.$$

3. VALÓSZÍNŰSÉGI VÁLTOZÓK

3.1 A VALÓSZÍNŰSÉGI VÁLTOZÓ FOGALMA

A gyakorlatban előforduló kísérletek túlnyomó többségében a kísérlet végrehajtásakor a bekövetkező elemi eseménnyel egyidejűleg egy vagy több numerikus érték is adódik. Ezeket az értékeket az elemi esemény egyértelműen meghatározza: Valahányszor egy adott elemi esemény bekövetkezik, mindannyiszor ugyanaz a számérték adódik. Ennek megfelelően a kísérlet konkrét végrehajtásától függetlenül is hozzárendelhetjük ezeket a számokat az egyes elemi eseményekhez. Ily módon egy olyan számértékű függvényt kapunk, amelyik az eseménytéren van értelmezve.

3.1. Definíció. Az elemi események halmazán, a H eseménytéren értelmezett valós értékű függvényeket **valószínűségi változók**nak nevezzük.

Jelölésükre a nagy betűket fogjuk használni: X, Y stb. Természetesen ezek mindig függvényeket jelentenek. Ha ezt külön is hangsúlyozni akarjuk, akkor az $X(e)$ írásmódot alkalmazzuk, annak megfelelően, hogy ezek a függvények az elemi események halmazán vannak értelmezve.

A továbbiakban valamely valószínűségi változó értékkészletét K-val fogjuk jelölni.

Az X valószínűségi változóval kapcsolatban különböző eseményekről beszélhetünk. Legyen $A \subseteq R$ tetszőleges halmaz. Ekkor az

$$\{e : e \in H, X(e) \in A\} \subseteq H$$

eseményt $X \in A$- val fogjuk jelölni. Ez azt az eseményt jelenti, hogy az X valószínűségi változó az A halmazbeli értéket vesz fel, vagy más szóval, hogy X értéke az A halmazba esik.

Ha $A \cap K = \emptyset$ (a két halmaznak nincs közös eleme), akkor $X \in A$ a lehetetlen esemény.

Ha $K \subseteq A$, akkor $X \in A$ a biztos esemény.

Ha tetszőleges c valós szám esetén $A = (-\infty, c)$, akkor $X \in (-\infty, c)$ akkor és csakis akkor teljesül, ha $X < c$. Ez pontosan az az esemény, hogy X c-nél kisebb értéket vesz fel. Ezt röviden $X < c$ eseménynek fogjuk mondani.

3.1. Példa. Tekintsük a kockadobás kísérletét. Ekkor $H = \{e_1, \dots, e_6\}$, ahol e_i jelenti azt az elemi eseményt, hogy *i*-t dobunk a kockával $i = 1, \dots, 6$. A kísérlet kimeneteleivel egyúttal adódó legnyilvánvalóbb numerikus érték maga a dobott szám, tehát az

$$X(e_i) = i, \ \ i = 1, \dots, 6.$$

valószínűségi változó. Ennek a valószínűségi változónak a definícióját röviden úgy is szoktuk mondani, hogy az *X* valószínűségi változó jelentse a dobott számot.

A fenti kísérlettel kapcsolatban azonban más valószínűségi változók is értelmezhetők. Tekintsük például a következőképpen definiált *Y* valószínűségi változót:

$$Y = Y(e_i) = \begin{cases} 1 & i = 1,3,5 \\ 2 & i = 2,4,6. \end{cases}$$

Ennek a függvénynek két lehetséges értéke van, az 1 és a 2. A valószínűségi változó az 1 értéket veszi fel, ha páratlan számot dobunk, és a 2 értéket veszi fel, ha páros számot dobunk.

Nézzünk a fenti valószínűségi változókkal néhány eseményt, és azok valószínűségeit.!

1) $X \leq 0$.

$$(X \leq 0) = \emptyset, \text{ mert X legkisebb értéke 1,}$$

$$P(X \leq 0) = 0.$$

2) $X = 5$.

$$(X = 5) = (e_5) = \text{"5-öt dobunk"},$$

$$P(X = 5) = \frac{1}{6}.$$

3) $X \in A$, ahol $A = (2,4,6)$.

$$(X \in A) = (X = 2) + (X = 4) + (X = 6) = (e_2) + (e_4) + (e_6) = \text{"páros számot dobunk"},$$

$$P(X \in A) = P(X = 2) + P(X = 4) + P(X = 6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

4) $Y < 3$.

$$(\mathrm{Y} < 3) = (\mathrm{Y} = 1) + (Y = 2) = (e_1, e_3, e_5) + (e_2, e_4, e_6) = (e_1, e_3, e_5, e_2, e_4, e_6) = H,$$

$$P(\mathrm{Y} < 3) = 1.$$

3.2. AZ ELOSZLÁSFÜGGVÉNY

Legyen X egy valószínűségi változó, és x tetszőleges valós szám. Tekintsük az $X<x$ eseményt. Ennek az eseménynek a valószínűsége nyilván függ az x-től .

3.2. Definíció. Azt a valós számok halmazán értelmezett függvényt, amelyik megadja, hogy az $X<x$ esemény valószínűsége hogyan függ x-től az X valószínűségi változó **eloszlásfüggvény**ének nevezzük:

$$F(x) = P(X < x) \quad x \in R.$$

Az eloszlásfüggvények jelölésére a *F, G ,H* stb betűket fogjuk használni.

3.2.1. Az eloszlásfüggvény tulajdonságai

Az eloszlásfüggvénynek a következő tulajdonságai vannak:

1) Az eloszlásfüggvény monoton nem csökkenő.

Ha $x_1 < x_2$, akkor

$$F(x_1) \le F(x_2).$$

2) Tetszőleges x_0 helyen léteznek a

$$\lim_{x\to x_0-0} F(x) \text{ és } \lim_{x\to x_0+0} F(x)$$

határértékek. Ha az x_0 helyen az F eloszlásfüggvény nem folytonos, akkor

$$F(x_0) = \lim_{x\to x_0-0} F(x).$$

3) $\lim\limits_{x\to-\infty} F(x) = 0$ és $\lim\limits_{x\to\infty} F(x) = 1$.

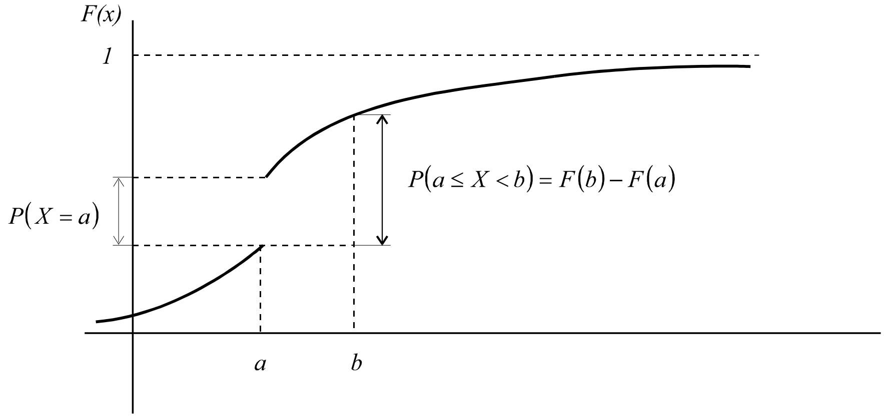
4) Tekintsünk egy tetszőleges $[a,b)$ intervallumot.

$$P(X \in [a,b)) = P(a \le X < b) = F(b) - F(a).$$

5)Tekintsük az $(X = a)$ eseményt, ahol a tetszőleges valós szám.

$$P(X = a) = \begin{cases} 0 & \text{ha } F(x) \text{ az } x = a \text{ helyen folytonos,} \\ \lim\limits_{x\to a+0} F(x) - F(a) & \text{ha } F(x) \text{ az } x = a \text{ helyen nem folytonos.} \end{cases}$$

Tehát, ha az X valószínűségi változó eloszlásfüggvénye folytonos az a helyen, akkor 0 annak a valószínűsége, hogy a változó ezt az értéket veszi fel. Ha pedig az eloszlásfüggvény nem folytonos az a helyen, akkor az a érték felvételének a valószínűségét az adja meg, hogy mekkora az eloszlásfüggvény ugrása ezen a helyen.



3.1. ÁBRA

3.3. ELOSZLÁSOK OSZTÁLYOZÁSA

A gyakorlati alkalmazások szempontjából fontos szerepet játszik a valószínűségi változók két osztálya: a diszkrét, illetve a a folytonos eloszlású változók osztálya.

3.3. Definíció. Az X valószínűségi változót **diszkrét eloszlásúnak**(röviden diszkrétnek) nevezzük, ha lehetséges értékei egy véges, vagy megszámlálhatóan végtelen $x_1, x_2, \ldots$ sorozatot alkotnak.

A $p_k = P(X = x_k)$ valószínűség az x_k érték felvételének a valószínűsége.

Az $(X = x_k)$ $(k = 1, 2, \ldots)$ teljes eseményrendszert alkotnak, amiből következik, hogy

$$\sum_k p_k = 1.$$

Azt mondjuk, hogy az $\{x_n\}$ és $\{p_n\}$ sorozatok az X valószínűségi változó eloszlását alkotják. Ezek ismeretében a valós számok tetszőleges A részhalmaza esetén meghatározható a $P(X \in A)$ valószínűség. Az $(X \in A)$ esemény csak úgy következhet be, ha az X valószínűségi változó olyan értéket vesz fel, amelyik benne van az A halmazban, azaz

$$P(X \in A) = \sum_{\substack{k \\ x_k \in A}} P(X = x_k) = \sum_{\substack{k \\ x_k \in A}} p_k .$$

Speciálisan

$$F(x) = P(X < x) = \sum_{\substack{k \\ x_k < x}} P(X = x_k) = \sum_{\substack{k \\ x_k < x}} p_k .$$

Tehát a $p_k(k = 1,2,\ldots)$ valószínűségek egyértelműen meghatározzák a diszkrét eloszlású változó eloszlásfüggvényét.

Ebben az esetben azt is szoktuk mondani, hogy az $\{x_k\},\{p_k\}$ sorozatok alkotják a valószínűségi változó eloszlását.

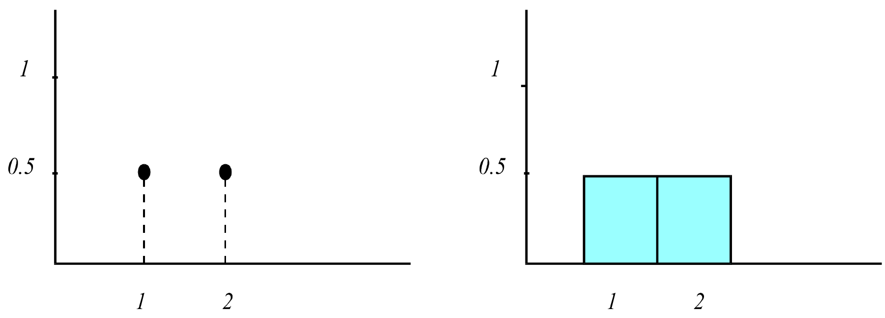
3.2. Példa. Tekintsük azt a kísérletet, hogy egy érmét feldobunk, és definiáljuk X-et a következő módon:

$$X = \begin{cases} 1 & \text{ha fejet dobunk,} \\ 2 & \text{ha ’r‡st dobunk.} \end{cases}$$

X eloszlását a következő táblázattal adhatjuk meg:

x_k	p_k
1□	0.5□
2□	0.5□

A lehetséges grafikus ábrázolások:



3.2. ÁBRA

A 3.2. ábrán a jobboldali ábrát úgy szerkesztjük meg, hogy a lehetséges értékek fölé olyan téglalapokat rajzolunk, amelyeknek területei egyenlők az értékek felvételének valószínűségeivel. Az így kapott ábrát a diszkrét eloszlás **hisztogramjának** nevezzük.

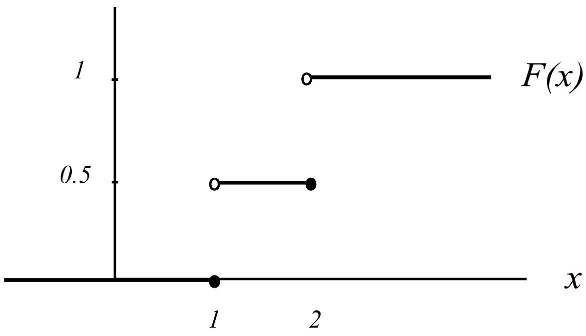
A valószínűségi változó eloszlásfüggvényét pedig a következő meggondolásokkal kaphatjuk.

$$X < x = \begin{cases} \varnothing & x \le 1 \\ X = 1 & 1 < X \le 2 \\ H & 2 < X \end{cases}$$

Ennek megfelelően

$$F(x) = P(X < x) = \begin{cases} 0 & x \le 1 \\ 0.5 & 1 < x \le 2 \\ 1 & 2 < x \end{cases}.$$

A függvényt 3.3 ábrán láthatjuk.



3.3. ÁBRA

Amit a fenti példában tapasztaltunk, az általában is igaz. Egy **diszkrét valószínűségi változó eloszlásfüggvény** un.. **tiszta ugró függvény**, amelynek a valószínűségi változó lehetséges értékeinél vannak szakadásai, ezeken a helyeken függvény annyit ugrik, mint amekkora az adott érték felvételének a valószínűsége. Bármely két lehetséges érték között pedig konstans a függvény.

3.4. Definíció. Az X valószínűségi változót **folytonos eloszlásúnak** röviden folytonosnak) nevezzük, ha létezik olyan, az egész számegyenesen integrálható, nemnegatív $f(x)$ függvény, hogy tetszőleges (véges, vagy végtelen) $[a,b)$ intervallum esetén

$$P(X \in [a,b)) = \int_a^b f(x)dx .$$

Az $f(x)$ függvényt a valószínűségi változó **sűrűségfüggvény**ének nevezzük.

3.3.1. A sűrűségfüggvény tulajdonságai

1) A sűrűségfüggvényt a változó $F(x)$ eloszlásfüggvényét egyértelműen meghatározza.

$$F(x)=P(X<x)=P(X\in(-\infty,x))=\int_{-\infty}^{x}f(t)dt\,.$$

2) A sűrűségfüggvény egész számegyenesen vett integrálja 1.

$$\int_{-\infty}^{\infty}f(x)dx=P(-\infty<X<\infty)=1.$$

3) Folytonos eloszlású valószínűségi változó eloszlásfüggvénye mindenütt folytonos, esetleg véges sok pontot kivéve differenciálható és

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

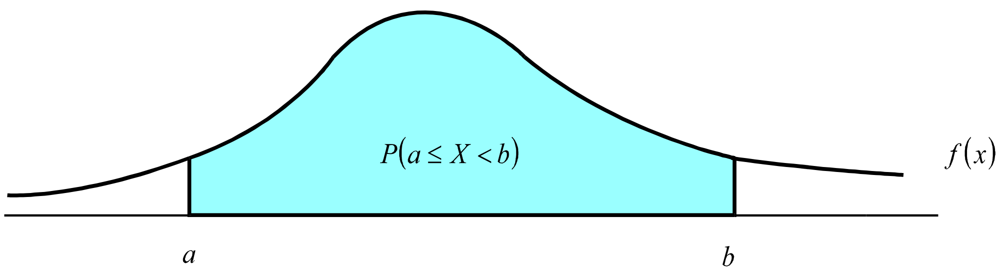
3.3.2. Folytonos eloszlású változók tulajdonságai

Legyen X folytonos eloszlású, $f(x)$ a sűrűségfüggvénye, $F(x)$ pedig az eloszlásfüggvénye.

1) Tetszőleges $x_0\in R$ esetén $P(X=x_0)=0$.

2) $P(X\in[a,b))=P(X\in[a,b])=P(X\in(a,b])=P(X\in(a,b))$,

3) A határozott integrál geometriai jelentésére tekintettel tetszőleges $[a,b]$ intervallumon a sűrűségfüggvény grafikonja alatti terület megadja annak a valószínűségét, hogy a valószínűségi változó az adott intervallumba eső értéket vesz fel. (3.4. ábra)



3.4. ÁBRA

3.3.3. Diszkrét eloszlás közelítése folytonossal

Ha valamely X diszkrét valószínűségi változó lehetséges értékeinek a száma igen nagy, (vagy például, amikor megszámlálhatóan végtelen lehetséges értéke van), akkor sok esetben célszerű un. folytonos közelítést alkalmaznunk, amely a következőt jelenti. Vegyük a kérdéses diszkrét eloszlás hisztogramját, és keressünk hozzá olyan $f(x)$ sűrűségfüggvényt, amelyik jól illeszkedik rá abban az értelemben, hogy tetszőleges $[a, b]$ intervallumon a sűrűségfüggvény alatti terület, és a hisztogram $[a, b]$ intervallumra eső területe jó közelítéssel egyenlő. Ha van ilyen sűrűségfüggvény, akkor azt mondjuk, hogy X közelítőleg folytonos eloszlású, amelynek $f(x)$ a sűrűségfüggvénye. Ilyen esetben

$$\int_a^b f(x)dx \approx P(X \in [a,b)).$$

3.4 ELOSZLÁSOK NUMERIKUS JELLEMZŐI

3.4.1. A várhatóérték

Valamely valószínűségi változó ***várhatóértéke*** egy, a változó eloszlása által meghatározott valós szám, amelyet $E(X)$-szel jelölünk

Diszkrét eset . Ha X lehetséges értékei $x_1, x_2, \ldots$, és ezeket $p_1, p_2, \ldots$ valószínűségekkel veszi fel, akkor

$$E(X) = \sum_k p_k x_k .$$

Folytonos eset. Ha X sűrűségfüggvénye $f(x)$, akkor

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx .$$

Azt, hogy a várhatóértéket miért éppen a fenti formulákkal értelmezzük diszkrét valószínűségi változó esetén szemléltetjük. Az egyszerűség kedvéért tegyük fel, hogy a X valószínűségi változónak véges sok lehetséges értéke van. Legyenek ezek $x_1, x_2, \ldots, x_r$, és $p_1, p_2, \ldots, p_r$ ezek felvételének valószínűségei. Végezzünk n megfigyelést X-re vonatkozóan. Ha a megfigyeléseink során az x_1 k_1 - szer, az x_2 k_2 - ször,..., és az x_r k_r - szer fordult elő ($k_1 + k_2 + \cdots + k_r = n$), akkor a megfigyelt értékek átlaga

$$\frac{k_1 x_1 + k_2 x_2 + \cdots + k_r x_r}{n} = \frac{k_1}{n} x_1 + \frac{k_2}{n} x_2 + \cdots + \frac{k_r}{n} x_r .$$

Mivel a fenti összegben ismételt megfigyeléssorozatok esetén a $\frac{k_i}{n}$ relatív gyakoriság az x_i érték felvételének p_i valószínűsége körül ingadozik ($i = 1,\dots,n$), ezért a megfigyelt értékek átlaga pedig az általunk értelmezett várhatóérték körül fog ingadozni.Az ingadozás mértéke nyilván annál kisebb lesz, minél hosszabb megfigyeléssorozatokat ismétlünk. Általában is megvan a várhatóértéknek ez a szemléletes jelentése: ha az adott valószínűségi változóra vonatkozóan n hosszúságú megfigyeléssorozatokat végzünk, akkor a megfigyelt értékek sorozatonként képzett átlagai a várhatóérték körül fognak ingadozni. Egy mefigyeléssorozatból képzett átlagra azt mondhatjuk, hogy közelítőleg a várhatóértékkel egyenlő, és ez a közelítés annál pontosabb , minél több megfigyelésünk van.

Érdemes külön megvizsgálni azt az esetet, amikor az X valószínűségi változót egy véges sokaságból történő véletlenszerű választással kapcsolatosan értelmezzük. Tegyük fel, hogy van N egyedünk, és ezeknek valamely számértékkel jellemezhető tulajdonságát (például személyek esetén életkor, éves jövedelem, egy átlagos héten fogyasztott tej mennyisége stb) vizsgáljuk. A sokaságból válasszunk ki véletlenszerűen egy egyedet és X jelentse a vizsgált tulajdonság mértékét a kiválasztott egyed esetén. A sokaság összetétele az adott tulajdonság szempontjából legyen a következő:

Érték	Egyedek száma
x_1	n_1
x_2	n_2
.	□
.	
.	
x_r	n_r
Összesen	N

Ekkor X diszkrét valószínűségi változó, amelynek lehetséges értékei $x_1, x_2, \dots, x_r$, és ezek felvételének valószínűségei $p_1 = \frac{n_1}{N}, p_2 = \frac{n_2}{N}, \dots, p_r = \frac{n_r}{N}$. Ennek megfelelően a várhatóértéke

$$E(X) = p_1 x_1 + \dots + p_r x_r = \frac{n_1}{N} x_1 + \dots + \frac{n_r}{N} x_r = \frac{n_1 x_1 + \cdots n_r x_r}{N},$$

ami nem más, mint a vizsgált tulajdonság átlagértéke az N egyedből álló sokaságra nézve.

3.3. Példa. Tegyük fel, hogy van 1000 ember, és ezeknek az életkorát vizsgáljuk. Legyen X egy véletlenszerűen kiválasztott ember életkora. Az 1000 ember életkor szerinti megoszlása legyen a következő:

Életkor	Egyedek száma
18	200
20	300
21	400
22	100
Összesen	1000

$$P(X=18)=\frac{200}{1000},P(X=20)=\frac{300}{1000},P(X=21)=\frac{400}{1000},P(X=22)=\frac{100}{1000}.$$

Ennek megfelelően

$$E(X)=18\cdot\frac{200}{1000}+20\cdot\frac{300}{1000}+21\cdot\frac{400}{1000}+22\cdot\frac{100}{1000}=$$

$$=\frac{18(200)+20(300)+21(400)+22(100)}{1000}=20.20,$$

ami nem más mint az 1000 ember életkorának átlaga.

Most pedig vegyünk 100 hosszúságú független megfigyeléssorozatokat *X*-re vonatkozóan. Egy megfigyeléssorozat azt jelenti, hogy véletlenszerűen kiválasztunk egy embert az 1000 közül, feljegyezzük az életkorát, és ezt ismételjük 100-szor. Négy megfigyeléssorozat eredményét az alábbi táblázatban foglaltuk össze:

	Esetszám			
Életkor	1.sorozat	2.sorozat	3.sorozat	4.sorozar
18	22	30	18	19
20	23	31	33	29
21	39	28	41	40
22	16	11	8	12
Átlag	20.27	19.90	20.21	20.26

(Az egyes sorozatokat számítógépes szimulációval állítottuk elő.)
Láthatjuk, hogy az egyes megfigyeléssorozatok átlagai a valószínűségi változó várhatóértéke (a populáció átlaga) körül ingadoznak.

3.4.1.1. A várhatóérték tulajdonságai

1. Ha $c \in R$, és X egy olyan valószínűségi változó, amelynek c az egyetlen lehetséges értéke, azaz $P(X=c)=1$, akkor

$$E(X)=cP(X=c)=c.$$

Az ilyen valószínűségi változó igazából konstans, mivel az értéke minden elemi eseménynél ugyanaz. A fenti tulajdonságot röviden úgy is szoktuk mondani, hogy "Konstans várhatóértéke önmagával egyenlő".

2. Ha $m_1 \le X \le m_2$,akkor

$$m_1 \le E(X) \le m_2$$

3. Ha a egy valós szám, akkor aX is valószínűségi változó, és

$$E(aX) = aE(X).$$

4. Ha X és Y két valószínűségi változó, akkor

$$E(X+Y) = E(X) + E(Y).$$

3.5. Definíció. Az X és Y valószínűségi változókat függetleneknek nevezzük, ha a valós számok tetszőleges A és B részhalmaza esetén az $X \in A$ és az $X \in B$ események függetlenek.

A fenti definíció gyakorlatilag azt jelenti, hogy a kísérlet elvégzésekor az X változó kapott értéke semmilyen hatással nincs arra, hogy az Y változónak milyen értéke adódik.

5. Ha X és Y független valószínűségi változók, akkor

$$E(X \cdot Y) = E(X)E(Y).$$

3.4.2. A szórás

Mint azt az előzőekben láttuk, a valószínűségi változó várhatóértéke a változó "átlagos" viselkedéséről ad információt. Nem mond viszont semmit arról, hogy az egyes megfigyelt értékek milyen mértékben térnek el ettől a várhatóértéktől, milyen mértékben szóródnak körülötte. A várhatóérték körüli szóródás mérőszáma, a váltózó eloszlásának egy másik numerikus jellemzője a ***szórás***.

Az X változó szórását $D(X)$-szel jelöljük, és definíció szerint

$$D(X) = \sqrt{E\left((X-m)^2\right)},$$

ahol $m = E(X)$.

A szórás négyzetét, azaz a $E\left((X-m)^2\right)$ mennyiséget **szórásnégyzetnek** vagy **varianciának** nevezzük. Ennek megfelelően a szokásos jelölései: $D^2(X)$ vagy $V(X)$.

3.4.2.1. A szórás tulajdonságai

1. Ha $c \in R$, és X egy olyan valószínűségi változó, amelynek c az egyetlen lehetséges értéke, azaz $P(X = c) = 1$, akkor

$$D(X) = 0.$$

2. Ha $c \in R$, akkor

$$D(cX) = |c|D(X)$$

3. Ha *X* és *Y* független valószínűségi változók, akkor

$$D^2(X+Y) = D^2(X) + D^2(Y).$$

4. $D^2(X) = E(X^2) - m^2$, ahol *m* a valószínűségi változó várhatóértéke.

3.4.2.2. Szórás számítása az eloszlás ismeretében

Diszkrét eset. Legyenek az *X* valószínűségi változó lehetséges értékei $x_1, x_2, \ldots,$ és ezek felvételének valószínűségei rendre $p_1, p_2, \ldots$

$$D^2(X) = E(X^2) - m^2 = \sum_i x_i^2 p_i - \left(\sum_i x_i p_i\right)^2.$$

Folytonos eset. Legyen $f(x)$ a *X* valószínűségi változó sűrűségfüggvénye. Bebizonyítható, hogy

$$D^2(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - m^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \left(\int_{-\infty}^{\infty} xf(x)dx\right)^2.$$

3.4.3. A medián

Valamely *X* valószínűségi változó mediánját a következőképpen értelmezzük. Legyen $F(x)$ a valószínűségi változó eloszlásfüggvénye.

Azt az m_e számot ,amelyre az

$$P(X < \mathrm{m_e}) \le \frac{1}{2}, \quad P(X > \mathrm{m_e}) \le \frac{1}{2}$$

egyenlőtlenségek teljesülnek, az *X* valószínűségi változó ***mediánjának*** nevezzük.

Ha az

$$F(x) = \frac{1}{2}$$

egyenletnek egy megoldása van, akkor ez egyenlő a valószínűségi változó, illetve az eloszlás mediánjával.

Ha az eloszlásfüggvény mindenütt folytonos és szigorúan monoton növekvő, akkor a fenti egyenletnek csak egy megoldás van. Ekkor a mediánra az jellemző, hogy

$$P(X < m_e) = F(m_e) = \frac{1}{2}.$$

Továbbá

$$P(X > m_e) = 1 - P(X \le m_e) = 1 - P(X < m_e)1 - \frac{1}{2} = \frac{1}{2}.$$

Diszkrét eloszlású változó esete. Legyenek $x_1, x_2, \ldots$ a valószínűségi változó lehetséges értékei, és tegyük fel, hogy ez egyúttal nagyság szerint növekvő sorrendet is jelent.Ekkor az eloszlásfüggvény egy tiszta ugró függvény, amelynek az $x_1, x_2, \ldots$ helyeken vannak az ugrásai.

a) Ha az egyenletnek van megoldása, akkor létezik egy olyan x_i, hogy

$F(x) = \frac{1}{2}$ az egész $(x_i, x_{i+1}]$ intervallumin teljesül, vagyis az egyenletnek végtelen sok megoldása van. Ekkor az egyértelműség kedvéért, megállapodás szerint ennek az intervallumnak a középpontját nevezzük mediánnak:

$$m_e = \frac{x_i + x_{i+1}}{2}$$

A folytonos esetben tapasztalt tulajdonság most is érvényes:

$$P(X < m_e) = F(m_e) = \frac{1}{2},$$

és

$$P(X > m_e) = 1 - P(X \le m_e) = 1 - (P(X < m_e) + P(X = m_e)) = 1 - \frac{1}{2} - 0 = \frac{1}{2}.$$

b) Az egyenletnek nincs megoldása, azaz nincs olyan hely, ahol az eloszlásfüggvény az 1/2 értéket veszi fel, az eloszlásfüggvény "átugorja" az $y = 1/2$ egyenest. Ez azt jelenti, hogy van olyan x_i lehetséges értéke a valószínűségi változónak, hogy

$$F(x_i) < \frac{1}{2} \quad \text{és} \quad \lim_{x \to x_i + 0} F(x) > \frac{1}{2},$$

azaz
$$P(X < x_i) < \frac{1}{2}$$

és

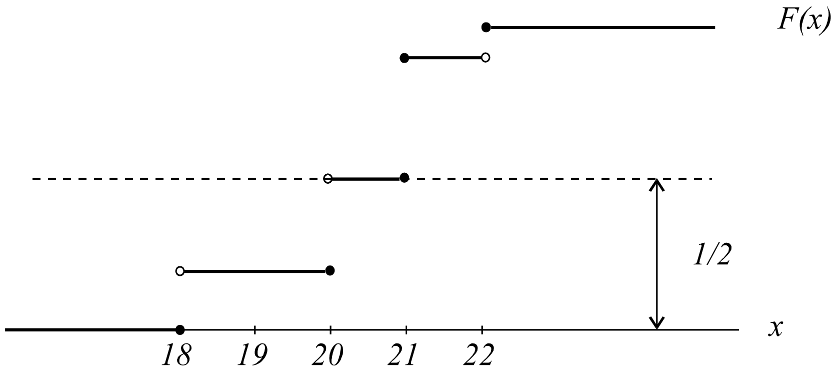
$$P(X > x_i) = 1 - P(X \le x_i) = 1 - \lim_{x \to x_i + 0} F(x) < \frac{1}{2}.$$

Ebben az esetben az x_i-t nevezzük az X valószínűségi változó mediánjának: $m_e = x_i$.

3.4. Példa. Tekintsük a 3.3. példában definiált X valószínűségi változót. X eloszlása

x_i	18	20	21	22
p_i	0.2	0.3	0.4	0.1

Tehát X eloszlásfüggvénye



A fenti ábrából látható, hogy az

$$F(x) = \frac{1}{2}$$

egyenletnek minden, a $(20,21]$ intervallumba eső valós szám megoldása, tehát az X mediánja,

$$m_e = 20.5.$$

Nézzük meg még, hogy a vizsgált sokaságban mi az életkor mediánja. Ennek meghatározásához a sokaság egyedeit életkor szerint növekvő sorrendbe rendezzük, és venni kell a két középső egyed életkorának átlagát, mivel a sokaság páros számú egyedből áll.

Sorsz.	1.	...	200.	201	...	500.	501.	...	900.	901.	...	1000.
Kor	18	...	18	20	...	20	21	...	21	22	...	22

Az életkor szerint növekvő sorrendben az 500. 20 éves, az 501. pedig 21 éves, tehát a sokaságban az életkor mediánja 20.5, ami megegyezik az általunk vizsgált valószínűségi változó mediánjával.

3.4.4. Kvartilisek, módusz

A mediánhoz teljesen hasonló módon értelmezhetjük a k_a alsó kvartilist és a k_f felső kvartilist.

Ha az

$$F(x) = \frac{1}{4} \quad \text{és} \quad F(x) = \frac{3}{4}$$

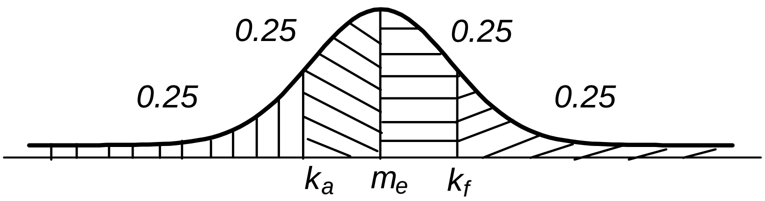
egyenleteknek van egyértelmű megoldásuk, akkor ezeket az eloszlás **alsó**, illetve **felső kvartilisének** nevezzük. Ha az egyenleteknek nincs megoldásuk, vagy ha a megoldások egy egész intervallumot tesznek ki, akkor a mediánhoz hasonló módon értelmezzük a kvartiliseket.

Ezek különbségét, a

$$k_f - k_a$$

értéket **interkvartilis terjedelemnek nevezzük.**

A kvartilisek és a medián a valószínűségi változó értékkészletét négy egyenlő valószínűségű részre osztják , ahogy ez egy folytonos eloszlás esetén a 3.1. ábrán is látható.



3.1. ÁBRA

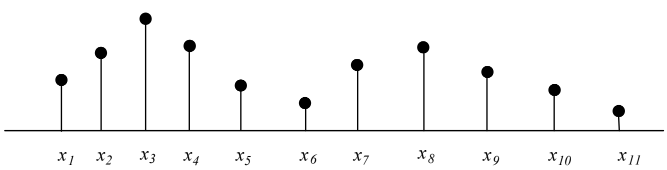
A módusz definíciójánál külön kezeljük a diszkrét és a folytonos eloszlású változozók esetét.

Folytonos eloszlás esetén az eloszlás móduszai a sűrűségfüggvény lokális maximumhelyei. Aszerint, hogy a sűrűségfüggvénynek egy, két, stb lokális maximuma van, beszélünk unimodális, bimodális stb eloszlásokról. A gyakorlatban előforduló legfontosabb folytonos eloszlások unimodálisak, azaz egy móduszúak.

Diszkrét esetben először a lehetséges értékeket rendezzük nagyság szerint növekvő sorrendbe, azaz feltesszük, hogy a lehetséges értékek $x_1, x_2, \ldots$ sorozata nagyság szerint növekvő sorrendet jelent. Tekintsük ennek megfelelően a valószínűségek

$$p_k = P(X = x_k), k = 1,2,\ldots$$

sorozatát. Ezen sorozat lokális maximumainak felelnek meg az eloszlás móduszai.. A móduszokra az m_d jelölést használjuk. (ld. a 3.2.ábrát.)



3.2. ÁBRA

Ebben az esetben az eloszlásnak két módusza van $m_d = x_3$ és $m_d = x_8$.

3.4.5. Szimmetrikusság, ferdeség

Az X valószínűségi változó várhatóértékét jelöljük m-mel, és eloszlásfüggvénye legyen $F(x)$.

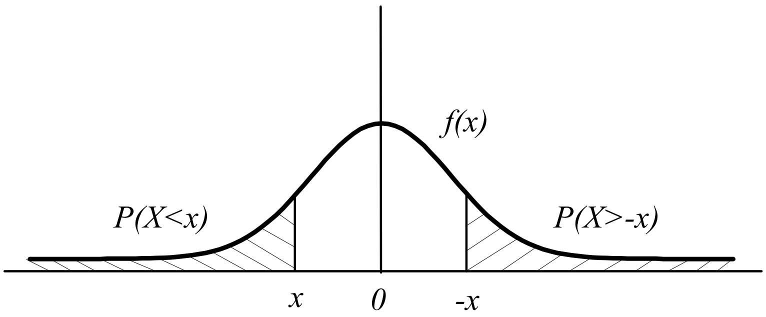
Az X eloszlását szimmetrikusnak nevezzük, ha az X-m és m-X valószínűségi változók eloszlása megegyezik, azaz ha tetszőleges $x \in R$ esetén

$$P(X - m < x) = P(m - X < x).$$

Ha az X változó folytonos eloszlású, akkor ez azt jelenti, hogy

$$F(m + x) = 1 - F(m - x)$$

és hogy a sűrűségfüggvénye szimmetrikus az $x = m$ egyenesre. $m = 0$ esetben a sűrűségfüggvény páros függvény.

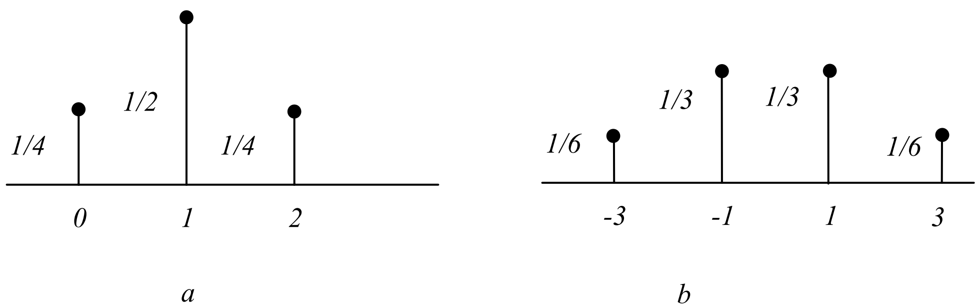


3.3. ÁBRA

A 3.3. ábrán egy páros sűrűségfüggvényt láthatunk. Geometriai okokból nyilvánvaló, hogy a satírozott területek megegyeznek, azaz

$$P(X < x) = P(X > -x) = P(-X < x)$$

Diszkrét eloszlások esetén a szimmetrikusságot a 3.4. ábrán szemléltetjük.



3.4. ÁBRA

Tekintsük először az . *a)* esetet.

$$E(X) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Könnyen látható, hogy az $X-1$ és az $1-X$ valószínűségi változók eloszlása megegyezik:

$$X-1 = \begin{cases} -1 \\ 0 \\ 1 \end{cases} \Leftrightarrow \begin{matrix} X=0 & P_{-1} = 1/4 \\ X=1 & P_0 = 1/2 \\ X=2 & P_1 = 1/4 \end{matrix} ,$$

és

$$1-X = \begin{cases} -1 \\ 0 \\ 1 \end{cases} \Leftrightarrow \begin{matrix} X=2 & P_{-1} = 1/4 \\ X=1 & P_0 = 1/2 \\ X=0 & P_1 = 1/4 \end{matrix} .$$

A *b)* esetben

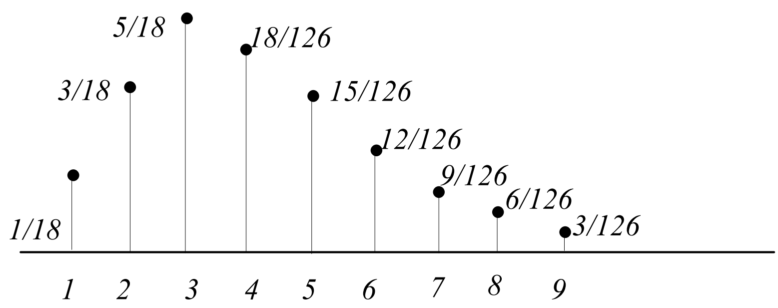
$$E(X) = (-3) \cdot \frac{1}{6} + (-1) \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 3 \cdot \frac{1}{6} = -\frac{1}{2} - \frac{1}{3} + \frac{1}{3} + \frac{1}{2} = 0 .$$

Az előzőhöz teljesen hasonló módon meggyőződhetünk róla, hogy az X és a $-X$ változók eloszlása megegyezik.

Bebizonyítható, hogy szimmetrikus eloszlások esetén a várhatóérték és a medián egyenlők. Ha a szimmetrikus eloszlásnak egyetlen módusza van, akkor mind a három numerikus jellemző egybeesik:

$$m = m_e = m_d .$$

Ha valamely valószínűségi változó eloszlása nem szimmetrikus, akkor **ferdének** nevezzük. A gyakorlatban gyakran találkozunk olyan egy móduszú eloszlásokkal, amelyek az jellemző, hogy a módusz egyik oldalán (tőle jobbra vagy balra) "hosszan elnyúlnak", a másik irányban pedig "röviden lefutnak". Ilyen eloszlásokat



3.5. ÁBRA

szemléltetünk a 3.5. és 3.6. ábrákon.

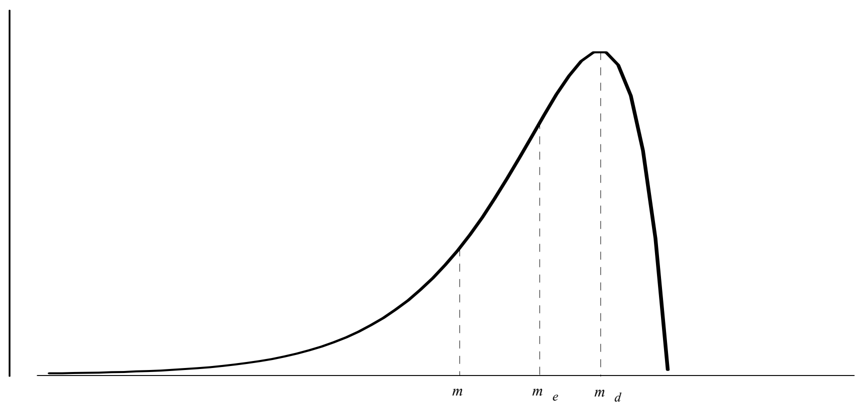
A 3.5. ábrán látható diszkrét eloszlást pozitívan ferdének mondjuk, mert a az eloszlás hosszan elnyúló része (az eloszlás hosszú farka) a módusztól jobbra, pozitív irányba esik. Az eloszlás egyetlen módusza $m_d = 3$. Mivel az eloszlásfüggvény az egész $(3,4]$ intervallumon $1/2$-del egyenlő, ezért a medián $m_e = 3.5$. Az eloszlás várhatóértéke pedig két tizedesre kerekítve

$$m = 1\cdot\frac{1}{18} + 2\cdot\frac{3}{18} + 3\cdot\frac{5}{18} + 4\cdot\frac{18}{126} + 5\cdot\frac{15}{126} + 6\cdot\frac{12}{126} + 7\cdot\frac{9}{126} +$$

$$+8\cdot\frac{6}{126} + 9\cdot\frac{3}{126} = 4.06$$

Általában is igaz, hogy pozitívan ferde eloszlások esetén

$$m_e < m$$



3.6. ÁBRA

A 3.6. ábrán egy negatív irányban ferde folytonos eloszlás sűrűségfüggvénye látható. Negatívan ferde eloszlások esetén $m < m_e$.

Az eloszlás ferdeségének jelzésére a

$$\gamma = \frac{E\left((X-m)^3\right)}{\left(D(X)\right)^3}$$

un. **ferdeségi együttható** szolgál. Szimmetrikus eloszlás esetén $\gamma = 0$. Ha $\gamma > 0$, akkor az eloszlás pozitívan ferde, ha $\gamma < 0$, akkor pedig negatívan. Minél nagyobb γ abszolút értéke, annál erősebben ferde az eloszlás.

Az eloszlás például pozitív irányú ferdesége azt jelenti, hogy az $(X-m)^3$ különbség várhatóértéke pozitív, ami abból adódik, hogy a változónak a várhatóértéktől való pozitív irányú eltérései valószínűbbek, mint az ugyanolyan nagyságú, de negatív irányúak. A negatív irányú ferdeség jelentése hasonló.

3.5 Kockázati döntések

Kockázati döntésről beszélünk, ha

1. két vagy több tevékenység közül kell választanom;
2. a döntés meghozatalakor nem ismerem egyértelműen, hogy a választott tevékenységet milyen (gazdasági) körülmények között kell majd végeznünk;
3. a különböző körülmények között az egyes tevékenységekkel más-más hasznot tudok elérni;
4. a döntéskor számba tudom venni, hogy elvileg milyen körülmények lehetségesek, és ismerem ezek előfordulási valószínűségeit.

Minden döntést illetően az elérhető haszon tekintetében van némi bizonytalanság. A döntés meghozatalakor az egyes tevékenységekkel elérhető haszon valószínűségi változó, amelynek ismerem az eloszlását. Ennek a valószínűségi változónak a várhatóértékét **várható haszonnak** nevezzük. Kockázati döntések esetén a **maximális várható haszon** elvén döntünk, ha azt a tevékenységet választjuk, amely esetén a várható haszon a legnagyobb.

A maximális várható haszon elvén történő döntéskor nem vagyunk tekintettel az egyes döntések kockázatára. A döntések kockázatát a haszon szórásával mérjük. Másik lehetséges döntési elv a **minimális kockázat** elve: azt a tevékenységet választjuk, amely esetén a haszon szórása a legkisebb. A két különböző elven meghozott döntés nem feltétlenül esik egybe.

Tegyük fel, hogy négy különböző tevékenység (T1, T2, T3 és T4) közül kell választanom. A tevékenység végrehajtásakor három különböző körülménnyel kell számolnom: K1, K2, K3, és ezek előfordulási valószínűségei rendre $p_1,\ p_2,\ p_3$. $H_1, H_2\ H_3, H_4$ valószínűségi változók az egyes tevékenységből származó hasznot jelentik. Ezek valószínűségi változók, mert például a T1 tevékenységből származó hasznom attól függően más és más lesz, hogy a lehetséges körülmények közül éppen mi következik be. Ezeknek a valószínűségi változóknak az eloszlásait és várhatóértékeit tartalmazza az alábbi táblázat.

	Körülmények			
	K1	K2	K3	
	Valószínűségeik			
Tevékenységek	p_1	p_2	p_3	**Várható haszon**
T1	h_{11}	h_{12}	h_{13}	$E(H_1) = h_{11}p_1 + h_{12}p_2 + h_{13}p_3$
T2	h_{21}	h_{22}	h_{23}	$E(H_2) = h_{21}p_1 + h_{22}p_2 + h_{23}p_3$
T3	h_{31}	h_{32}	h_{33}	$E(H_3) = h_{31}p_1 + h_{32}p_2 + h_{33}p_3$
T4	h_{41}	h_{42}	h_{43}	$E(H_4) = h_{41}p_1 + h_{42}p_2 + h_{43}p_3$

A maximális várható haszon elvén alapuló döntésnél azt a tevékenységet kell választanom, amely esetén a haszon várhatóértéke a legnagyobb.

Valamilyen módon mérni kellene, hogy az egyes döntések milyen kockázattal járnak. A haszon várhatóértéke, mint döntési kritérium hosszú távon a legnagyobb átlagos hasznot eredményezi. Sok esetben a döntéshozók a rövid távú érdekeket is figyelembe veszik. Ilyenkor fel kell mérniük a döntéseik kockázatát is. A várható érték kritériumnál a kockázatot nem vesszük figyelembe

A kockázat egy lehetséges mérőszáma a változó terjedelme, ami nem más, mint a változó legnagyobb és legkisebb értékének különbsége. E mérőszám szerint, annak a döntésnek nagyobb a kockázata, amely esetén a haszon változó terjedelme nagyobb.

A kockázatnak a változó terjedelménél jobb mérőszáma a haszon szórása, mivel a szórósnál azt is figyelembe vesszük, hogy az értékkészlet egyes értékei milyen valószínűséggel fordulhatnak elő.

Lehetséges olyan döntés is, amelynél a kockázatot akarjuk minimalizálni. A **minimális kockázat elve** alapján ha döntünk, akkor azt a tevékenységet kell választani, amelyik esetén a haszon változó szórása a legkisebb.

Általában a maximális várható haszon elvén hozott döntés és minimális kockázat alapján hozott döntés nem azonos. Sokszor a döntés milyensége nagyban függ a döntéshozó habitusától. Egy konzervatívabb döntéshozó, vagy szervezet valószínű, hogy inkább a kockázat minimalizálására törekszik és inkább a kisebb várható nyereséget fogja választani, ha annak kisebb a kockázata. A "szerencsejátékos" típusú döntéshozó, pedig éppen ellenkezőleg.

3.5. Példa. Egy adott összeget akarunk befektetni. Három befektetési lehetőség közül választhatunk: részvény, hötvény, vagy ingatlan. A gazdasági körülményektől függően az egyes befektetési formáknak mások lesznek a hozamaik. A gazdasági előrejelzések szerint a gazdasági körülményeket illetően lassú, normál illetve gyors növekedés lehetséges, és ezek valószínűségei rendre 0.3, 1.5 illetve 0.2. A befektetendő összegünk hozamai(valamely egységben kifejezve) az egyes esetekben a következőképpen alakulnak:

	Gazdasági körülmények		
Befektetés	Lassú növekedés	Normál növekedáés	Gyors növekedés
Részvény	-100	70	120
Kötvény	40	50	90
IIngatlan	-150	40	180

A negatív hozamok veszteségeket jelentenek. Az a kérdés, hogy milyen befektetés mellett kell döntenünk, hogy a várható hozam maximális legyen. Határozzuk meg az egyes befektetések kockázatát is.

Bármit döntünk is, a befektetés hozama valószínûségi változó, mert értéke függ attól, hogy milyenek lesznek a gazdasági körülmények. Legyen X_R, X_K és X_I a befektetésünk hozama részvény, kötvény, illetve ingatla vásárlás esetén. Mindhárom valószínûségi változó eloszlása ismert, ezek várhatóértéke a várható hozam az egyes döntéseink esetén. Ha maximális várható hozamra törekszünk, akkor azt a befektetési formát kell választanunk, amely esetén ez a várhatóérték a legnagyobb. A döntésünk kockázatát pedig ezen változók szórása méri. Meg kell tehát határoznunk mindhárom valószínûségi változó várhatóértékét és szórását.

Mindhárom változónak 3 lehetséges értéke van, és az ezekhez tartozó valószínûségek mindhárom esetben ugyanazok.

p_i	X_R	X_K	X_I	X_R^2	X_K^2	X_I^2
0.3	-100	40	-150	10000	1600	225000
0.5	70	50	40	4900	2500	1600
0.2	120	90	180	144000	8100	324000
$E(\bullet)$	29	**55**	11	8330	3350	14030
$D(\bullet)$	86.54	18.03	117.94			

$$E(X_R) = 0.3 \cdot (-100) + 0.5 \cdot 70 + 0.2 \cdot 120 = 29,$$

$$E(X_K) = 0.3 \cdot 40 + 0.5 \cdot 50 + 0.2 \cdot 90 = 55,$$

$$E(X_I) = 0.3 \cdot (-150) + 0.5 \cdot 40 + 0.2 \cdot 180 = 11,$$

$$E\left(X_R^2\right) = 0.3 \cdot 1000 + 0.5 \cdot 4900 + 0.2 \cdot 14400 = 8330,$$

$$E\left(X_K^2\right) = 0.3 \cdot 1600 + 0.5 \cdot 2500 + 0.2 \cdot 8100 = 3350,$$

$$E\left(X_I^2\right) = 0.3 \cdot 22500 + 0.5 \cdot 1600 + 0.2 \cdot 32400 = 14030$$

$$D(X_R) = \sqrt{E\left(X_R^2\right) - (E(X_R))^2} = \sqrt{8330 - 29^2} = \sqrt{7489} = 86.54,$$

$$D(X_K) = \sqrt{E\left(X_K^2\right) - (E(X_K))^2} = \sqrt{3350 - 55^2} = \sqrt{325} = 18.03,$$

$$D(X_I) = \sqrt{E(X_I^2) - (E(X_I))^2} = \sqrt{14030 - 11^2} = \sqrt{13909} = 117.94\,.$$

A várható hozamunk akkor lesz a legnagyobb, ha kötvényt vásárolunk. Most egyúttal ennek a legkisebb a kockázata is.

4. NEVEZETES DISZKRÉT ELOSZLÁSOK

4.1 BINOMIÁLIS ELOSZLÁS

4.1. Definíció. A X változót *n,p* (*n* természetes szám, 0<*p*<1) paraméterű **binomiális eloszlásúnak** nevezzük, ha lehetséges értékei 0,1,2,..., *n*, és

$$p_k = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \left(k = 0,1,\ldots,n\right)$$

Rövid jelölése: $X \sim B(n, p)$.

Az alábbiakban példát fogunk nézni binomiális eloszlású változóra.

Legyen adott egy $p > 0$ valószínűségű A esemény. Végezzünk azonos módon, egymástól függetlenül n kísérletet az A bekövetkezésére vonatkozóan. Minden esetben csak az érdekel bennünket, hogy az A bekövetkezett-e vagy sem. Az ilyen kisérletsorozatot **Bernoulli kisérletsorozatnak** nevezzük. A Bernoulli kisérletsorozattal kapcsolatban definiáljuk az X valószínűség változatát a következő módon:

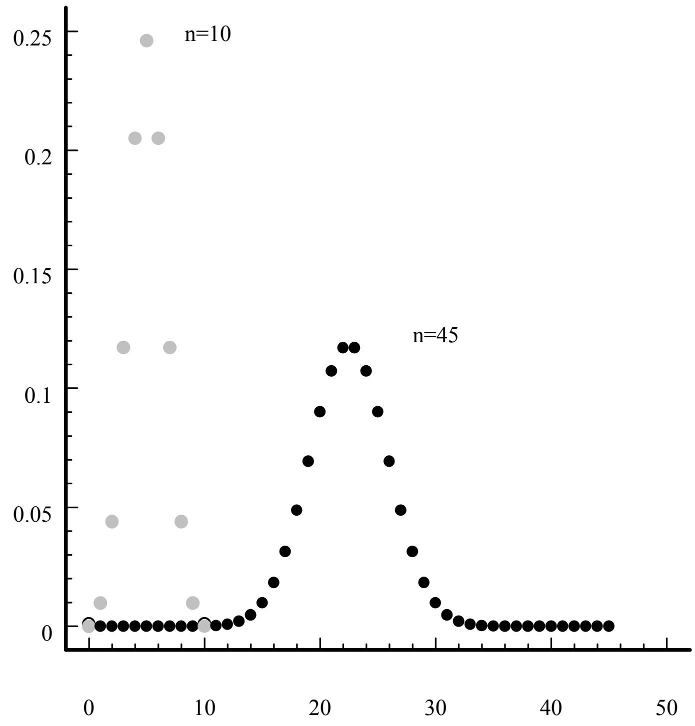
X = az A bekövetkezéseinek ("sikerek") száma az n ismétlés során.

X nyilván egy diszkrét valószínűségi változó, amelynek lehetséges értékei:

$$x_0 = 0, x_1 = 1, \ldots, x_n = n \,,$$

és ezek felvételének valószínűségei a 2.12. Tétel szerint:

$$p_k = P\left(X = k\right) = \binom{n}{k} p^k (1-p)^{n-k} \; k = 0,1,\ldots,n \;\; ,$$



4.1. ÁBRA

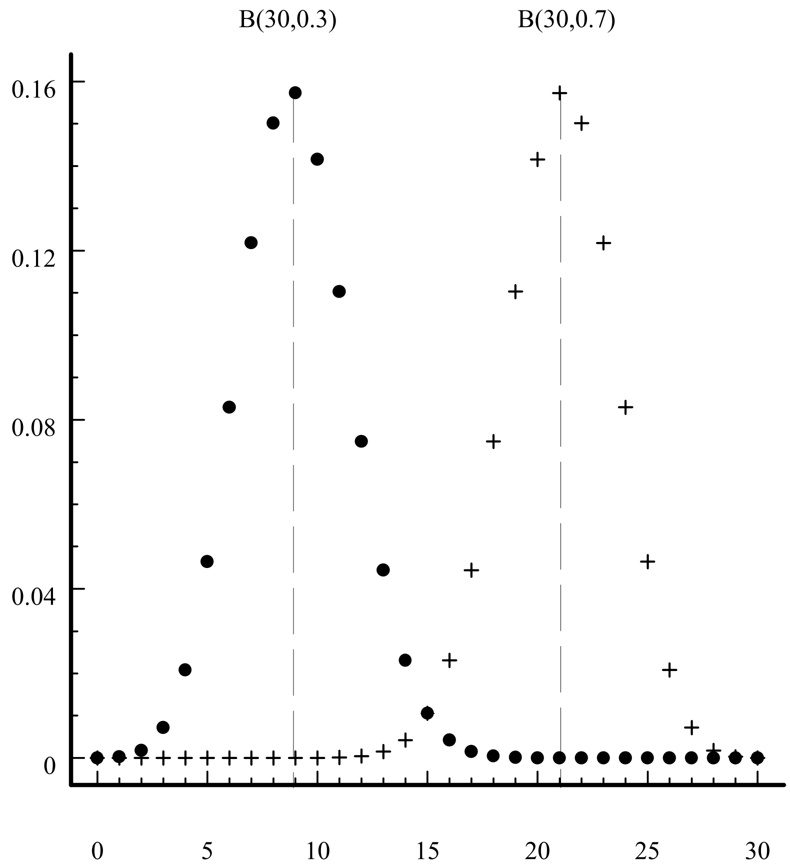
A 4.1. ábrán a $B(10{,}0.5)$ és $B(45{,}0.5)$ eloszlásokat ábrázoltuk

Látható, és egyébként egyszerű számolással igazolható, hogy p=0.5 esetén az eloszlások szimmetrikusak.

Ha n páros, akkor a változónak páratlan számú lehetséges értéke van, és ezek közül a középső felvételének a legnagyobb a valószínűsége.

Ha n páratlan, akkor a változónak páros számú lehetséges értéke van, és ezek közül a középső kettő felvételének a legnagyobb a valószínűsége, és ezek a valószínűségek egymással egyenlők.

Ha $p \neq 0.5$, akkor az eloszlások nem szimmetrikusak



4.2. ÁBRA

A $B(30,0.3)$ és a $B(30,0.7)$ eloszlások láthatók a 4.2. ábrán. Az első esetben a 9-es, a második esetben a 21-es érték felvételének a legnagyobb a valószínűsége. Ez például azt jelenti, hogy ha egy 0.3 valószínűségű A esemény bekövetkezésére vonatkozóan 30 független kísérletet végzünk, akkor az a legvalószínűbb, hogy a A esemény 9-szer következik be. Ha viszont az A esemény valószínűsége 0.7, akkor az a legvalószínűbb, hogy a 30 kísérlet közül 21-ben következik be az A esemény.

Egyszerű számolással meggyőződhetünk róla, hogy általában

$$p_{k-1} \leq p_k$$

teljesül, ha

$$k \leq (n+1)p \,,$$

azaz $B(n,p)$ eloszlás esetén a $[(n+1)p]$ érték felvételének a legnagyobb a valószínűsége. Ha $(n+1)p$ egész, akkor két maximális van a valószínűségek között: $p_{(n+1)p-1} = p_{(n+1)p}$ a legnagyobbak.

Ha X $B(n,p)$ eloszlású valószínűségi változó,akkor megmutatható, hogy

$$E(X) = np,$$

és

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

4.1. Példa. Tízszer feldobunk egy kockát.
a)Mi a valószínűsége annak, hogy négyszer fogunk 6-ot dobni?
b) Mi a hatos dobások számának legvalószínűbb értéke?
c) Mi a hatos dobások számának várhatóértéke?

Jelentse X a 10-es dobássorozatban a hatos dobások számát. A X valószínûségi változó B(10, 1/6) eloszlású.

a) Meg kell határoznunk az X=4 esemény valószínűségét.

$$P(X=4) = \binom{10}{4}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^{10-4} = 0.054266.$$

b) A B(10, 1/6) eloszlás esetén a $\left[(10+1)\frac{1}{6}\right] = \left[\frac{11}{6}\right] = 1$ érték a legvalószínűbb, tehát az a legvalószínűbb, hogy a tíz dobás során egyszer fogunk 6-ot dobni.

c)

$$E(X) = 10 \cdot \frac{1}{6} = \frac{5}{3},$$

vagyis a 10-es dobássorozatban a 6-os dobások számának várhatóértéke 5/3. Ez azt jelent, hogy ha sok 10-es dobássorozatot végzünk, akkor az átlagosan egy sorozatra jutó 6-os dobás közelítőleg 5/3 lesz. A közelítés annál pontosabb lesz minél több dobássorozatot végzünk.

4.1.1. A binomiális eloszlás és a visszatevéses mintavételezés

Legyen $p100\%$ egy gyártmánytételben a selejtarány. Vegyünk ebből a tételben egy n elemű mintát visszatevéssel úgy, hogy a soron következő mintaelem kiválasztásánál a tétel mindegyik elemét azonos valószínűséggel választhatjuk. S legyen az az esemény, hogy a kiválasztott gyártmány selejt. Ekkor $P(S) = p$, mivel minden egyes kiválasztásnál a teljes tételből választunk. N_S legyen az n kiválasztott között a selejtek száma, és R_S a selejtarány. Ekkor a N_S valószínűségi változó $B(n, p)$ eloszlású, mert éppen azt mutatja, hogy az n független ismétlés során a S esemény hányszor következett be, Tehát a mintában lévő selejtek számának várhatóértéke és szórása

$$E(N_S) = np, D(N_S) = \sqrt{np(1-p)}.$$

Mivel

$$R_S = \frac{N_S}{n},$$

$$E(R_S) = E\left(\frac{N_S}{n}\right) = \frac{1}{n}E(N_S) = \frac{1}{n}np = p,$$

és

$$D^2(R_S) = D^2\left(\frac{N_S}{n}\right) = \frac{1}{n^2}D^2(N_S) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n},$$

$$D(R_S) = \sqrt{\frac{p(1-p)}{n}}.$$

4.2. Példa. Egy kávécsomagokból álló gyártmánytételben a súlyhiányos csomagok aránya 5%. A tételből visszatevéssel 20 elemű mintát veszünk. Mi annak a valószínűsége, hogy a kiválasztott mintábna a súlyhiányos csomagok aránya nem haladja meg a 10%-ot?

Jelentse X a kiválasztottak között a súlyhiányosak számát. Az X 20 és 0.05 paraméterű binomiális eloszlású valószínûségi változó, és az $\frac{X}{20} \le 0.1$ esemény valószínûségére vagyunk kíváncsiak. Nyilván

$$\left(\frac{X}{20} \le 0.1\right) = (X \le 2) = (X = 0) + (X = 1) + (X = 2),$$

tehát

$$P\left(\frac{X}{20} \le 0.1\right) = P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2).$$

A fenti valószínűséget numerikusan például a következő két módon határozhatjuk meg.

a)

$$P(X = 0) = \binom{20}{0}0.05^0 \cdot 0.95^{20} = 0.05^0 \cdot 0.95^{20} = 0.358486,$$

$$P(X = 1) = \binom{20}{1}0.05^1 \cdot 0.95^{19} = 20 \cdot 0.05^1 \cdot 0.95^{19} = 0.377354,$$

$$P(X = 2) = \binom{20}{2}0.05^2 \cdot 0.95^{18} = 190 \cdot 0.05^2 \cdot 0.95^{18} = 0.188677,$$

Vagyis

$$P\left(\frac{X}{20} \le 0.1\right) = 0.924517.$$

4.24 A HIPERGEOMETRIKUS ELOSZLÁS

4.2. Definíció. Legyenek *M,N,n* pozitív egész számok, *n,M<N.* Az X valószínűségi változót (*M,N,n* paraméterű) **hipergeometrikus eloszlásúnak** nevezzük, ha lehetséges értékei 0,1,...,n és

$$p_k = P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \quad (k = 0,...,n)$$

Megmutatható, hogy

$$E(X) = n\frac{M}{N},$$

és

$$D(X) = \sqrt{n\frac{M}{N}\left(1-\frac{M}{N}\right)\left(1-\frac{n-1}{N-1}\right)}.$$

4.2.1. A hipergeometrikus eloszlás és a visszatevés nélküli mintavételezés

Legyen adott egy N gyártmányból álló tétel, amelyben S a selejtek száma, vagyis a selejtarány $p = S/N$. Ebből a tételből visszatevés nélkül válasszunk ki *n* gyártmányt úgy, hogy bármely *n* gyártmány kiválasztásának ugyanaz a valószínűsége. Legyen N_S az *n* kiválasztott között a selejtek száma, és R_S pedig a selejtarány.

Az N_S ilyen visszatevés nélküli mintavételezéskor *N,S,n* paraméterű hipergeometrikus eloszlású

A lehetséges értékei nyilván 0,1,...,*n*, és az $N_S = k$ $(k = 0,1,\ldots,n)$ esemény valószínűségét a klasszikus képlettel határozhatjuk meg. Az összes esetek száma $\binom{N}{n}$, mivel a *N* gyártmányból ennyi különböző módon lehetséges *n*-et kiválasztani.

Az $N_S = k$ esemény szempontjából kedvező minden olyan eset, amikor *k* gyártmányt a *S* számú selejt közül választunk, és a többi *n-k* gyártmányt pedig az *N-S* jó gyártmány közül. Tehát a kedvező esetek száma $\binom{S}{k}\binom{N-S}{n-k}$, azaz

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

A mintában lévő selejtek számának, illetve a mintabeli selejtarány várhatóértéke

$$E(N_S) = n\frac{S}{N} = np,$$

$$E(R_S) = E\left(\frac{N_S}{n}\right) = \frac{1}{n}E(N_S) = \frac{1}{n}np = p,$$

ugyanaz, mint visszatevéses mintavételezésnél.

A szórásokban azonban lesz különbség a két mintavételezés között, ugyanis megmutatható, hogy

$$D(N_S) = \sqrt{np(1-p)\left(1-\frac{n-1}{N-1}\right)},$$

és így

$$D(R_S) = D\left(\frac{N_S}{n}\right) = \frac{1}{n}D(N_S) = \sqrt{\frac{p(1-p)}{n}\left(1-\frac{n-1}{N-1}\right)}.$$

Tekintsünk egyre nagyobb és nagyobb egyedszámú tételeket úgy, hogy közben a $p = \frac{S}{N}$ selejtarány állandó. Mivel rögzített n mintanagyság mellett $N \to \infty$ esetén $\left(1-\frac{n-1}{N-1}\right) \to 1$, és az előző feltevés miatt $\sqrt{\frac{p(1-p)}{n}}$ állandó, a visszatevés nélküli mintavételezés estén kapott szórásokra érvényes, hogy

$$\lim_{N\to\infty} D(N_S) = \sqrt{np(1-p)},$$

és

$$\lim_{N\to\infty} D(R_S) = \sqrt{\frac{p(1-p)}{n}}.$$

Tehát elegendően nagy egyedszámú tételek esetén az n elemű mintában adódó selejtszám, illetve selejtarány szórása teintetében sincs különbség a visszatevés nélküli, illetve a visszatevéses mintavételezés között.

A gyakorlatban ez azt jelenti, hogy ha a tétel nagyságához képest elhanyagolható a minta nagysága, akkor a két mintavételezés között a selejtszám, és a selejtarány várhatóértékét és szórását tekintve nincs különbség. Ez már jó közelítéssel teljesül minden olyan esetben, amikor $n \le 0.05N$, vagyis ha a mintába a tételnek legfeljebb 5%-át választjuk ki.

4.3. A POISSON ELOSZLÁS

4.3. Definíció. Az X valószínűségi változót $\lambda > 0$ paraméterű **Poisson eloszlásúnak** nevezzük, ha lehetséges értékei a a nemnegatív egész számok és

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

Bebizonyítható, hogy ha X Poisson eloszlású, akkor

$$E(X) = \lambda,$$

és

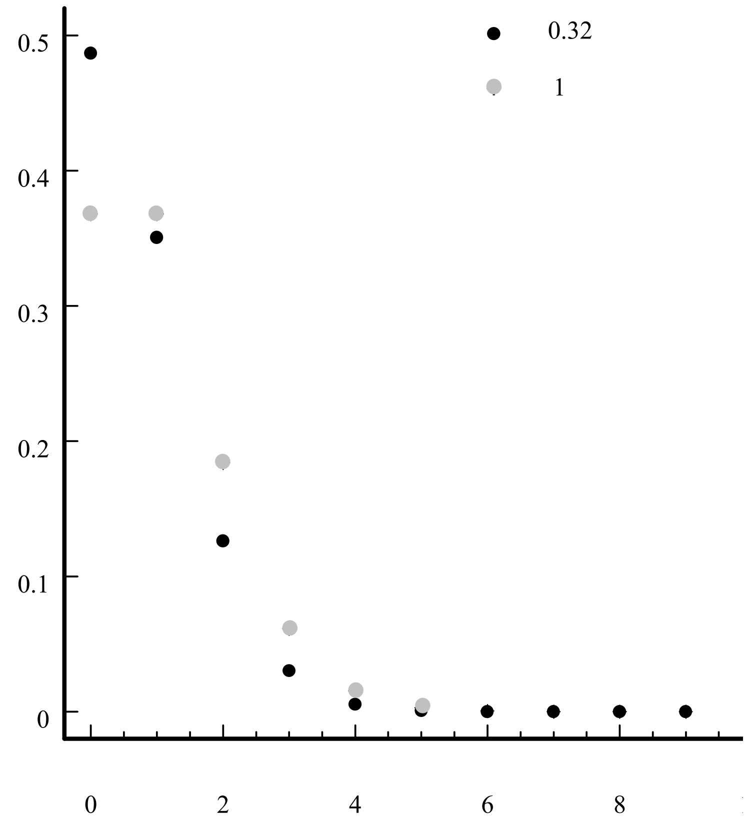
$$D(X) = \sqrt{\lambda}$$

Vagyis a Poisson eloszlás paramétere nem más, mint az eloszlás várhatóértéke, illetve varianciája.

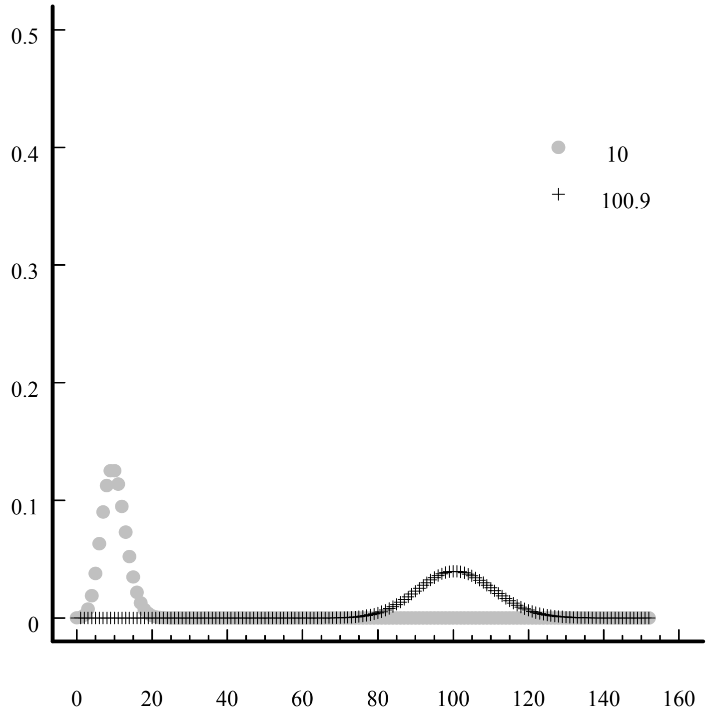
A 4.3 és 4.4 ábrán különböző paraméterű Poisson eloszlásokat láthatunk.

A 4.3. ábrán a $\lambda \le 1$ esetet szemléltetjük. Ha $\lambda < 1$, akkor az X=0 valószínűsége a legnagyobb. Ha $\lambda = 1$, akkor az eloszlás első két tagja egyenlő, és ezek, a 0 és az 1 értékek felvételeinek a valószínűségei a legnagyobbak.

A 4.4. ábrán látható eloszlások esetén $\lambda > 1$. Ekkor a lehetséges értékek közül a $[\lambda]$ a legvalószínűbb, az eloszlást alkotó valószínűségek közül $p_{[\lambda]}$ a maximális. Ha λ egész szám, akkor még az is igaz, hogy $p_\lambda = p_{\lambda-1}$, vagyis két legvalószínűbb érték van.



4.3. ÁBRA



4.4. ÁBRA

4.3.1. A Poisson folyamat

Tekintsünk valamely véletlenszerű időpontokban bekövetkező eseményt. $X(t)$ jelentse a $[0,t)$ időintervallumban az esemény bekövetkezéseinek a számát. (A 0 időpont a megfigyeléseink kezdetét jelenti.) Az esemény bekövetkezései **Poisson folyamatot** alkotnak, ha teljesülnek a következők

1. Annak a valószínűsége, hogy valamely kis $\Delta t > 0$. hosszúságú időintervallumban az esemény bekövetkezik közelítőleg arányos Δt-vel.

$$P(X(t+\Delta t) - X(t) > 0) = \mu \cdot \Delta t, \quad \mu > 0.$$

2. Az $X(t+\Delta t) - X(t)$ valószínűségi változó eloszlása csak Δt-től függ és t-től független. Az $X(t)$ értelmezése szerint ez a különbség nem más, mint a $[t, t+\Delta t)$ időintervallumban az esemény bekövetkezéseinek a száma. Tehát ennek eloszlása csak az intervallum hosszától függ, és független attól, hogy az idő folyamán hol tekintjük ezt az időintervallumot.

3. Annak a valószínűsége, hogy egy adott időpillanatban az esemény egynél többször következzen be nulla..

4. Diszjunkt időintervallumokban az esemény bekövetkezésének számai egymástól függetlenek. Azaz, ha $0 < t_1 < t_2 < \ldots < t_n$ tetszőleges időpontok, akkor az

$$X(t_1), X(t_2)-X(t_1), \ldots, X(t_n)-X(t_{n-1})$$

valószínűségi változók függetlenek.

Ha a fentiek teljesülnek, akkor megmutatható, hogy bármely t>0 eseténaz $X(t)$ valószínűségi változó μt paraméterű Poisson eloszlású. Vagyis a t idő alatt bekövetkező események számának várhatóértéke

$$E(X(t)) = \mu \cdot t\ .$$

Nyilván az egységnyi idő alatti bekövetkezések száma (t=1) μ paraméterű Poisson eloszlású és μ az egységnyi idő alatti bekövetkezések számának várhatóértéke.

4.3. Példa. Egy cég irodahá telefonközpontjába fél óránként átlagosan 4.9 telefonhívás érkezik.

a) Határozzuk meg annak a valószínűségét, hogy 10 és 11 óra között 6-szor fogjál telefonon keresni a z irodaházat.

b) Határozzuk meg annak a valószínűségét, hogy egy adott órányi időtartam altt legalább 12 hívást kap a központ.

A telefonközpontba fél óra alatt beérkező hívások száma 4.9 paraméterű Poisson eloszlású. Legyen X az egy tetszőleges órányi időintervallumban beérkező hívások száma. X 9.8 paraméterű Poisson eloszlású valószínûségi változó.

a)

$$P(X=8) = \frac{9.8^8}{8!} e^{-9.8} = 0.117004.$$

b)

$$P(X \geq 12) = 1 - P(X < 12) = 1 - P(X \leq 11) = 1 - \sum_{k=0}^{11} \frac{9.8^k}{k!} \cdot e^{-9.8}\ .$$

A fenti valószínûséget számítógép használata nélkül igen fáradságos kiszámítani (egy 12 tagú összeg minden elemét ki kell számítani, majd ezeket összegezni!).

5. NEVEZETES FOLYTONOS ELOSZLÁSOK

5.1 A NORMÁLIS ELOSZLÁS

5.1. Definíció. A X folytonos eloszlású valószínűségi változót **normális eloszlásúnak** nevezzük, ha sűrűségfüggvénye

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}},$$

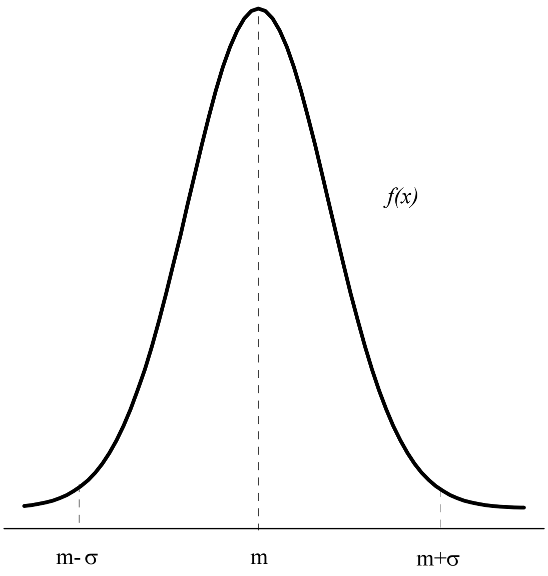
ahol $m \in R$ és $\sigma > 0$ az eloszlás paraméterei.

Mivel az e^{-x^2} függvénynek nincs elemi függvényekkel felírható primitív függvénye, az eloszlás eloszlásfüggvénye zárt formulával nem írható fel, viszont az őt megadó

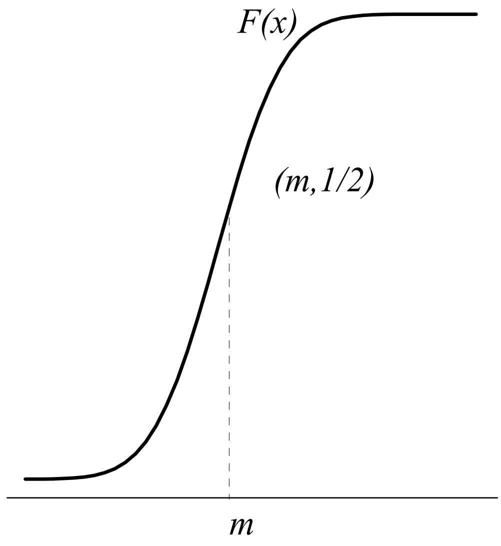
$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{1}{2}\frac{(t-m)^2}{\sigma^2}} dt .$$

improprius integrál értéke bármely x valós szám esetén közelítő numerikus módszerrel tetszőleges pontossággal kiszámítható.

Ezeket a függvényeket láthatjuk a 5.1 a., illetve 5.1. b ábrákon.



5.1. a ÁBRA



5.1. b ÁBRA

A sűrűségfüggvény görbéjét Gauss-görbének, vagy haranggörbének is szokás nevezni. Ez a görbe az $x = m$ egyenesre szimmetrikus, az $x = m$ helyen van az egyetlen lokális maximuma, tehát m az X várhatóértéke (és mediánja):

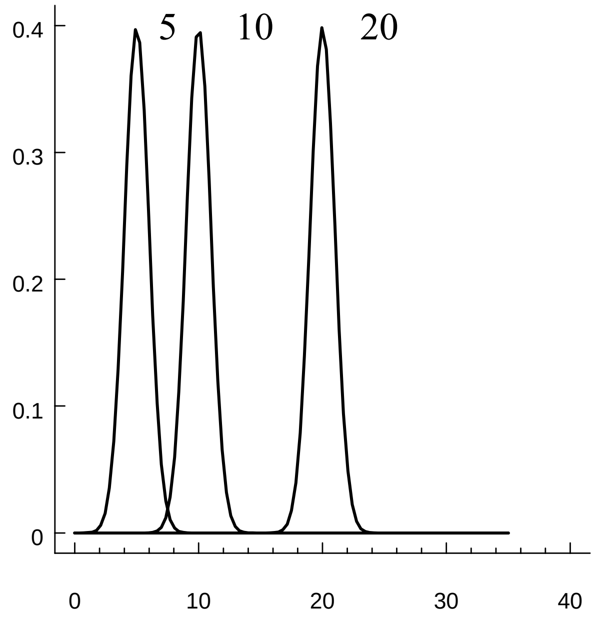
$$m = E(X).$$

Az $f(x)$ függvény második deriváltját vizsgálva meggyőződhetünk róla, hogy az $x = m \pm \sigma$ helyeken pedig inflexiós pontjai vannak.

Megmutatható, hogy a sűrűségfüggvény másik paramétere, σ a valószínűségi változó szórásával egyenlő:

$$\sigma = D(X).$$

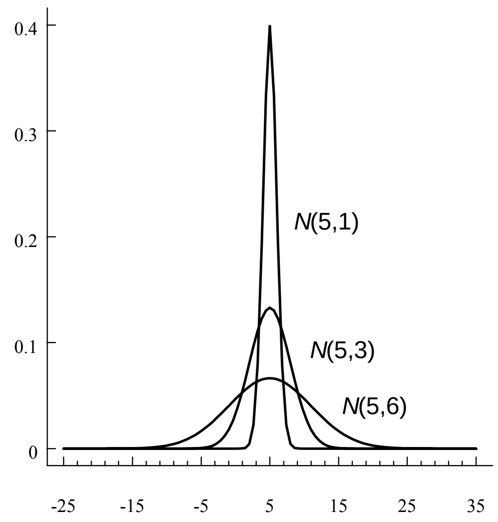
Az 5.2. ábrán az $N(5,1)$, $N(10,1)$, $N(20,1)$ eloszlások sűrűségfüggvényeit ábrázoltuk.



5.2. ÁBRA

Ebből az ábrából jól látható, hogy a várhatóértékkel hogyan változik a valós számoknak az a részhalmaza, amelyikbe nagy valószínűséggel beleesnek a változó értékei.

Az 5.3. ábrán a szórás változásának hatását szemléltetjük. Mindhárom sűrűségfüggvény $m = 5$ várhatóértékű eloszláshoz tartozik, tehát a megfelelő változók értékei nagy valószínűséggel az 5 -ös érték környezetébe fognak esni. A szórás növekedésével



5.3. ÁBRA

azonban az 5-ös értéknek ez a környezete egyre szélesebb.

Bebizonyíthatók a normális eloszlás következő tulajdonságai.

1) Ha $X \ N(m,\sigma)$ eloszlású akkor tetszőleges $Y = aX + b$ lineáris transzformáltja is normális eloszlású $a \cdot m$ várhatóértékkel és $|a| \cdot \sigma$ szórással

2) Ha $X_1 \sim N(m_1, \sigma_1)$, és $X_2 \sim N(m_2, \sigma_2)$, és X_1, X_2 függetlenek, akkor tetszőleges $a, b \in R$ esetén az $aX_1 + bX_2$ valószínűségi változó is normális eloszlású, amelynek várhatóértéke $am_1 + bm_2$, és szórása $\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2}$.

5.1.1. A standard normális eloszlás

Legyen X egy tetszőleges valószínűségi változó, $E(X) = m, D(X) = \sigma$., és tekintsük az $Y = \dfrac{X - m}{\sigma}$ transzformáltját.

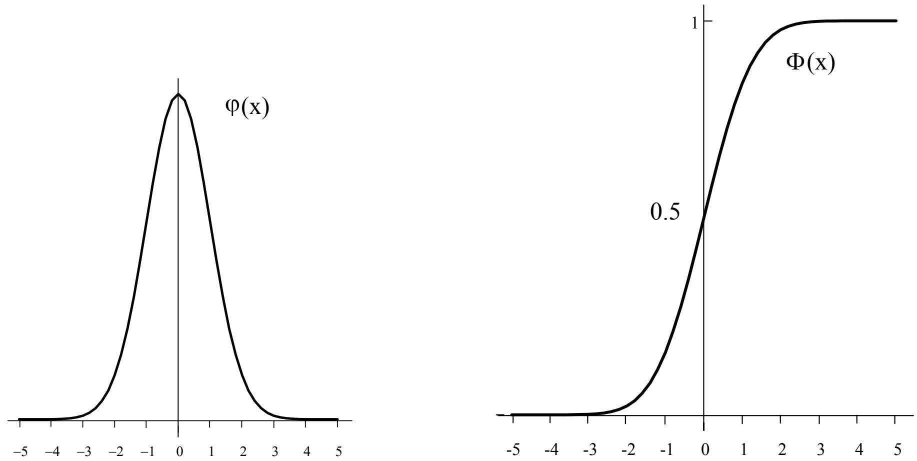
$$E(Y) = E\left(\frac{1}{\sigma}X - \frac{m}{\sigma}\right) = \frac{1}{\sigma}E(X) - \frac{m}{\sigma} = 0,$$

és

$$D^2(Y) = D^2\left(\frac{1}{\sigma}X - \frac{m}{\sigma}\right) = D^2\left(\frac{1}{\sigma}X\right) + D^2\left(\frac{m}{\sigma}\right) = \frac{1}{\sigma^2}D^2(X) + 0 = 1$$

Ezt a transzformációt (a változóból levonjuk a várhatóértékét és elosztjuk a szórásával) **standardizálásnak** nevezzük. Az így kapott Y változót az X **standardizáltjának** nevezzük.

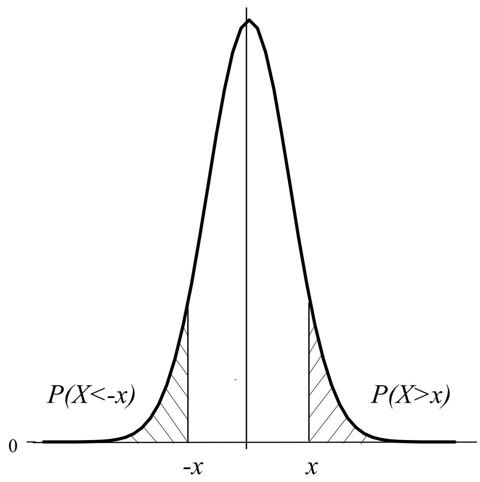
A normális eloszlás fent említett 1) tulajdonsága miatt ha X$\sim N(m,\sigma)$, akkor a standardizáltja $N(0,1)$ eloszlású. Az $N(0,1)$ eloszlást **standard normális eloszlásnak** nevezzük, sűrűségfüggvényét φ-vel, eloszlásfüggvényét Φ-vel jelöljük.(5.4. ábra)



5.4. ÁBRA

A standard normális eloszlás sűrűségfüggvénye páros függvény, ezért geometriai okokból az 5.5. ábra alapján nyilvánvaló, hogy ha X standard normális eloszlású, akkor tetszőleges $x > 0$ esetén

$$P(X < -x) = P(X > x)$$



5.5. ÁBRA

A fenti azonosság a standard normális eloszlás eloszlásfüggvényére vonatkozóan a következőt jelenti:

$$\Phi(-x) = P(X < -x) = P(X > x) = 1 - P(X \le x) = 1 - P(X < x) = 1 - \Phi(x)$$

A következőkben megmutatjuk, hogy milyen kapcsolat van egy tetszőleges normális eloszlás eloszlásfüggvénye és a standard normális eloszlás eloszlásfüggvénye között. Legyen $X \sim N(m,\sigma)$ és Y a standardizáltja. Mivel tetszőleges $x \in R$ esetén

$$X < x = X - m < x - m = \frac{X-m}{\sigma} < \frac{x-m}{\sigma} = Y < \frac{x-m}{\sigma}$$

ezért, ha $F(x)$ az $N(m,\sigma)$ eloszlás eloszlásfüggvénye, akkor

$$F(x) = P(X < x) = P\left(Y < \frac{x-m}{\sigma}\right) = \Phi\left(\frac{x-m}{\sigma}\right).$$

Ennek az összefüggésnek az a gyakorlati jelentősége, hogy tetszőleges normális eloszlás eloszlásfüggvényének az értékeit ki tudjuk számolni, ha a standard normális eloszlás eloszlásfüggvényének az értékeit ismerjük.

Ebből rögtön adódik az is, hogy tetszőleges $[a,b)$ intervallum esetén

$$P(X \in [a,b)) = F(b) - F(a) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right).$$

5.1. Példa. Legyen X $N(100,10)$ eloszlású.
a) Határozzuk meg annak a valószínűségét, hogy az X az átlagától legfeljebb 5%-kal tér el.
b) Melyi az az érték, amelynél nagyobb érték felvételének valószínűsége 0.8?

a) Legyen $F(x)$ az $N(100,10)$ eloszlás eloszlásfüggvénye.

$$P(|X-100|\le 5)=P(-5\le X-100\le 5)=P(95\le X\le 105)=F(105)-F(95).$$

Ha csak a standard normális eloszlásfüggvényére vonatkozó függvénytáblázatunk van, akkor a megoldás:

$$F(105)-F(95)=\Phi\left(\frac{105-100}{10}\right)-\Phi\left(\frac{95-100}{10}\right)=\Phi(0.5)-\Phi(-0.5)=$$

$$=\Phi(0.5)-(1-\Phi(0.5))=2\cdot\Phi(0.5)-1=2\cdot 0.6915-1=$$

$$=1.3830-1=0.3830.$$

b) Jelöljük K-val a kérdéses értéket, amelyre

$$P(X>K)=0.8$$

teljesül.

$$P(X>K)=1-P(X\le K)=1-P(X<\mathrm{K})=0.8\,.$$

Tehát K a

$$P(X<\mathrm{K})=F(K)=0.2$$

összefüggésből határozható meg.

Ha csak a standard normális eloszlásfüggvényére vonatkozó függvénytáblázatunk van, akkor

$$F(K)=\Phi\left(\frac{K-100}{10}\right)=0.2$$

alapján $\frac{K-100}{10}$ nem más, mint a standard normális eloszlás 0.2- hez tartozó kritikus értéke, amit a Φ függvény visszakeresésével határozhatunk meg. Mivel a 0.2 függvényérétk nem szerepel a táblázatban, fel kell használnunk a

$$\Phi\left(\frac{K-100}{10}\right)=1-\Phi\left(-\frac{K-100}{10}\right)=1-\Phi\left(\frac{100-K}{10}\right)$$

összefüggést, tehát a

$$1-\Phi\left(\frac{100-K}{10}\right)=0.2$$

egyenlet alapján

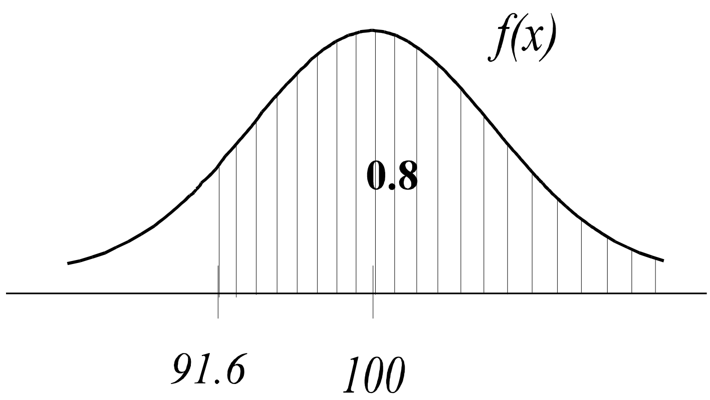
$$\Phi\left(\frac{100-K}{10}\right)=0.8$$

a 0.8 függvényértték visszakeresésével (a 0.8-hez legközelebb eső függvényérték a 0.7995)

$$\frac{100-K}{10}=0.84,$$

ahonnan

$$K=91.6.$$



5.1.2. Az "egy σ" és "két σ " szabályok

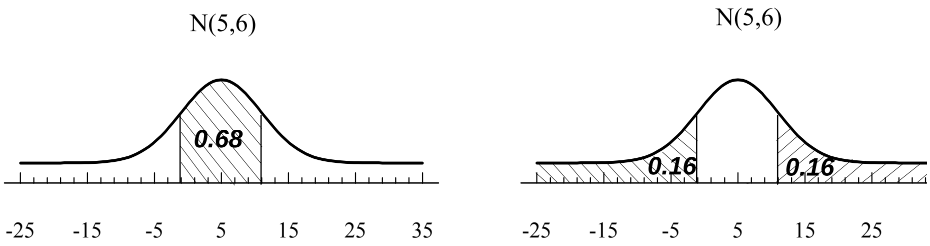
Legyen $X\sim N(m,\sigma)$. Határozzuk meg a $X\in[m-\sigma,m+\sigma]$ eseménynek a valószínűségét. Ez az esemény azt jelenti, hogy a változó értéke pozitív, vagy negatív irányban legfeljebb egy szórásnyival tér el a várhatóértéktől. Az előzőek szerint

$$P(X\in[m-\sigma,m+\sigma])=F(m+\sigma)-F(m-\sigma)=$$

$$=\Phi\left(\frac{m+\sigma-m}{\sigma}\right)-\Phi\left(\frac{m-\sigma-m}{\sigma}\right)=\Phi(1)-\Phi(-1)=$$

$$=\Phi(1)-(1-\Phi(1))=2\Phi(1)-1=2\cdot 0.8413-1=0.68$$

Ennek megfelelően az u.n. "egy σ szabály": 0.68 annak a valószínűsége, hogy egy normális eloszlású valószínűségi változó értéke a várhatóértékétől legfeljebb egy szórásnyival térjen el, azaz 0.32 annak a valószínűsége, hogy (vagy pozitív, vagy



5.5. ÁBRA

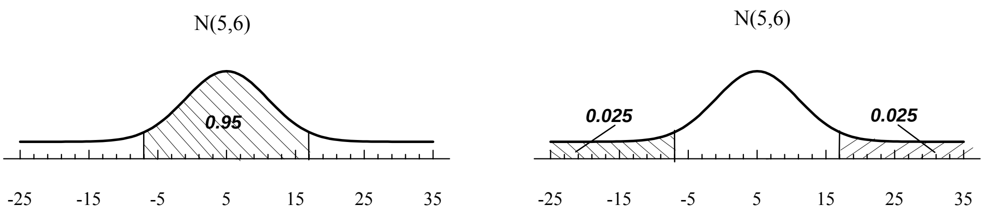
negatív irányban) a szórás egyszeresénél többel térjen el a várhatóértékétől. (5.5. ábra)

A "két σ szabály "pedig az $X \in [m-2\sigma, m+2\sigma]$ esemény valószínűségére vonatkozik.(5.6. ábra)

$$P(X \in [m-2\sigma, m+2\sigma]) = F(m+2\sigma) - F(m-2\sigma) =$$

$$= \Phi\left(\frac{m+2\sigma-m}{\sigma}\right) - \Phi\left(\frac{m-2\sigma-m}{\sigma}\right) =$$

$$= \Phi(2) - (1-\Phi(2)) = 2\Phi(2) - 1 = 2 \cdot 0.9772 - 1 = 0.95$$



5.6. ÁBRA

5.2. Példa. Egy gép furatokat készít. A furatok átmérője normális eloszlású, melynek várhatóértéke 60 mm, szórása pedig 1.2 mm. Az egyes furatok átmérői egymástól függetlenek.

a)Mi annak a valószínűsége, hogy 10 egymás után készült furat között lesz olyan, amelyiknek az átmérője a 60 mm-től egy szórásnyinál jobban fog különbözni?
b) Melyik az a *K*, amely esetén 0.05 annak a valószínûsége, hogy 10 egymás után készült furat között lesz olyan, amelyiknek az átmérője a 60 mm-től *K*-nál többel fog különbözni?
c) Melyik az a *K*, amely esetén 0.95 annak a valószínûsége, hogy 10 egymás után készült furat mindegyikének az átmérője a 60 mm-től legfeljebb *K*-val t fog különbözni?

Jelentse X a furat átmérőjét. Esetünkben X~N(60,1.2), és $F(x)$ legyen ennek a normális eloszlásnak az eloszlásfüggvénye.

a) Y jelentse a 10 egymás után készült furat között azoknak a számát, amelyek átmérője a 60 mm-től egy szórásnyinál jobban különbözik. Ha $p = P(|X-60| > 1.2)$, akkor Y azt jelenti, hogy a 10 eset közül hányban következett be a *p* valószínûségű esemény, vagyis Y~B(10, p).

Az $Y > 0$ esemény valószínűségére vagyunk kíváncsiak.

Először is meg kell határoznunk a *p* valószínûséget. Az "egy σ szabály" serint

$$p = P(|X-60| > 1.2) = 1 - 0.68 = 0.32,$$

tehát Y~B(10,0.32).

$$P(Y > 0) = 1 - P(Y \le 0) = 1 - P(Y = 0),$$

mivel binomiális valószínûségi változó esetén az $Y \le 0$ esemény akkor és csakis akkor következik be, ha *Y*=0. Tehát

$$P(Y > 0) = 1 - \binom{10}{0} \cdot 0.32^0 \cdot (1-0.32)^{10-0} = (1-0.32)^{10-0} = 0.9789.$$

b) Legyen Y a 10 egymás után készült furat között azoknak a száma, amelyek átmérője a 60 mm-től K-nál többel fog különbözni. Ha $p = P(|X-60| > K)$, akkor Y~B(10, p).

Most annak a valószínűsége adott, hogy a $p = P(|X-60| > K)$ valószínûségű esemény a 10 eset között bekövetkezik. A kérdés tehát az, hogy milyen *K* esetén lesz a *p* olyan, hogy a B(10, *p*) eloszlású valószínûségi változóra teljesül, hogy

$$P(Y > 0) = 0.05.$$

Mivel *Y* binomiális eloszlású

$$P(Y > 0) = 1 - P(Y \le 0) = 1 - P(Y = 0) = 1 - \binom{10}{0} \cdot p^0 \cdot (1-p)^{10-0} = 1 - (1-p)^{10} = 0.05,$$

tehát *p* az

$$1 - (1-p)^{10} = 0.05$$

egenletből határozható meg.

$$1 - p = \sqrt[10]{0.95},$$

$$p = 0.0051.$$

Tehát az $|X-60| > K$ esemény valószínûségének 0.0051-nek kell lennie.

$$P(|X-60| > K) = P(X-60 > K) + P(X-60 < -K) =$$

$$= 1 - P(X-60 \le K) + P(X-60 < \text{-}K) =$$

$$= 1 - P(X \le K + 60) + P(X < -K + 60) = 1 - F(K + 60) - F(-K + 60) =$$

$$1 - \Phi\left(\frac{K}{1.2}\right) + \Phi\left(\frac{-K}{1.2}\right) =$$

$$= 1 - \Phi\left(\frac{K}{1.2}\right) + \left(1 - \Phi\left(\frac{K}{1.2}\right)\right) = 2 \cdot \left(1 - \Phi\left(\frac{K}{1.2}\right)\right),$$

ahol Φ a standard normális eloszlás eloszlásfüggvénye.

Vagyis K-t a

$$2 \cdot \left(1 - \Phi\left(\frac{K}{1.2}\right)\right) = 0.0051$$

összefüggésből határozhatjuk meg.

$$\Phi\left(\frac{K}{1.2}\right) = 0.9975.$$

Meghatározzuk azt az értéket, ahol a Φ függvény a 0.9975 értéket veszi fe1.

$$\Phi(2.8070) = 0.9975,$$

tehát

$$\frac{K}{1.2} = 2.8070.$$

amiből

$$K = 1.2 \cdot 2.8070 = 3.3684.$$

Tehát annak a valószínűsége 0.05, hogy 10 egymás után készült furat között lesz olyan, amelyiknek az átmérője a 60 mm-től 3.3684 mm-nél többel fog különbözni

c) Legyen Y a 10 egymás után készült furat között azoknak a száma, amelyek átmérője a 60 mm-től legfeljebb K-val fog különbözni. Ha $p = P(|X - 60| \le K)$, akkor $Y \sim B(10, p)$.

Most az $Y = 10$ esemény valószínűsége adott. A kérdés az, hogy milyen K esetén lesz a p olyan, hogy a B(10, p) eloszlású valószínûségi változóra teljesül, hogy

$$P(Y = 10) = \binom{10}{10} \cdot p^{10} \cdot (1-p)^0 = p^{10} \cdot = 0.95,$$

tehát p az

$$p^{10} \cdot = 0.95$$

egyenletből határozható meg.

$$p = \sqrt[10]{0.95},$$

$$p = 0.9949.$$

Tehát az $|X - 60| \le K$ esemény valószínûségének 0.9949-nak kell lennie.

$$P(|X - 60| \le K) = P(-K < X - 60 < \mathrm{K}) = P(-K + 60 < X < \mathrm{K} + 60) =$$

$$= F(K + 60) - F(-K + 60) = \Phi\left(\frac{K}{1.2}\right) - \Phi\left(\frac{-K}{1.2}\right) =$$

$$\Phi\left(\frac{K}{1.2}\right) - \left(1 - \Phi\left(\frac{K}{1.2}\right)\right) = 2 \cdot \Phi\left(\frac{K}{1.2}\right) - 1,$$

ahol Φ a standard normális eloszlás eloszlásfüggvénye.

Vagyis K-t a

$$2 \cdot \Phi\left(\frac{K}{1.2}\right) - 1 = 0.9949$$

egyenletből határozhatjuk meg.

$$\Phi\left(\frac{K}{1.2}\right) = \frac{1.9949}{2} = 0.9975.$$

Meghatározzuk azt az értéket, ahol a Φ függvény a 0.9975 értéket veszi fe1.

$$\Phi(2.8070) = 0.9975,$$

tehát

$$\frac{K}{1.2} = 2.8070,$$

amiből

$$K = 3.3684.$$

vagyis annak a valószínűsége 0.95, hogy 10 egymás után készült furat midegyikének az átmérője a 60 mm-től legfeljebb 3.3684 mm-rel fog különbözni.

5.1.3. Az exponenciális eloszlás

5.2. Definíció A X valószínűségi változót $\lambda\,(>0)$ paraméterű exponenciális eloszlásúnak nevezzük, ha eloszlásfüggvénye

$$F(x)=\begin{cases}0, & \text{ha } x\le 0\\ 1-e^{-\lambda x}, & \text{ha } x>0\end{cases}.$$

Ennek megfelelően a λ paraméterű exponenciális eloszlás folytonos eloszlás, melynek sűrűségfüggvénye

$$f(x)=\begin{cases}0, & \text{ha } x<0\\ \lambda e^{-\lambda x}, & \text{ha } x>0\end{cases}.$$

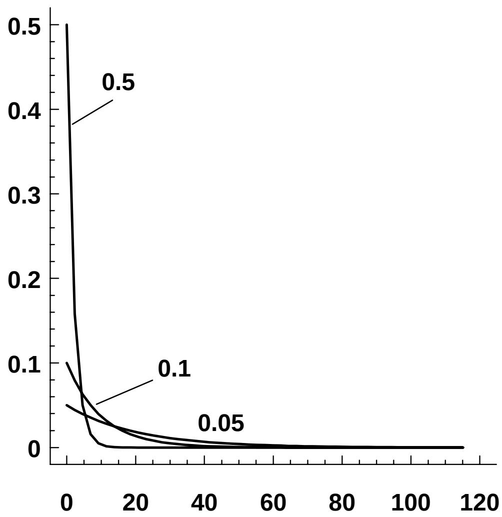
Megmutatható, hogy

$$E(X)=\frac{1}{\lambda},$$

és

$$D(X)=\frac{1}{\lambda}.$$

Néhány exponenciális eloszlás sűrűségfüggvényét az 5.7. ábrán láthatjuk.



5.7. ÁBRA

Bebizonyítható, hogy valamely T valószínűségi változó akkor és csakis akkor exponenciális eloszlású, tetszőleges $t,\Delta t>0$ esetén teljesül, hogy

$$P(T\ge t+\Delta t \mid T\ge t)=P(T\ge \Delta t)$$

Az exponenciális eloszlásnak ezt a tulajdonságát "örökifjúságnak" szoktuk nevezni. Ha ugyanis pl. T valamelyin típusú berendezés, alkatrész stb. élettartamát jelöli, akkaor a fenti tulajdonság azt jelenti, hogy ha az illető már legalább t ideje létezik(működik), akkor annak a valószínűsége, hogy még legalább további Δt idejig fog létezni(működni) egyenlő azzal, hogy az illető egyáltalán legalább Δt ideig létezik(működik).

Az exponenciális eloszlás gyakorlati alkalmazását illetően két fontos területet említünk meg.

Élettartamok

Az olyan berendezések, alkatrészek élettartama, amelyek tönkremenetelében az öregedés nem játszik szerepet (tönkremenésüket egyéb véletlen körülmények okozzák) exponenciális eloszlású. Hogy az öregedésnek nincs szerepe az azt jelenti, hogy ha a berendezésünk már legalább t ideje működik, akkor annak a valószínűsége, hogy még legalább további Δt ideig fog működni az ugyanannyi, minthogy egyáltalán az élettartama legalább Δt lesz. (Az emberi élettartam nyilván nem ilyen. Ugyanis annak a valószínűsége, hogy egy 90 éves ember még további 10 évet fog élni az nem ugyanakkora, mint hogy egy ember megéri a 10. születésnapját.)

Várakozási idők

Tekintsünk véletlenszerű időpontokban bekövetkező eseményeket, ahol a bekövetkezések Poisson folyamatot alkotnak. Ekkor tetszőleges t idő alatti bekövetkezések száma $\mu \cdot t$ paraméterű Poisson eloszlású valószínűségi változó, ahol μ az *egységnyi idő* alatti bekövetkezések számának a várhatóértéke. Legyen T *ugyanilyen egységekeben* mérve az esemény két egymás utáni bekövetkezése között eltelt időt. Ez az időtartam nyilván véletlenszerű. T-t várakozási időnek szokás nevezni, mert ennyi időt kell várnunk az esemény legközelebbi bekövetkezéséig. Bebizonyítható, hogy a T valószínűségi változó μ paraméterű exponenciális eloszlású. Az esemény két egymás utáni bekövetkezése között eltelt idő várhatóértéke tehát $1/\mu$.

5.3. Példa. Egy áruház pénztárához 9 és 12 óra közötti időszakban óránként átlagosan 54 vásárló érkezik. Mi a valószínűsége annak, hogy az adott időszakban két vásárló beérkezése között 1.5 percnél rövidebb idő telik el?

Jelölje T a két vásárló beérkezése között eltelt időt.

a) Azt tudjuk, hogy az 1 óra alatt beérkező vásárlók száma 54 paraméterű Poisson eloszlású, tehát az egy perc alatt beérkező vásárlók száma $\frac{54}{60} = 0.9$ paraméterű Poisson eloszlású. (egységnyi idő =1 perc!) Ha T jelenti a két vásárló beérkezése között eltelt idő percekben mérve, akkor T 0.9 paraméterű exponenciális eloszlású és a T<1.5 esemény valószínûsége a kérdés. Legyen $F(x)$ a 0.9 paraméterű exponenciális eloszlás eloszlásfüggvénye. Ekkor

$$P(T < 1.5) = F(1.5) = 1 - e^{-0.9 \cdot 1.5} = 1 - 0.2592 = 0.7408.$$

b) Ha T jelenti a két vásárló beérkezése között eltelt időt órákban mérve, akkor mivel az egy óra alatt beérkező vásárlók száma 54 paraméterű Poisson eloszlású (egységnyi idõ =1 óra!), a T 54 paraméterű exponenciális eloszlású valószínûségi változó , és 1.5 perc=1.5/60 óra miatt most a $T < \frac{1.5}{60}$ esemény valószínûségére vagyunk kíváncsiak. Ha most $F(x)$ az 54 paraméterű exponenciális eloszlás eloszlásfüggvénye, akkor

$$P\left(T < \frac{1.5}{60}\right) = F\left(\frac{1.5}{60}\right) = F\left(\frac{1}{40}\right) = 1 - e^{-54\cdot\frac{1}{40}} = 0.7408\,.$$

A NORMÁLIS ELOSZLÁS FÜGGVÉNYTÁBLÁZATA

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

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.0000	0.5000	0.3600	0.6406	0.7200	0.7642
0.0100	0.5040	0.3700	0.6443	0.7300	0.7673
0.0200	0.5080	0.3800	0.6480	0.7400	0.7704
0.0300	0.5120	0.3900	0.6517	0.7500	0.7734
0.0400	0.5160	0.4000	0.6554	0.7600	0.7764
0.0500	0.5199	0.4100	0.6591	0.7700	0.7794
0.0600	0.5239	0.4200	0.6628	0.7800	0.7823
0.0700	0.5279	0.4300	0.6664	0.7900	0.7852
0.0800	0.5319	0.4400	0.6700	0.8000	0.7881
0.0900	0.5359	0.4500	0.6736	0.8100	0.7910
0.1000	0.5398	0.4600	0.6772	0.8200	0.7939
0.1100	0.5438	0.4700	0.6808	0.8300	0.7967
0.1200	0.5478	0.4800	0.6844	0.8400	0.7995
0.1300	0.5517	0.4900	0.6879	0.8500	0.8023
0.1400	0.5557	0.5000	0.6915	0.8600	0.8051
0.1500	0.5596	0.5100	0.6950	0.8700	0.8079
0.1600	0.5636	0.5200	0.6985	0.8800	0.8106
0.1700	0.5675	0.5300	0.7019	0.8900	0.8133
0.1800	0.5714	0.5400	0.7054	0.9000	0.8159
0.1900	0.5753	0.5500	0.7088	0.9100	0.8186
0.2000	0.5793	0.5600	0.7123	0.9200	0.8212
0.2100	0.5832	0.5700	0.7157	0.9300	0.8238
0.2200	0.5871	0.5800	0.7190	0.9400	0.8264
0.2300	0.5910	0.5900	0.7224	0.9500	0.8289
0.2400	0.5948	0.6000	0.7257	0.9600	0.8315
0.2500	0.5987	0.6100	0.7291	0.9700	0.8340
0.2600	0.6026	0.6200	0.7324	0.9800	0.8365
0.2700	0.6064	0.6300	0.7357	0.9900	0.8389
0.2800	0.6103	0.6400	0.7389	1.0000	0.8413
0.2900	0.6141	0.6500	0.7422	1.0100	0.8438
0.3000	0.6179	0.6600	0.7454	1.0200	0.8461
0.3100	0.6217	0.6700	0.7486	1.0300	0.8485
0.3200	0.6255	0.6800	0.7517	1.0400	0.8508
0.3300	0.6293	0.6900	0.7549	1.0500	0.8531
0.3400	0.6331	0.7000	0.7580	1.0600	0.8554
0.3500	0.6368	0.7100	0.7611	1.0700	0.8577

A NORMÁLIS ELOSZLÁS FÜGGVÉNYTÁBLÁZATA

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

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1.0800	0.8599	1.4400	0.9251	1.8000	0.9641
1.0900	0.8621	1.4500	0.9265	1.8100	0.9649
1.1000	0.8643	1.4600	0.9279	1.8200	0.9656
1.1100	0.8665	1.4700	0.9292	1.8300	0.9664
1.1200	0.8686	1.4800	0.9306	1.8400	0.9671
1.1300	0.8708	1.4900	0.9319	1.8500	0.9678
1.1400	0.8729	1.5000	0.9332	1.8600	0.9686
1.1500	0.8749	1.5100	0.9345	1.8700	0.9693
1.1600	0.8770	1.5200	0.9357	1.8800	0.9699
1.1700	0.8790	1.5300	0.9370	1.8900	0.9706
1.1800	0.8810	1.5400	0.9382	1.9000	0.9713
1.1900	0.8830	1.5500	0.9394	1.9100	0.9719
1.2000	0.8849	1.5600	0.9406	1.9200	0.9726
1.2100	0.8869	1.5700	0.9418	1.9300	0.9732
1.2200	0.8888	1.5800	0.9429	1.9400	0.9738
1.2300	0.8907	1.5900	0.9441	1.9500	0.9744
1.2400	0.8925	1.6000	0.9452	1.9600	0.9750
1.2500	0.8944	1.6100	0.9463	1.9700	0.9756
1.2600	0.8962	1.6200	0.9474	1.9800	0.9761
1.2700	0.8980	1.6300	0.9484	1.9900	0.9767
1.2800	0.8997	1.6400	0.9495	2.0000	0.9773
1.2900	0.9015	1.6500	0.9505	2.0100	0.9778
1.3000	0.9032	1.6600	0.9515	2.0200	0.9783
1.3100	0.9049	1.6700	0.9525	2.0300	0.9788
1.3200	0.9066	1.6800	0.9535	2.0400	0.9793
1.3300	0.9082	1.6900	0.9545	2.0500	0.9798
1.3400	0.9099	1.7000	0.9554	2.0600	0.9803
1.3500	0.9115	1.7100	0.9564	2.0700	0.9808
1.3600	0.9131	1.7200	0.9573	2.0800	0.9812
1.3700	0.9147	1.7300	0.9582	2.0900	0.9817
1.3800	0.9162	1.7400	0.9591	2.1000	0.9821
1.3900	0.9177	1.7500	0.9599	2.1100	0.9826
1.4000	0.9192	1.7600	0.9608	2.1200	0.9830
1.4100	0.9207	1.7700	0.9616	2.1300	0.9834
1.4200	0.9222	1.7800	0.9625	2.1400	0.9838
1.4300	0.9236	1.7900	0.9633	2.1500	0.9842

A NORMÁLIS ELOSZLÁS FÜGGVÉNYTÁBLÁZATA

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

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2.1600	0.9846	2.5200	0.9941	2.8800	0.9980
2.1700	0.9850	2.5300	0.9943	2.8900	0.9981
2.1800	0.9854	2.5400	0.9945	2.9000	0.9981
2.1900	0.9857	2.5500	0.9946	2.9100	0.9982
2.2000	0.9861	2.5600	0.9948	2.9200	0.9982
2.2100	0.9864	2.5700	0.9949	2.9300	0.9983
2.2200	0.9868	2.5800	0.9951	2.9400	0.9984
2.2300	0.9871	2.5900	0.9952	2.9500	0.9984
2.2400	0.9875	2.6000	0.9953	2.9600	0.9985
2.2500	0.9878	2.6100	0.9955	2.9700	0.9985
2.2600	0.9881	2.6200	0.9956	2.9800	0.9986
2.2700	0.9884	2.6300	0.9957	2.9900	0.9986
2.2800	0.9887	2.6400	0.9959	3.0000	0.9987
2.2900	0.9890	2.6500	0.9960	3.0100	0.9987
2.3000	0.9893	2.6600	0.9961	3.0200	0.9987
2.3100	0.9896	2.6700	0.9962	3.0300	0.9988
2.3200	0.9898	2.6800	0.9963	3.0400	0.9988
2.3300	0.9901	2.6900	0.9964	3.0500	0.9989
2.3400	0.9904	2.7000	0.9965	3.0600	0.9989
2.3500	0.9906	2.7100	0.9966	3.0700	0.9989
2.3600	0.9909	2.7200	0.9967	3.0800	0.9990
2.3700	0.9911	2.7300	0.9968	3.0900	0.9990
2.3800	0.9913	2.7400	0.9969	3.1000	0.9990
2.3900	0.9916	2.7500	0.9970	3.1100	0.9991
2.4000	0.9918	2.7600	0.9971	3.1200	0.9991
2.4100	0.9920	2.7700	0.9972	3.1300	0.9991
2.4200	0.9922	2.7800	0.9973	3.1400	0.9992
2.4300	0.9925	2.7900	0.9974	3.1500	0.9992
2.4400	0.9927	2.8000	0.9974	3.1600	0.9992
2.4500	0.9929	2.8100	0.9975	3.1700	0.9992
2.4600	0.9931	2.8200	0.9976	3.1800	0.9993
2.4700	0.9932	2.8300	0.9977	3.1900	0.9993
2.4800	0.9934	2.8400	0.9977	3.2000	0.9993
2.4900	0.9936	2.8500	0.9978	3.2100	0.9993
2.5000	0.9938	2.8600	0.9979	3.2200	0.9994
2.5100	0.9940	2.8700	0.9979	3.2300	0.9994

A NORMÁLIS ELOSZLÁS FÜGGVÉNYTÁBLÁZATA

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

x	$\Phi(x)$	x	$\Phi(x)$
3.2400	0.9994	3.6400	0.9999
3.2500	0.9994	3.6500	0.9999
3.2600	0.9994	3.6600	0.9999
3.2700	0.9995	3.6700	0.9999
3.2800	0.9995	3.6800	0.9999
3.2900	0.9995	3.6900	0.9999
3.3000	0.9995	3.7000	0.9999
3.3100	0.9995	3.7100	0.9999
3.3200	0.9995	3.7200	0.9999
3.3300	0.9996	3.7300	0.9999
3.3400	0.9996	3.7400	0.9999
3.3500	0.9996	3.7500	0.9999
3.3600	0.9996	3.7600	0.9999
3.3700	0.9996	3.7700	0.9999
3.3800	0.9996	3.7800	0.9999
3.3900	0.9997	3.7900	0.9999
3.4000	0.9997	3.8000	0.9999
3.4100	0.9997	3.9100	0.9999
3.4200	0.9997	3.9200	0.9999
3.4300	0.9997	3.9300	0.9999
3.4400	0.9997	3.9400	0.9999
3.4500	0.9997	3.9500	0.9999
3.4600	0.9997	3.9600	0.9999
3.4700	0.9997	3.9700	0.9999
3.4800	0.9997	3.9800	0.9999
3.4900	0.9998	3.9900	0.9999
3.5000	0.9998	4.0000	0.9999
3.5100	0.9998		
3.5200	0.9998		
3.5300	0.9998		
3.5400	0.9998		
3.5900	0.9998		
3.6000	0.9998		
3.6100	0.9998		
3.6200	0.9999		
3.6300	0.9999		

TARTALOMJEGYZÉK

1. **VÉLETLEN ESEMÉNYEK** 1
1.1 Elemi esemény, összetett esemény 1
1.2. Műveletek eseményekkel 3
1.3 Az eseményműveletek tulajdonságai 6
2. **VALÓSZÍNŰSÉGSZÁMÍTÁSI ALAPOK** 11
2.1 Valószínűségek 11
2.2 A valószínűség axiómái 12
2.3A véletlen kísérlet matematikai modellje 13
2.4 Klasszikus valószínűségi kísérlet 14
2.5 A valószínűség tulajdonságai 16
2.6 Feltételes valószínűség, események függetlensége 17
2.6.1.A feltételes valószínűség 17
2.6.2.A feltételes valószínűség tulajdonságai 19
2.6.3.Események függetlensége 20
2.6.4 Független kísérletek 21
2.7 A valószínűségek szorzási szabálya 22
2.8 Teljes valószínűség tétele, Bayes -étel 23
3. **VALÓSZÍNŰSÉGI VÁLTOZÓK** 26
3.1 A valószínűségi változó fogalma 26
3.2. Az eloszlásfüggvény 28
3.2.1.Az eloszlásfüggvény tulajdonságai 28
3.3. Eloszlások osztályozása 29
3.3.1.A sűrűségfüggvény tulajdonságai 32
3.3.2.Folytonos eloszlású változók tulajdonságai 32
3.3.3.Diszkrét eloszlás közelítése folytonossal 33

3.4 Eloszlások numerikus jellemzői 33

3.4.1.A várhatóérték 33

3.4.1.1.A várhatóérték tulajdonságai 35

3.4.2.A szórás 36

3.4.2.1.A szórás tulajdonságai 36

3.4.2.2.Szórás számítása az eloszlás ismeretében 37

3.4.3.A medián 37

3.4.4.Kvartilisek, módusz 39

3.4.5.Szimmetrikusság, ferdeség 41

3.5 Kockázati döntések 44

4. **NEVEZETES DISZKRÉT ELOSZLÁSOK** 49

4.1 Binomiális eloszlás 49

4.1.1.A binomiális eloszlás és a visszatevéses mintavételezés 52

4.2 A hipergeometrikus eloszlás 54

4.2.1.A hipergeometrikus eloszlás és a visszatevés nélküli mintavételezés 54

4.3. A Poisson eloszlás 56

4.3.1.A Poisson folyamat 58

5. **NEVEZETES FOLYTONOS ELOSZLÁSOK** 60

5.1 A normális eloszlás 60

5.1.1.A standard normális eloszlás 63

5.1.2.Az "egy σ" és "két σ " szabályok 66

5.1.3.Az exponenciális eloszlás 70