

### 3. A VALÓSZÍNŰSÉG-ELMÉLET ALAPJAI

Ebben a függelékben azokat a valószínűség-elméleti alapfogalmakat foglalkoztatjuk össze, amelyekre a mérések kiértékeléséhez szükség van. A 3.1. alfejezet a területen teljesen kezdők számára készült. A későbbiek a területet ismerők számára is hasznos ismétlés lehet. Figyelmeztetjük azonban az Olvasót, hogy a rövidség kedvéért itt számos egyszerűsítésre kényszerülünk, tehát ennek a függeléknek a tanulmányozása nem helyettesíti a valószínűség-elmélet alapos megtanulását. Az irodalom rendkívül gazdag magyar nyelven is [1]. Az irodalomjegyzék csak példaképpen ajánl néhány könyvet.

#### 3.1. Alapfogalmak

##### Esemény és valószínűség

##### A valószínűség definíciója

Amikor *kísérletet* végzünk, annak *kimenetele* legtöbbször nem jósolható meg biztonsággal, mert a *véletlentől* függ. A valószínűség-elmélet tárgya az ilyen kísérletek elemzése. Ha csak egyetlen kísérletet végzünk, annak kimeneteléről alig lehet valamit mondani, viszont az elmélet kijelentései egyre megbízhatóbbá válnak, ahogy egyre többször ismétljük meg a kísérletet. Úgy is fogalmazhatunk, hogy *a valószínűség-elmélet a véletlen tömegjelenségekkel foglalkozik*.

A kísérlet minden lehetséges kimenetelét *elemi eseménynek* nevezzük. Példák elemi eseményekre:

- két kockával való dobáskor a két kockán kapott számokból alkotott számpár: (2, 3), (5, 1) stb.;
- bridge-osztáskor a négy kézben levő 13–13 lap együttese;
- céllövéskor a golyó becsapódási helyének a céltábla középpontjától való távolsága;
- lottóhúzáskor a kijövő számötös;
- radioaktív bomláskor az 1 s alatt elbomlott atomok száma.

Az összes lehetséges elemi események együttesét tekintjük az  $\Omega$  halmaz elemeinek.  $\Omega$  részhalmazait *eseményeknek* nevezzük. Természetesen minden elemi esemény egyben esemény is. Együttesüket *eseménytérnek* nevezzük. Példák eseményekre:

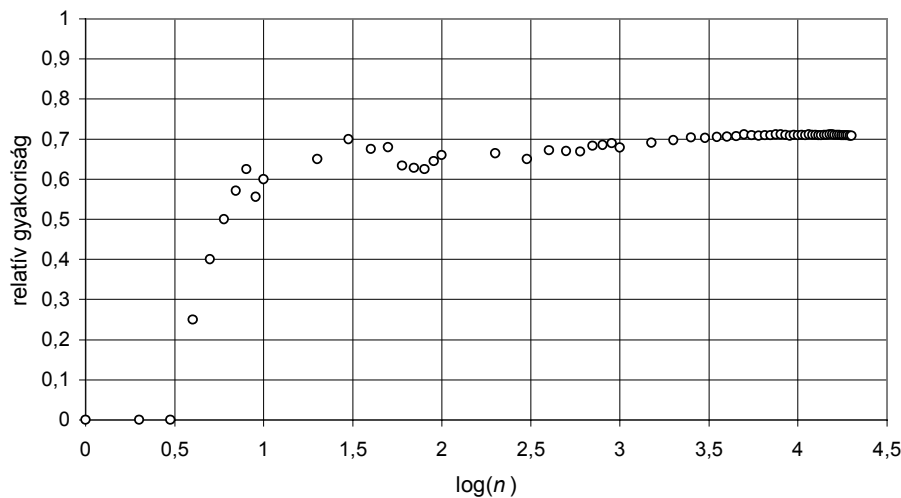
- két kockával való dobáskor a két kockán kapott szám egymással egyenlő:  $\{(1, 1); (2, 2); (3, 3); (4, 4); (5, 5); (6, 6)\}$ ;

- bridge-osztáskor egy kézben van mind a négy ász;
- céllövéskor a 10-es kör, vagyis a golyó becsapódási helyének a céltábla középpontjától való távolsága kisebb, mint a 10-es kör sugara;
- lottóhúzáskor 5 darab négytalalatos szelvény van;
- radioaktív bomláskor az 1 s alatt elbomlott atomok száma kisebb, mint 1000.

A definícióból nyilvánvaló, hogy az  $\Omega$  esemény biztosan bekövetkezik. Ezért ezt *biztos eseménynek* nevezzük.  $\Omega$  valódi részhalmazai a kísérleteknek csak egy részében következnek be. Tegyük fel, hogy  $n$  kísérletet végeztünk, és az  $A$  esemény  $k$ -szor következett be. A  $k/n$  hányadost *relatív gyakoriságnak* nevezzük. Azt tapasztaljuk, hogy erősen ingadozik, amíg  $n$  kicsi, de  $n$  növekedésével stabilizálódik, és egy határértékhez tart. Ezt illusztráljuk a 3.1. ábrán:  $n \geq 5000$ -re a relatív gyakoriság gyakorlatilag stabilizálódik 0,7 közelében. Ezt a határértéket nevezzük *valószínűségnek*:

$$p(A) = \lim_{n \rightarrow \infty} \frac{k}{n}. \quad (3.1)$$

Egy külön fejezetet igényelne a konvergencia természetének elemzése, így ettől el kell tekintenünk. Mindössze annyit jegyzünk meg, hogy ezt az összefüggést a *nagy számok törvényének* nevezzük, amelynek több változata létezik.<sup>1</sup>



3.1. ábra. A relatív gyakoriság konvergenciája a valószínűséghez ( $p = 0,7$ )

Azt az eseményt, amelyre a valószínűség vonatkozik, (3.1) mintájára argumentumként jelöljük, ha szükséges. Bonyolultan definiálható események valószínűségét a következőképpen is jelölhetjük:

$$p = P\{r \leq R_{10}\},$$

<sup>1</sup> A fent említett egyszerűsítések egyike, hogy így definiáljuk a valószínűséget. A modern valószínűség-elmélet egészen másképp közelíti meg a dolgot. Az alábbiak megértéséhez azonban elégséges lesz a (3.1) képlet szerinti definíció. A konvergencia jellegét a későbbiekben még megvilágítjuk.

ahol  $R_{10}$  a 10-es kör és  $r$  a céltáblába fúródó golyó helyének a sugara. Ezzel annak az eseménynek a valószínűségét írtuk fel, hogy a céllövőnek 10-es kört sikerült lőnie.

### Független és egymást kizáró események

Az eseményekkel kapcsolatban meg kell ismerkednünk néhány fogalommal. Kimondunk továbbá néhány alapvető tételt – de bizonyítás nélkül.  $A$  és  $B$  egymást kizáró események, ha közös részük az üres halmaz. Más szóval egyszerre nem következhetnek be. Ha a kockadobásban az egyes kockákon kijött számokat  $i$ -vel és  $j$ -vel jelöljük, akkor a következő két esemény kizárja egymást:

$$A = \{i + j \leq 4\} \quad \text{és} \quad B = \{i + j \geq 10\}.$$

Nem zárja ki egymást viszont a következő két esemény:

$$A = \{i + j \leq 4\} \quad \text{és} \quad B = \{i + j = 2k\}, \quad (3.2)$$

ahol  $k$  egész szám. Közös részük ugyanis nem üres:

$$AB = \{(1, 1); (1, 3); (2, 2); (3, 1)\}. \quad (3.2a)$$

Ezek után ki tudjuk mondani a következő tételt:

3.1. TÉTEL. Ha  $A$  és  $B$  egymást kizáró események,

$$p(A + B) = p(A) + p(B). \quad (3.3)$$

Ha nem egymást kizárók, a tétel módosul:

$$p(A + B) = p(A) + p(B) - p(AB). \quad (3.3a)$$

A 3.1. TÉTEL általánosítható tetszőleges számú, egymást páronként kizáró eseményre. Számuk lehet megszámlálhatóan végtelen is.

Azt az eseményt, amelyben  $A$  nem következik be, felülhúzással jelöljük. Mivel az  $A$  esemény vagy bekövetkezik, vagy sem,

$$A + \bar{A} = \Omega,$$

A 3.1. TÉTEL szerint ebből viszont az adódik, hogy

$$p(\bar{A}) = p(\Omega) - p(A) = 1 - p(A). \quad (3.3b)$$

A következő alapvető fogalom az *események függetlensége*:

3.1. DEFINÍCIÓ. Az  $A$  és  $B$  eseményeket akkor mondjuk függetlennek, ha

$$p(AB) = p(A) \cdot p(B). \quad (3.4)$$

E definíció szerint tehát független események együttes bekövetkezésének a valószínűségét megkapjuk, ha külön-külön való bekövetkezésük valószínűségét összeszorozzuk. Az események általában nem függetlenek. Ezért szükségünk van a *feltételes valószínűség* fogalmára:

3.2. DEFINÍCIÓ. Az  $A$  eseménynek a  $B$  eseményre vonatkozó feltételes valószínűségét a következő képlet adja meg:

$$p(A|B) = \frac{p(AB)}{p(B)}. \quad (3.5)$$

Az itt szereplő  $p(A|B)$  feltételes valószínűség fogalmát egy példával világítjuk meg. Tekintsük a (3.2) szerint definiált eseményeket. Az együttes bekövetkezésüket jelentő eseményt (3.2a)-ban felírtuk. Nos, a kérdéses feltételes valószínűséget szintén a (3.1) határértékkel definiáljuk, de az ott szereplő  $n$ -be csak azokat a kísérleteket számítjuk bele, amelyekben a  $B$  esemény bekövetkezett.  $k$ -ba természetesen azokat a kísérleteket számoljuk bele, amelyekben az  $A$  esemény is bekövetkezett. Nagyon gyakran könnyebb a feltételes valószínűséget kiszámítani, mint az együttes bekövetkezését. Ezért hasznos a (3.5) képlet. Végül megjegyezzük, hogy független eseményekre vonatkozóan a feltételes valószínűség a feltétel nélküli valószínűséggel egyezik meg. (3.4) alapján ugyanis írhatjuk:

$$p(A|B) = \frac{p(AB)}{p(B)} = \frac{p(A) \cdot p(B)}{p(B)} = p(A).$$

#### Azonos valószínűségű elemi események

Levezetünk egy hasznos összefüggést, amely akkor érvényes, amikor az elemi események száma véges és valószínűségük azonos. Ha az  $\Omega$  halmaz elemeinek a száma  $N$ , akkor az egyes elemi események  $p_0$  valószínűségére a 3.1. TÉTEL általánosítása szerint fennáll:

$$Np_0 = p(\Omega) = 1,$$

amiből

$$p_0 = \frac{1}{N}.$$

Ha az  $A$  eseményt alkotó elemi események száma  $K$ , akkor ugyanezzel a megfontolással kapjuk:

$$p(A) = Kp_0 = \frac{K}{N}. \quad (3.6)$$

Szavakban:

3.2. TÉTEL. Ha az elemi események száma véges, és valószínűségük azonos, akkor bármely esemény valószínűségét megadja a kedvező események és az összes események számának (vagyis  $K$ , illetve  $N$ ) hányadosa.

Nézzük például a (3.2) szerint definiált események valószínűségét! Az összes események száma:  $N = 36$ . Az  $A$  eseményt a következő elemi események alkotják:

$$A = \{(1, 1); (1, 2); (1, 3); (2, 1); (2, 2); (3, 1)\},$$

amelyek száma  $K = 6$ , tehát (3.6) alapján

$$p(A) = \frac{6}{36} = \frac{1}{6}.$$

Könnyen megszámlálhatjuk, hogy a  $B$  esemény számára kedvező események száma  $K = 18$ , tehát (3.6) alapján

$$p(B) = \frac{18}{36} = \frac{1}{2}.$$

A fogalmak jobb megértése céljából vizsgáljuk meg, hogyan teljesül a feltételes valószínűség definíciójával szolgáló (3.5) összefüggés. (3.2a) szerint az  $AB$  együttes esemény számára a kedvező esetek száma  $K = 4$ , vagyis

$$p(A|B) = \frac{4/36}{1/2} = \frac{2}{9}.$$

Ugyanezt közvetlenül is megkaphattuk volna. Ha ugyanis  $\Omega$ -t leszűkítjük  $B$ -re, akkor ezen az összes események száma már csak 18. Az ezen a részhalmazon az  $A$  esemény számára kedvező elemi események száma (3.2a) alapján 4. Így a feltételes valószínűség  $4/18 = 2/9$ .

### Geometriai valószínűség

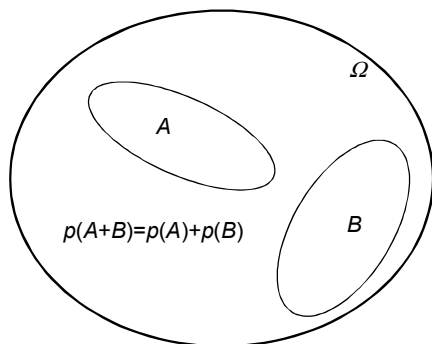
A geometriai valószínűség rokon az előző szakaszban tárgyaltakkal. Legyenek az elemi események egy véges  $T$  területű ponthalmaz pontjai.<sup>2</sup> Feltesszük, hogy a kísérletben minden pont azonos valószínűséggel jön ki. Ezt pontosabban is meg kell fogalmaznunk: ha az  $A$  eseménynek megfelelő alakzat területe  $t(A)$ , akkor  $A$  bekövetkezésének a valószínűsége

$$p(A) = \frac{t(A)}{T}. \quad (3.7)$$

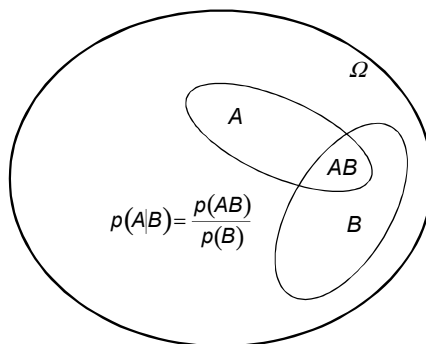
Ez a *geometriai valószínűség* fogalma. Erre vonatkozóan a 3.2a. és 3.2b. ábrák szemléltetik az egymást kizáró eseményeket és a feltételes geometriai valószínűséget. Az előbbi ábráról leolvasható, hogy az  $(A + B)$  ponthalmaz területe a két terület összege, tehát valószínűségük (3.7) szerint összeadódik. Az is látszik az ábráról, hogy események halmazelméleti összeadása a “VAGY” logikai kapcsolatnak felel meg: vagy  $A$ , vagy  $B$  következik be. A másik ábráról viszont az látszik, hogy a halmazelméleti szorzás az “ÉS” logikai kapcsolatot jelenti: mind  $A$ , mind  $B$  bekövetkezik. Az együttes bekövetkezés  $p(AB)$

<sup>2</sup> Az alábbiak könnyen átvihetők egy- vagy háromdimenziós alakzatokra is. A szemléletesség kedvéért választottuk a síkbeli ponthalmazokat.

valószínűsége a két ponthalmaz közös részének a területével arányos. A feltételes valószínűséget így a következőképpen foghatjuk fel: *feltéve*, hogy a kísérlet kimenetele által kiválasztott pont a  $B$  halmazba esik, keressük az  $A$  esemény bekövetkezésének a valószínűségét. Ekkor (3.7)-ben  $T$  szerepét  $p(B)$ , a  $T(A)$  területét pedig  $p(AB)$  játssza. Ezzel adódik (3.5).



3.2a. ábra. Egymást kizáró események geometriai valószínűsége



3.2b. ábra. Feltételes geometriai valószínűség szemléltetése

### Valószínűségi változó, eloszlásfüggvény

Amikor méréseket végzünk, a kísérlet valamilyen fizikai mennyiség mérését jelenti. A kísérlet kimenetétől függően a mérés eredménye más és más lehet. A mérés minden elemi eseményhez egy vagy több számot rendel hozzá. A mérések kiértékelése szempontjából tehát alapvető ennek a hozzárendelésnek ismerete. Így jutunk el a *valószínűségi változó* fogalmához, amelyet a mérésekben játszott szerepénél általánosabban határozzunk meg:

3.3. DEFINÍCIÓ. A valószínűségi változó az  $\Omega$  eseménytéren értelmezett mérhető<sup>3</sup> függvény.

A fogalmat két példával világítjuk meg a fentiek közül:

- Amikor két kockával való dobáskor az  $(i, j)$  számpár jön ki, ezek bármilyen függvénye valószínűségi változó, például  $i + j$ ,  $i/j$  stb.
- Céllövés esetében az  $\Omega$  eseményteret a céltábla pontjai alkotják. A céllövő szempontjából a legfontosabb valószínűségi változó a golyó becsapódási pontjának a tábla középpontjától való  $r$  távolsága.

Gyakran beleesünk abba a fogalmazási hibába, hogy nem különböztetjük meg az elemi eseményeket a valószínűségi változó hozzájuk rendelt értékétől. Mérések esetében ez gyakran megbocsátható pongyolaság. Vegyük példának a szoba hőmérsékletének a mérését. Ha a mérést többször megismételjük, általában különböző eredményeket kapunk, tehát a mért hőmérséklet a véletlentől függ. Hajlamosak vagyunk az elemi események  $\Omega$  halmazát a kapott hőmérsékletértékekkel azonosítani. Tudjuk persze, hogy nem erről van szó. Némi gondolkodás után azonban belátjuk, hogy ebben az esetben nem is olyan

<sup>3</sup> A mérhető függvény fogalmát a későbbiekben határozzuk meg.

könnyű az elemi eseményeket definiálni, hiszen azok számos tényező együttesét jelentik – természetesen attól függően, hogyan végeztük a mérést: mikor mérünk, nyitva voltak-e az ablakok, milyen a hőmérő pontossága, volt-e fűtés, stb. Ha ezek a tényezők mind szerepet játszanak a mérés kiértékelésében, kell törődnünk az elemi események pontos definíciójával. Ellenkező esetben az említett pongyolaság megbocsátható.

A valószínűségi változók legfontosabb jellemzője az *eloszlásfüggvény*:

3.4. DEFINÍCIÓ. A valószínűségi változó  $F(x)$  eloszlásfüggvénye annak a valószínűsége, hogy a változó  $\xi$  értéke kisebb  $x$ -nél:

$$F(x) = P\{\xi < x\}. \quad (3.8)$$

A kapcsos zárójelen belül szereplő reláció kijelöli az  $\Omega$  eseménytér egy részhalmazát. Ezeket a valószínűségi változó *nívóhalmazainak* nevezzük, amelyek definíció szerint maguk is események. A 3.3. DEFINÍCIÓban szereplő *mérhetőség* azt jelenti, hogy a nívóhalmazokhoz  $x$  minden értékénél kell tudni valószínűséget definiálni. Nagyon nehéz matematikai feladat nemmérhető halmazt konstruálni, így a mérések kiértékelésében mindig feltesszük, hogy a szereplő valószínűségi változók mérhető függvények. Ebben a jegyzetben – kevés kivételtől eltekintve – általában görög betűkkel jelöljük a valószínűségi változókat, az eloszlásfüggvény változóját pedig egy neki megfelelő latin betűvel, mint ezt (3.8)-ban is tettük.

A definícióból következik, hogy egy eloszlásfüggvény mindig monoton növekvő. Ami folytonosságát illeti, ez függ a valószínűségi változó jellegétől. A minket érdeklő esetekben a valószínűségi változók kétfélek lehetnek: *diszkrét* és *folytonos* változók.

3.5. DEFINÍCIÓ. A  $\xi$  valószínűségi változó diszkrét, ha értékei csak a megszámlálhatóan sok  $x_k$  szám valamelyike lehet ( $k = 1, 2, \dots$ ).

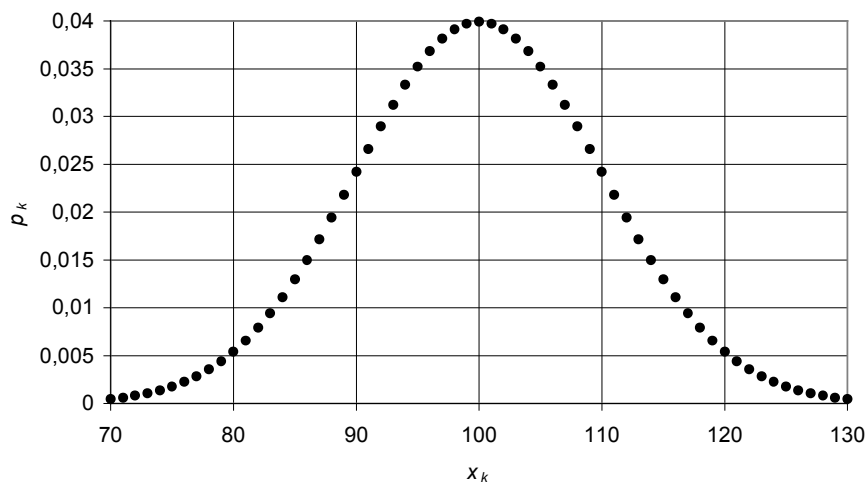
Ebben az esetben az eloszlásfüggvény két szomszédos  $x_k$  közötti intervallumban állandó, de az  $x = x_k$  pontokban ugrása van. Az ugrás nagyságát  $p_k$ -val jelöljük, ami annak a valószínűségét adja meg, hogy a  $\xi = x_k$  esemény bekövetkezzen:

$$p_k = P\{\xi = x_k\}. \quad (3.9a)$$

Diszkrét eloszlások esetében tehát az eloszlásfüggvényt a következőképpen írhatjuk fel:

$$F(x) = P\left\{ \sum_{k: x_k < x} p_k \right\}. \quad (3.9b)$$

Be lehet látni, hogy ez függvény *balról folytonos*.



3.3. ábra. Diszkrét eloszlás grafikonja

A folytonos valószínűségi változó fogalmát legegyszerűbb a diszkrét eloszlásokból kiindulva meghatározni. Ebben Gnyegyenko és Hincsin [1] gondolatmenetét követjük.<sup>4</sup> A 3.3. ábrán egy diszkrét eloszlásra ábrázoltuk a  $p_k$  valószínűségeket a változó értékészletét alkotó  $x_k$  értékek függvényében. Az ábrán látható pontok egy folytonos görbévé látszanak összeolvadni. Ezt a következőképpen tudjuk matematikailag is megfogalmazni. Kijelölünk egy  $[x, x+dx)$  intervallumot, és összegezzük az ebbe eső  $x_k$  értékekhez tartozó valószínűségeket. Legyen  $f(x)$  ezek átlagértéke:

$$\sum_{k: x \leq x_k < x+dx} p_k = f(x)dx.$$

Ha az  $x_k$  értékek minden határon túl sűrűsödnek az  $x$ -tengelyen, akkor ezzel eljutunk a folytonos valószínűségi változó fogalmához. Az  $f(x)$  függvény a  $\xi$  valószínűségi változó *sűrűségfüggvénye*. Ha a 3.3. ábrán látható pontokat egy folytonos görbével kötjük össze, akkor az  $F(x)$  eloszlásfüggvény ennek  $-\infty$ -tól  $x$ -ig terjedő része alatt alatti területet adja meg. Így jutunk a következő definícióhoz:

3.6. DEFINÍCIÓ. A  $\xi$  valószínűségi változó folytonos, ha eloszlásfüggvénye felírható az

$$F(x) = \int_{-\infty}^x f(t)dt \quad (3.10a)$$

integrál alakjában.<sup>5</sup>

<sup>4</sup> Ez a gondolatmenet azoknak szól elsősorban, akik még nem szereztek kellő jártasságot az integrálok területén.

<sup>5</sup> A valószínűség-elméletben ennél jóval általánosabban definiálják a folytonos eloszlásokat. Az itt adott definíció valójában a *totálfolytonos* valószínűségi változók definíciója. Az Olvasótól elvárt matematikai előismeretekre való tekintettel egyszerűsítettük a definíciót.



Az  $\Omega$  halmaz ekkor vagy egy (véges vagy végtelen) intervallum, vagy ilyenek egyesítése. (3.10a)-ból következik, hogy az  $f(x)$  sűrűségfüggvény az eloszlásfüggvény deriváltja:

$$f(x) = \frac{dF(x)}{dx}. \quad (3.10b)$$

A továbbiakban – hacsak lehet – először a diszkrét változók alapján fogjuk a fogalmakat bevezetni, és csak ezután adjuk meg ezeknek a folytonos eloszlásokra vonatkozó megfelelőit.

A 3.5. és 3.6. DEFINÍCIÓKBÓL következik, hogy

$$\lim_{x \rightarrow +\infty} F(x) = 1,$$

hiszen  $F(+\infty)$  annak a valószínűségét jelenti, hogy  $\xi$  egyáltalán felvesz valamilyen valós értéket, ami nyilvánvalóan a biztos eseménnyel azonos. Korábbi összefüggéseink alapján ez a következőt is jelenti. Diszkrét valószínűségi változók esetében

$$\sum_{k=1}^{\infty} p_k = 1, \quad (3.11a)$$

illetve folytonos valószínűségi változók esetében

$$\int_{-\infty}^{+\infty} f(x) dx = 1. \quad (3.11b)$$

### Várható érték és szórás

Egy *diszkrét* valószínűségi változó értékét  $n$ -szer megmértük. Tegyük fel, hogy az  $x_k$  érték  $l_k$ -szor jött ki. Ha  $n \rightarrow \infty$ , akkor definíció szerint

$$p_k = \lim_{n \rightarrow \infty} \frac{l_k}{n}.$$

Vegyük a kapott eredmények átlagát:

$$\bar{\xi} = \frac{\sum_k l_k x_k}{n}.$$

Az összegzés itt  $k$  minden, a mérésekben előforduló értékre kiterjed. Ennek a mennyiségnek  $n \rightarrow \infty$  mellett vett határértékét nevezzük a  $\xi$  valószínűségi változó *várható értékének*:

$$a = M(\xi) = \lim_{n \rightarrow \infty} \bar{\xi} = \sum_{k=1}^{\infty} p_k x_k. \quad (3.12a)$$

Folytonos valószínűségi változó esetében – analóg megfontolásokkal – a következő definíciót kapjuk a várható értékre:

$$a = M(\xi) = \int_{-\infty}^{+\infty} xf(x)dx . \quad (3.12b)$$

Ezekben a definíciókban természetesen feltételezzük, hogy a végtelen összeg, illetve integrál konvergens. A várható értéket gyakran egyszerűen csak *átlagértéknek* nevezzük, hiszen ennek a határértéke. Az  $M(\dots)$  jelölés is erre utal: “mean” angolul átlagot jelent.<sup>6</sup>

Mivel a valószínűségi változónak az egyes kísérletekben kapott értéke az átlagtól eltérhet, szükség van egy olyan jellemzőre is, amely ennek a nagyságát jellemzi. Első ötletként erre kézenfekvő a  $(\xi - a)$  különbség átlagát választani. Mint könnyen belátható, ez minden esetben zérus. Nem zérus azonban a különbség négyzetének az átlaga, amelyet *szórásnégyzetnek* nevezünk:

$$\sigma^2 = D^2(\xi) = \sum_{k=1}^{\infty} (x_k - a)^2 p_k , \quad (3.13a)$$

illetve folytonos valószínűségi változók esetében

$$\sigma^2 = D^2(\xi) = \int_{-\infty}^{+\infty} (x - a)^2 f(x)dx . \quad (3.13b)$$

A “szórásnégyzet” mellett használatos még a *variancia* vagy a *diszperzió* kifejezés is.<sup>7</sup> Négyzetgyökét *szórásnak* nevezzük, szokásos jelölése  $\sigma$ . A fenti definíciókban hallgatólagosan ismét feltettük, hogy a fellépő integrálok, illetve összegek konvergenssek. Ha (3.13)-ban a négyzetes tényezőt kifejtjük, egyszerűen levezethetjük a

$$D^2(\xi) = M(\xi^2) - [M(\xi)]^2 \quad (3.14)$$

összefüggést. Mivel a szórásnégyzet mindig pozitív, azt is kiolvashatjuk ebből a képletből, hogy egy valószínűségi változó négyzetének az átlaga nagyobb, mint a változó átlagának a négyzete.

Egyszerűen beláthatjuk, hogy a várható érték és a szórás arányosan változik, ha a valószínűségi változót egy  $c$  állandóval beszorozzuk:

$$M(c\xi) = cM(\xi) = ca \quad \text{és} \quad D^2(c\xi) = c^2 D^2(\xi) = c^2 \sigma^2 .$$

A szórás az eloszlásfüggvény hasznos jellemzője, amelyet meghatározni a mérések kiértékelésének központi feladata. Ebben az alfejezetben csak a

<sup>6</sup> Számos más nyelven is “m” az első betű: moyenne (francia), Mittelwert (német), medio (spanyol) stb.

<sup>7</sup> Az utóbbira utal a  $D^2(\dots)$  jelölésben a “D” betű.

*Csebisev-egyenlőtlenséget* idézzük, amely azt fejezi ki, hogy a várható értéktől a szóráshoz képest nagy eltérések nem valószínűek:

3.3. TÉTEL. Ha a  $\xi$  valószínűségi változó várható értéke  $a$ , szórása  $\sigma$ , akkor tetszőleges pozitív  $\lambda$ -ra fennáll, hogy

$$P\{|\xi - a| > \lambda\sigma\} \leq \frac{1}{\lambda^2}. \quad (3.15)$$

A tételt diszkrét valószínűségi változóra látjuk be, de érvényes folytonos valószínűségi változók esetében is. (3.13a) alapján írhatjuk:

$$\sigma^2 = \sum_{k=1}^{\infty} (x_k - a)^2 p_k \geq \sum_{k:|x_k - a| > \lambda\sigma} (\lambda\sigma)^2 p_k = (\lambda\sigma)^2 P\{|\xi - a| > \lambda\sigma\}$$

Ebből egyszerűen következik a bizonyítandó tétel. Ha  $\lambda$ -t úgy választjuk meg, hogy  $\lambda\sigma = \varepsilon$  legyen, akkor az egyenlőtlenség a

$$P\{|\xi - a| > \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}. \quad (3.15a)$$

alakra hozható. Gyakran ebben az alakjában alkalmazzuk.

### Magasabb momentumok

A (3.12) és (3.13) képletekkel definiált várható érték, illetve szórásnégyzet általánosításaként definiálhatunk további momentumokat. Diszkrét valószínűségi változó  $n$ -edik *momentuma*

$$M_n = M(\xi^n) = \sum_{k=1}^{\infty} x_k^n p_k, \quad (3.16a)$$

ha ez a sor konvergens. A (3.13) képletekkel definiált szórásnégyzet ún. *centrális momentum*, amelynek természetes általánosítása a

$$C_n = M[(\xi - a)^n] = \sum_{k=1}^{\infty} (x_k - a)^n p_k \quad (3.16b)$$

$n$ -edik centrális momentum, ha ez a sor konvergens.

Ugyanezeknek a mennyiségeknek a definícióját könnyen vihetjük át folytonos eloszlásokra is:

$$M_n = M(\xi^n) = \int_{-\infty}^{+\infty} x^n f(x) dx \quad (3.16c)$$

és

$$C_n = M[(\xi - a)^n] = \int_{-\infty}^{+\infty} (x - a)^n f(x) dx, \quad (3.16d)$$

ha ezek az integrálok léteznek.

## Többszámú eloszlások<sup>8</sup>

### Együttes eloszlásfüggvény

Az eddigieket általánosíthatjuk több valószínűségi változó esetére. Jelöljük ezeket  $\xi_1, \xi_2, \dots, \xi_n$ -vel. Az egyszerűség kedvéért mindegyikről feltesszük, hogy folytonosak. Nem okoz nehézséget a diszkrét változók esetére való áttérés. *Együttes eloszlásfüggvényüket* (3.8) mintájára definiáljuk:

3.7. DEFINÍCIÓ. A  $\xi_1, \xi_2, \dots, \xi_n$  valószínűségi változók együttes eloszlásfüggvényét a következő képlet adja meg:

$$F(x_1, x_2, \dots, x_n) = P\{\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n\}. \quad (3.17)$$

Deriváltja az *együttes sűrűségfüggvény* [vö. (3.10b)]:

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}. \quad (3.18)$$

A sűrűségfüggvénynek a következő értelmet lehet adni. Jelöljünk ki az  $n$ -dimenziós térben egy  $dV = dx_1 dx_2 \dots dx_n$  infinitezimális térfogatelemet, és keressük annak a valószínűségét, hogy a  $(\xi_1, \xi_2, \dots, \xi_n)$  számok által a térben kijelölt pont ebbe esik. (3.18) szerint ezt első rendben az

$$f(x_1, x_2, \dots, x_n) dV = f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$$

kifejezés adja meg, ahol  $x_1 = \xi_1, x_2 = \xi_2, \dots, x_n = \xi_n$ . Amikor a geometriai valószínűséget definiáltuk, feltételeztük, hogy ez a valószínűség független a  $(\xi_1, \xi_2, \dots, \xi_n)$  ponttól. Ha tehát a rendelkezésre álló térfogat  $V_\Omega$ , akkor (3.7) szerint a geometriai valószínűség  $dV/V_\Omega$ , ami azt jelenti, hogy ekkor az együttes sűrűségfüggvény

$$f(x_1, x_2, \dots, x_n) = \frac{1}{V_\Omega}.$$

Ezt az eloszlást *egyenletes eloszlásnak* nevezzük. Nyilvánvaló, hogy csak akkor tudjuk értelmezni, amikor  $\Omega$  térfogata véges.

A definícióból következik, hogy az együttes sűrűségfüggvénynek  $\Omega$ -ra vett integrálja 1. Tegyük fel, hogy ismerjük a  $(\xi_1, \xi_2, \dots, \xi_n)$  valószínűségi változók együttes sűrűségfüggvényét, de egyikük (például  $\xi_1$ ) számunkra érdektelen. Hogyan lehet ebből a maradék  $(n-1)$  valószínűségi változó együttes sűrűségfüggvényét kiszámítani? Ha az eloszlásfüggvényben  $x_1$  helyére  $+\infty$ -t helyettesítünk, akkor a (3.17) szerint független  $\xi_1$  értékétől, és végeredményben

<sup>8</sup> Tekintve, hogy ebben a jegyzetben a többszámú eloszlások közül többnyire csak a Gauss-eloszlásra (lásd alább) van szükség, amely folytonos eloszlás, a fogalmakat folytonos valószínűségi változókra definiáljuk.

$(\xi_2, \dots, \xi_n)$  eloszlásfüggvényévé válik. A sűrűségfüggvény tekintetében ez az  $x_1$  változó szerint való integrálást jelent:

$$f(x_2, \dots, x_n) = \int_{-\infty}^{+\infty} f(x_1, x_2, \dots, x_n) dx_1.$$

Az ilyen típusú integrálokat a fennmaradó  $(\xi_2, \dots, \xi_n)$  változók *perem-sűrűségfüggvényének* nevezzük.

Az események függetlenségére vonatkozó definíciót egyszerűen átvihetjük a valószínűségi változókra is:

3.8. DEFINÍCIÓ. A  $\xi_1, \xi_2, \dots, \xi_n$  valószínűségi változókat függetlennek nevezzük, ha együttes eloszlásfüggvényük a következőképpen bontható tényezőkre:

$$F(x_1, x_2, \dots, x_n) = F_1(x_1)F_2(x_2) \dots F_n(x_n). \quad (3.19a)$$

(3.18)-ből következik, hogy ekkor az együttes sűrűségfüggvény is hasonlóképpen bontható tényezőkre:

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2) \dots f_n(x_n). \quad (3.19b)$$

### Várható érték és szórás

Az egyes valószínűségi változók várható értékét és szórását az egyetlen valószínűségi változó esetére használt meghatározások szerint definiálhatjuk. (3.12b) mintájára  $\xi_i$  várható értéke (átlaga):

$$a_i = M(\xi_i) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} x_i f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n, \quad (3.20a)$$

( $i = 1, 2, \dots, n$ ). Hasonlóan definiálhatjuk a szórásnégyzetet is:

$$\begin{aligned} \sigma_i^2 &= D^2(\xi_i) = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} (x_i - a_i)^2 f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n, \end{aligned} \quad (3.20b)$$

( $i = 1, 2, \dots, n$ ). Egyszerűen belátható, hogy ez egyben  $\xi_i$  perem-eloszlásának az átlaga, illetve szórásnégyzete is. Ez az észrevétel azt jelenti, hogy egyetlen változó várható értékének és szórásának a meghatározásához elegendő a minket érdeklő változót egyedül megfigyelni, hiszen a többi változó hatása benne van a perem-eloszlásban.

### Kovariancia

Többdimenziós eloszlások esetében fellép a *kovariancia*, amely egyetlen változó esetében – értelemszerűen – nem definiálható.

3.9. DEFINÍCIÓ. A  $\xi_i$  és  $\xi_j$  valószínűségi változók kovarianciáját a

$$\begin{aligned} \text{cov}(\xi_i, \xi_j) &= M[(\xi_i - a_i)(\xi_j - a_j)] = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} (x_i - a_i)(x_j - a_j) f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned} \quad (3.21)$$

képlet adja meg.

Nyilvánvaló, hogy  $i = j$  esetén a kovariancia azonos a szórásnégyzettel.  $i \neq j$  esetén viszont a kovariancia némi felvilágosítást ad a két változó függetlenségére vonatkozóan. Fennáll ugyanis az

3.4. TÉTEL. Ha a  $\xi_i$  és  $\xi_j$  valószínűségi változók függetlenek, kovarianciájuk eltűnik.

Helyettesítsük (3.19b)-t (3.21)-be:

$$\begin{aligned} \text{cov}(\xi_i, \xi_j) &= \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} (x_i - a_i)(x_j - a_j) f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_i - a_i)(x_j - a_j) f(x_i, x_j) dx_i dx_j = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_i - a_i)(x_j - a_j) f_i(x_i) f_j(x_j) dx_i dx_j = \\ &= \int_{-\infty}^{+\infty} (x_i - a_i) f_i(x_i) dx_i \cdot \int_{-\infty}^{+\infty} (x_j - a_j) f_j(x_j) dx_j = 0. \end{aligned}$$

*A tétel megfordítása nem érvényes:* két valószínűségi változó kovarianciája úgy is eltűnhet, hogy azok nem függetlenek.

A kovariancia nagyságát korlátozza az

3.5. TÉTEL. A  $\xi_i$  és  $\xi_j$  valószínűségi változók kovarianciájának abszolút értéke nem lehet nagyobb, mint szóráruk szorzata:

$$|\text{cov}(\xi_i, \xi_j)| \leq \sigma_i \sigma_j. \quad (3.22)$$

Legyen  $\xi = \xi_i - a_i$  és  $\eta = \xi_j - a_j$ . Nyilván ekkor  $M(\eta) = M(\xi) = 0$ .  $\lambda$  tetszőleges valós számértéke mellett fennáll a

$$0 \leq D^2(\xi - \lambda\eta) = D^2(\xi) - 2\lambda M(\eta\xi) + \lambda^2 D^2(\eta),$$

egyenlőtlenség. Ennek a polinomnak a determinánsa nyilván nemnegatív, vagyis

$$|M(\eta\xi)| \leq D(\xi) \cdot D(\eta),$$

amint a tétel állítja. Egyenlőség akkor és csak akkor állhat fent, ha van olyan  $\lambda$ , amelyre

$$\xi - \lambda\eta = \text{konstans}.$$

A most bizonyított tételt *Schwarz-féle egyenlőtlenség*nek nevezzük.<sup>9</sup>

A kovariancia és a szórások

$$r = \frac{\text{cov}(\xi_i, \xi_j)}{\sigma_i \sigma_j}$$

hányadosát *korrelációs együtthatónak* nevezzük. A 3.5. TÉTEL szerint  $r$  abszolút értéke nem lehet nagyobb 1-nél.

### Feltételes sűrűségfüggvény

(3.5) alapján definiálhatjuk a feltételes sűrűségfüggvényt. Tekintsünk két valószínűségi változót:  $\xi$  és  $\eta$ . Együttes sűrűségfüggvényük  $f(x, y)$ .  $\eta$  *perem-sűrűségfüggvénye*

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

A 3.2. DEFINÍCIÓ alkalmazásához az  $A$  és  $B$  események legyenek a következők:

$$A = \{x \leq \xi < x + dx\} \quad \text{és} \quad B = \{y \leq \eta < y + dy\}.$$

Definíció szerint

$$P(AB) = f(x, y) dx dy \quad \text{és} \quad P(B) = f(y) dy.$$

A (3.5) szerinti feltételes valószínűség ezzel így írható:

$$P(A|B) = \frac{f(x, y) dx dy}{f(y) dy} = \frac{f(x, y)}{f(y)} dx = f(x|\eta = y) dx.$$

Ezen alapul az

3.10. DEFINÍCIÓ. A  $\xi$  valószínűségi változónak az  $\eta$  valószínűségi változóra vonatkozó feltételes sűrűségfüggvénye

$$f(x|\eta = y) = \frac{f(x, y)}{f(y)}. \quad (3.23)$$

<sup>9</sup> A tételt sok nemzet tekinti magáénak: a franciák Cauchy-ról, az oroszok Bunyakovszkijről nevezték el, ami nem csoda, hiszen levezetése annyira egyszerű, hogy többen is megkaphatták egymástól függetlenül. Így Cauchy-Schwarz-Bunyakovszkij-féle egyenlőtlenségnek kellene neveznünk. Az egyszerűség kedvéért a legrövidebb nevű szerzőt választottuk.

A feltétel jelölését néha egyszerűsítjük:  $f(x|y)$ .

### Vektori jelölésmód

A többdimenziós eloszlások esetében kényelmes a vektori jelölésmód. A  $\xi_1, \xi_2, \dots, \xi_n$  valószínűségi változókat a  $\bar{\xi}$  vektor komponenseinek tekintjük. Hasonlóan az  $\mathbf{x}$  és  $\mathbf{a}$  vektorokban egyesítjük az  $(x_1, x_2, \dots, x_n)$ , illetve az  $(a_1, a_2, \dots, a_n)$  változókat. Ezekkel a jelölésekkel a (3.20a) képletek az egyszerűbb

$$\mathbf{a} = M(\bar{\xi}) = \int \mathbf{x} f(\mathbf{x}) d\mathbf{x} \quad (3.24)$$

alakban írhatók fel. Itt az integrált az  $f(\mathbf{x})$  függvény teljes értelmezési tartományra ki kell terjeszteni. Ha az integrálási tartomány más, akkor azt értelemszerűen jelöljük.

A  $\bar{\xi}$  vektor két komponense között a (3.21)-ben definiált kovarianciát minden lehetséges  $(i, j)$  indexpárra képezzük, és a kapott kovarianciákat a  $\mathbf{B}$  kovarianciamátrix  $(i, j)$  elemének tekintjük:

$$B_{ij} = \text{cov}(\xi_i, \xi_j). \quad (3.25)$$

Ha szükséges, a kovarianciamátrix  $\mathbf{B}$  jelöléséhez indexben tüntetjük fel, melyik véletlen vektorhoz tartozik. A definícióból következik, hogy  $\mathbf{B}$  szimmetrikus. A vektori jelölésmód segítségével megmutatjuk, hogy ennél több is igaz:

3.6. TÉTEL. Minden kovarianciamátrix pozitív szemidefinit, szimmetrikus, vagyis tetszőleges  $\mathbf{z}$  vektorra fennáll, hogy

$$\mathbf{z}^T \mathbf{B} \mathbf{z} \geq 0. \quad (3.26)$$

(3.21) szerint  $\xi_i$  és  $\xi_j$  kovarianciája a  $(\xi_i - a_i)(\xi_j - a_j)$  szorzat várható értéke.

Ez a szorzat azonban tekinthető a  $(\bar{\xi} - \mathbf{a})(\bar{\xi} - \mathbf{a})^T$  diád  $(i, j)$  elemének is. Ebből következik, hogy a (3.25) képlet vektori alakja

$$\mathbf{B} = M \left[ (\bar{\xi} - \mathbf{a})(\bar{\xi} - \mathbf{a})^T \right], \quad (3.27)$$

amire sűrűn hivatkozunk ebben a jegyzetben. A  $\mathbf{z}^T (\bar{\xi} - \mathbf{a})$  skalárszorzat zérus várható értékű valószínűségi változó bármilyen konstans<sup>10</sup>  $\mathbf{z}$  vektor esetében. Négyzetének várható értéke nem lehet negatív:

$$0 \leq M \left\{ \left[ \mathbf{z}^T (\bar{\xi} - \mathbf{a}) \right]^2 \right\} = M \left\{ \mathbf{z}^T (\bar{\xi} - \mathbf{a})(\bar{\xi} - \mathbf{a})^T \mathbf{z} \right\} =$$

<sup>10</sup> Ebben az összefüggésben a "konstans" azt jelenti, hogy *nem valószínűségi változó*.



$$= \mathbf{z}^T \mathbf{M} \left\{ (\bar{\xi} - \mathbf{a})(\bar{\xi} - \mathbf{a})^T \right\} \mathbf{z} = \mathbf{z}^T \mathbf{B} \mathbf{z}.$$

Éppen ezt kellett bizonyítani.

### Transzformált változók várható értéke és kovarianciája

Mérések kiértékelésekor a közvetlenül mért mennyiségekből további valószínűségi változókat számítunk ki, más szóval *transzformáljuk* őket. Az alábbiakban a *lineáris* transzformációkat tekintjük, vagyis meghatározzuk a transzformált mennyiségek várható értékét és kovarianciamátrixát. Írjuk a transzformációt az

$$\bar{\eta} = \mathbf{A} \bar{\xi}$$

alakba. Várható értékét (3.24)-ből kapjuk:

$$\mathbf{b} = \mathbf{M}(\bar{\eta}) = \mathbf{M}(\mathbf{A} \bar{\xi}) = \int \mathbf{A} \mathbf{x} f(\mathbf{x}) d\mathbf{x} = \mathbf{A} \int \mathbf{x} f(\mathbf{x}) d\mathbf{x} = \mathbf{A} \mathbf{a}. \quad (3.28)$$

Hasonlóan egyszerű kiszámítani a transzformált változók kovarianciamátrixát. (3.27) alapján írhatjuk:

$$\begin{aligned} \mathbf{B}_{\bar{\eta}} &= \mathbf{M} \left[ \mathbf{A} (\bar{\xi} - \mathbf{a})(\bar{\xi} - \mathbf{a})^T \mathbf{A}^T \right] = \\ &= \mathbf{A} \mathbf{M} \left[ (\bar{\xi} - \mathbf{a})(\bar{\xi} - \mathbf{a})^T \right] \mathbf{A}^T = \mathbf{A} \mathbf{B}_{\bar{\xi}} \mathbf{A}^T. \end{aligned} \quad (3.29)$$

Ha  $\mathbf{A}$  annak az  $\mathbf{U}$  mátrixnak a transzponáltja, amely (2.3) szerint a kovarianciamátrixot diagonálisra transzformálja, akkor az  $\bar{\eta}$  vektor komponensei *korrelálatlanok*. Mint fentebb említettük, ez nem feltétlenül jelent függetlenséget is.

Az  $\mathbf{A}$  transzformációs mátrixról nem szükséges kikötnünk, hogy négyzetes legyen. Szélső esetben lehet akár egy vektor is. Az imént kapott képletek alapján fontos tételeket bizonyíthatunk be, ha  $\mathbf{A} = \bar{\omega}^T = (1, 1, \dots, 1)$ . Ekkor a transzformáció egyetlen valószínűségi változót eredményez:

$$\eta = \bar{\omega}^T \bar{\xi} = \sum_{i=1}^n \xi_i.$$

(3.28) alapján érvényes az

3.7. TÉTEL. Valószínűségi változók összegének a várható értékét tagonként lehet képezni:

$$\mathbf{M} \left( \sum_{i=1}^n \xi_i \right) = \sum_{i=1}^n \mathbf{M}(\xi_i). \quad (3.30)$$

$\eta$  szórásnégyzetét (3.29) alapján számíthatjuk ki. Az eredményt nem írjuk fel az általános esetben. Fontos tételt kapunk azonban független valószínűségi vál-

tozókra. Ebben az esetben ugyanis a kovarianciamátrix diagonális. Fennáll tehát az

3.8. TÉTEL. Független valószínűségi változók összegének a szórásnégyzetét tagonként lehet képezni:

$$D^2\left(\sum_{i=1}^n \xi_i\right) = \sum_{i=1}^n D^2(\xi_i). \quad (3.31)$$

Fontos hangsúlyozni, hogy ez a tétel csak független (pontosabban: korrelálatlan) valószínűségi változókra érvényes.

### 3.2. Nevezetes eloszlások

Ebben az alfejezetben olyan ismert eloszlások sűrűségfüggvényét, továbbá várható értékét és szórását adjuk meg, amelyek a mérések kiértékelésében fontos szerepet játszanak. Először az egydimenziós eloszlásokat tekintjük, majd áttérünk többdimenziós eloszlásokra. Az utóbbi kört leszűkítjük a többdimenziós Gauss-eloszlásra.

#### Egydimenziós eloszlások

##### Egyenletes eloszlás

A geometriai valószínűséggel kapcsolatban már utaltunk az egyenletes eloszlásra. Akkor mondjuk, hogy a  $\xi$  valószínűségi változó egyenletes eloszlású a  $[0, \theta]$  intervallumban, ha sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{1}{\theta}, & \text{ha } 0 \leq x \leq \theta, \\ 0 & \text{ha } x < 0, \text{ vagy } x > \theta. \end{cases} \quad (3.32)$$

Az egyenletes eloszlás gyakrabban fordul elő, mint gondolnánk. Mindenesetre arról nevezetes, hogy a mérések kiértékelésére *általában* kidolgozott módszerek sorra csődöt mondanak, amikor a mért adatok egyenletes eloszlásúak. Ennek oka abban rejlik, hogy a valószínűségi változónak az a tartománya, ahol a sűrűségfüggvény 0-tól különbözik, függ  $\theta$ -tól.  $\theta$  általában ismeretlen paraméter, és a mérést gyakran azért végezzük, hogy értékéről felvilágosítást kapjunk.

Tetszőleges  $\xi$  valószínűségi változót lehet egyenletes eloszlásúvá transzformálni. Érvényes ugyanis az

3.9. TÉTEL. Legyen a  $\xi$  valószínűségi változó eloszlásfüggvénye  $F(x)$ . Ekkor  $F(\xi)$  egyenletes eloszlású valószínűségi változó a  $[0, 1]$  intervallumban.

Jelöljük  $F(\xi)$  eloszlásfüggvényét  $\Phi(x)$ -szel. Ekkor definíció szerint fennáll:

$$\Phi(x) = P\{F(\xi) < x\} = P\{\xi < F^{-1}(x)\} = F[F^{-1}(x)] = x,$$

ahol a “-1” felső indexszel az inverz függvényt jelöltük. Ennek alapján  $F(\xi)$  sűrűségfüggvénye azonosan 1. Mivel  $F(\xi)$  felső határa 1, ezzel a tételt bebizonyítottuk.

A (3.12) és (3.13) képletek alapján egyszerűen kapjuk az egyenletes eloszlás várható értékét és szórásnégyzetét:

$$M(\xi) = \int_0^{\theta} x \frac{dx}{\theta} = \frac{\theta}{2} \quad (3.33a)$$

és

$$D^2(\xi) = \int_0^{\theta} \left(x - \frac{\theta}{2}\right)^2 \frac{dx}{\theta} = \frac{\theta^2}{12}. \quad (3.33b)$$

### Binomiális eloszlás

Legyen  $p$  az  $A$  esemény valószínűsége. A kísérletet  $n$ -szer megismételjük.  $k$ -val jelöljük azoknak a kísérleteknek a számát, amelyekben  $A$  bekövetkezik. A  $k$  valószínűségi változó eloszlása a *binomiális eloszlás*. Egyszerűen kiszámíthatjuk annak  $p_k$  valószínűségét, hogy  $A$  pontosan  $k$ -szor következik be, amihez az is kell, hogy  $A$  pontosan  $(n - k)$ -szor *ne* következzen be. Egy ilyen kísérletsorozat valószínűsége

$$p^k (1 - p)^{n-k}.$$

Az elvégzett  $n$  kísérlet eredménye  $k$  darab “igen”-ből és  $(n - k)$  darab “nem”-ből áll aszerint, hogy az  $A$  esemény bekövetkezett-e vagy sem. Mivel a keresett  $p_k$  valószínűség szempontjából közömbös ezek sorrendje, a fenti valószínűséget meg kell szorozni a kedvező kísérletsorozatok számával:

$$p_k = \binom{n}{k} p^k (1 - p)^{n-k}. \quad (3.34)$$

Egyszerűen beláthatjuk, hogy  $k$  várható értéke és szórásnégyzete

$$M(k) = np \quad (3.34a)$$

és

$$D^2(k) = np(1 - p). \quad (3.34b)$$

A binomiális eloszlás alapján meg tudjuk világítani a (3.1) képletben szereplő határérték jellegét. Ott említettük, hogy a valószínűség-elmélet modern megfogalmazásában ez a határérték nem több, mint a valószínűség meghatározására szolgáló “mérési utasítás”. Magát a valószínűséget ettől függetlenül definiálják, de ennek levezetése meghaladná jegyzetünk kereteit. Nos, tegyük fel, hogy egy ilyen definíció létezik. A 3.3. TÉTELben kimondott Csebisev-egyenlőtlenség (3.15a) alakja alapján ekkor írhatjuk:

$$P\left\{\left|\frac{k}{n} - p\right| > \varepsilon\right\} \leq \frac{\sigma^2}{\varepsilon^2} = \frac{D^2(k)}{n^2 \varepsilon^2} = \frac{p(1-p)}{n \varepsilon^2}.$$

Ebből következik a

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{k}{n} - p\right| > \varepsilon\right\} = 0$$

határérték. Ezt a fajta konvergenciát nevezzük *sztochasztikus konvergenciának*: ahogy a kísérletek  $n$  száma nő, egyre valószínűtlenebbek az olyan "igen-nem" sorozatok, amelyekben a relatív gyakoriság egy előírt  $\varepsilon$ -nál jobban eltér a valószínűségtől.

A binomiális eloszlásnak két közelítő alakja érdemel említést: a *Poisson-eloszlás* és a *Gauss-eloszlás*.

### Poisson-eloszlás

Rögzítjük az  $a = np$  várható értéket, miközben  $n \rightarrow \infty$ . Meg lehet mutatni, hogy ekkor  $p_k$  határértéke

$$p_k = e^{-a} \frac{a^k}{k!}. \quad (3.35)$$

Ezt nevezzük *Poisson-eloszlásnak*. Jellegzetessége, hogy várható értéke és szórásnégyzete azonos:

$$M(k) = D^2(k) = a. \quad (3.35a)$$

### Gauss-eloszlás

A binomiális eloszlás másik határértéke a *Gauss- vagy normáeloszlás*, amelynek a sűrűségfüggvénye

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}, \quad (3.36)$$

ahol

$$M(\xi) = a \quad \text{és} \quad D^2(\xi) = \sigma^2. \quad (3.36a)$$

Ez az eloszlás a mérések kiértékelésében alapvető szerepet játszik, ezért ebben a jegyzetben sokszor találkozunk vele.

### **Többdimenziós Gauss-eloszlás**

A többdimenziós eloszlások közül csak a többdimenziós Gauss-eloszlással foglalkozunk. Ha a  $\bar{\xi}$  vektor várható értéke, illetve kovarianciamátrixa

$$\mathbf{M}(\bar{\boldsymbol{\xi}}) = \mathbf{a} \quad \text{és} \quad \mathbf{B}_{\bar{\boldsymbol{\xi}}} = \mathbf{M} \left[ (\bar{\boldsymbol{\xi}} - \mathbf{a})(\bar{\boldsymbol{\xi}} - \mathbf{a})^T \right], \quad (3.37a)$$

a  $\bar{\boldsymbol{\xi}}$  vektor sűrűségfüggvénye

$$f(\mathbf{x}) = C_0 \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{a})^T \mathbf{B}_{\bar{\boldsymbol{\xi}}}^{-1} (\mathbf{x} - \mathbf{a}) \right\}, \quad (3.37)$$

ahol

$$C_0 = \frac{[\det(\mathbf{B}_{\bar{\boldsymbol{\xi}}}^{-1})]^{1/2}}{(2\pi)^{n/2}} = \frac{[\det(\mathbf{B}_{\bar{\boldsymbol{\xi}}})]^{-1/2}}{(2\pi)^{n/2}}. \quad (3.37b)$$

A sűrűségfüggvény definíciójához fel kell tételeznünk, hogy a kovarianciamátrix invertálható. E szakasz végén visszatérünk a szinguláris kovarianciamátrix esetére.

A 3.9. TÉTEL szerint független valószínűségi változók kovarianciája eltűnik, de ennek a megfordítása nem feltétlenül érvényes. Nos, Gauss-eloszlás esetében a kovariancia eltűnése függetlenséget jelent. Ezt a következőképpen láthatjuk be. Ha a kovarianciák eltűnnek, akkor a  $\mathbf{B}$  kovarianciamátrix diagonális:

$$\mathbf{B}_{\bar{\boldsymbol{\xi}}} = \text{diag}(\sigma_i^2), \quad \mathbf{B}_{\bar{\boldsymbol{\xi}}}^{-1} = \text{diag}\left(\frac{1}{\sigma_i^2}\right),$$

amit (3.37)-be helyettesítve kapjuk az

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(2\pi)^{n/2} \sqrt{\prod_{i=1}^n \sigma_i^2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \frac{(x_i - a_i)^2}{\sigma_i^2} \right\} = \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left\{ -\frac{(x_i - a_i)^2}{2\sigma_i^2} \right\} \end{aligned} \quad (3.37c)$$

sűrűségfüggvényt. Az itt szereplő tényezők az egyes valószínűségi változók külön-külön vett sűrűségfüggvényei. A most bizonyított kijelentést egy tétel formájában is kimondjuk:

3.10. TÉTEL. Gauss-eloszlású valószínűségi változók akkor és csak akkor függetlenek, ha kovarianciamátrixuk diagonális.

Befejezésül megbeszéljük, mi a helyzet akkor, amikor a kovarianciamátrix szinguláris. Ez azt jelenti, hogy van olyan nemzérus  $\mathbf{z}$  vektor, amelyre

$$\mathbf{z}^T \mathbf{B} \mathbf{z} = 0,$$

ami csak úgy lehetséges, hogy

$$\mathbf{z}^T \bar{\xi} = \text{konstans},$$

vagyis a valószínűségi változók nem lineárisan függetlenek. Ha a  $\mathbf{B}$  mátrix rangja  $k$ , az ilyen  $\mathbf{z}$  vektorok között lehet  $(n - k)$  darab lineárisan függetlent találni. Írjuk fel ezt általánosan is: létezik egy olyan  $[n \times (n - k)$  rendű]  $\mathbf{Z}_{n, n-k}$  mátrix, amelyre

$$\mathbf{Z}^T \bar{\xi} = \text{konstans},$$

és rangja  $(n - k)$ . Ez azt jelenti, hogy a  $\bar{\xi}$  vektor lehetséges értékei nem töltik ki a teljes  $n$ -dimenziós teret, hanem annak csak egy  $(n - k)$ -dimenziós alterét.

$n = 3$  és  $k = 2$  esetén, például, ez azt jelenti, hogy a  $\bar{\xi}$  vektor nem a teljes térben, hanem csak egy síkon változhat. Egyébként ez a sík a  $\mathbf{z}$  vektorra merőleges. Ha az  $f(x_1, x_2, x_3)$  sűrűségfüggvényt ennek ellenére három változó függvényének tekintjük, vagyis háromdimenziós  $dV = dx_1 dx_2 dx_3$  térfogatelemekre vonatkoztatjuk, akkor a sűrűségfüggvény azonosan 0. *Síkbeli* alakzatokra vonatkozó *térbeli* integrálok ugyanis eltűnnek.

Levonhatjuk tehát azt a következtetést, hogy a sűrűségfüggvény *nem definiálható*, amikor  $\mathbf{B}$  szinguláris. Ebben az esetben a  $\bar{\xi}$  vektor komponenseinek a számát csökkentenünk kell. A 2.3. DEFINÍCIÓ szerint  $\mathbf{B}$ -nek van egy nem-szinguláris  $k$ -adrendű almátrixa. A  $\bar{\xi}$  vektornak ehhez tartozó komponensei már lineárisan független valószínűségi változók, amelyekre értelmezhető a sűrűségfüggvény. A fennmaradó  $(n - k)$  darab valószínűségi változót az előbbiekkel ki lehet fejezni, tehát rájuk vonatkozóan bármilyen mennyiséget (várható értéket, kovarianciát, valószínűségeket stb.) ki lehet számítani. Ennek részleteibe azonban nem megyünk bele.

### 3.3. A Gauss-eloszlásból származtatott eloszlások

A mérések kiértékelésekor több olyan valószínűségi változóval van dolgunk, amelyek a Gauss-eloszlásból származtathatók. A méréskiértékelés irodalmában három eloszlás játszik különösen fontos szerepet:  $\chi^2$ -eloszlás, Student-eloszlás és Fisher-eloszlás. Egy negyedik, a  $\varphi$ -eloszlás jelenik meg a 9. fejezetben tárgyalt pontelhagyásos módszer alkalmazásakor.

#### $\chi^2$ -eloszlás

Ha  $(\xi_1, \xi_2, \dots, \xi_n)$  független Gauss-eloszlású valószínűségi változók, amelyek várható értéke 0, szórása 1, akkor definíció szerint

$$\chi_n^2 = \sum_{i=1}^n \xi_i^2. \quad (3.38)$$

Ennek a valószínűségi változónak a sűrűségfüggvénye

$$k_n(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad (3.38a)$$

ahol  $\Gamma(x)$  az ún. gammafüggvény:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt. \quad (3.39)$$

A (3.38)-ban megjelenő tagok  $n$  számát a  $\chi^2$ -eloszlás *szabadsági fokának* nevezzük.  $\chi^2$  várható értéke és szórásnégyzete:

$$M(\chi^2) = n \quad \text{és} \quad D^2(\chi^2) = 2n. \quad (3.38b)$$

### Student-eloszlás

Legyen  $\xi$  egy 0 várható értékű és 1 szórású, Gauss-eloszlású valószínűségi változó, amely független  $\chi_n^2$ -től. Ekkor a

$$t_n = \frac{\xi}{\sqrt{\frac{\chi_n^2}{n}}} \quad (3.40)$$

hányadost *Student-törtnek* nevezzük. Sűrűségfüggvénye

$$s_n(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad (3.40a)$$

várható értéke és szórása

$$M(t_n) = 0 \quad \text{és} \quad D^2(t_n) = \frac{n}{n-2}, \quad n > 2. \quad (3.40b)$$

A (3.40a) sűrűségfüggvény ugyan értelmezhető  $n = 1$ -re és  $n = 2$ -re is, de ezekben az esetekben nem létezik a szórás.  $n$ -et a Student-eloszlás esetében is a szabadsági fokok számának nevezzük.

### Fisher-eloszlás

A Fisher-eloszlás két, egymástól független  $\chi^2$ -változó hányadosa:

$$F_{mn} = \frac{\chi_m^2/m}{\chi_n^2/n}. \quad (3.41)$$

Sűrűségfüggvénye

$$f_{mn}(x) = \frac{n}{m} \frac{\Gamma\left(\frac{m+n}{2}\right) \left(\frac{nx}{m}\right)^{\frac{m}{2}-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right) \left(\frac{nx}{m}+1\right)^{\frac{m+n}{2}}}, \quad (3.41a)$$

várható értéke és szórásnégyzete

$$M(F_{mn}) = 1 \quad \text{és} \quad D^2(F_{mn}) = \frac{m+2}{m} \cdot \frac{n^2}{(n-2)(n-4)} - 1. \quad (3.41b)$$

### $\varphi$ -eloszlás

A  $\varphi$ -hányados emlékeztet a Fisher-hányadosra:

$$\varphi_{mn} = \frac{\sqrt{\chi_m^2/m}}{\chi_n^2/n}, \quad (3.42)$$

ahol  $\chi_m^2$  és  $\chi_n^2$  egymástól független  $\chi^2$ -változók. A Fisher-hányadoshoz képest az egyetlen különbség a számlálóban szereplő négyzetgyök. Bonyolultsága miatt nem írjuk fel a sűrűségfüggvényét, sem várható értékét és szórását.

### \*3.4. Korrelációs ellipszoid

A *korrelációs ellipszoid* hasznos segédeszköz véletlen vektorok összehasonlításában.

3.11. DEFINÍCIÓ. Ha  $M(\vec{\xi}) = \mathbf{a}$ , a  $\vec{\xi}$  véletlen vektor korrelációs ellipszoidját azoknak az  $\mathbf{x}$  vektoroknak a végpontjai alkotják, amelyek kielégítik az

$$(\mathbf{x} - \mathbf{a})^T \mathbf{B}_{\vec{\xi}}^{-1} (\mathbf{x} - \mathbf{a}) = 1 \quad (3.43)$$

egyenletet.

A korrelációs ellipszoid nevezetessége, hogy felületén a sűrűségfüggvény állandó. Ez közvetlenül belátható a többdimenziós Gauss-eloszlás (3.37) szerinti definíciója alapján.

Legyen  $\vec{\Omega}$  egységvektor, és számítsuk ki az  $\vec{\Omega}^T \vec{\xi}$  valószínűségi változó szórásnégyzetét:

$$r^2 = D^2(\vec{\Omega}^T \vec{\xi}) = M\left(\vec{\Omega}^T (\vec{\xi} - \mathbf{a}) (\vec{\xi} - \mathbf{a})^T \vec{\Omega}\right) = \vec{\Omega}^T \mathbf{B}_{\vec{\xi}} \vec{\Omega}. \quad (3.44)$$

Megmutatjuk, hogy a korrelációs ellipszoid tetszőleges  $\mathbf{x}$  pontjához lehet találni olyan  $\vec{\Omega}$  vektort, hogy  $|\mathbf{x} - \mathbf{a}| = r$ . Megfontolásainkat a 3.4. ábrán lehet nyo-



mon követni, amely két változó ( $n = 2$ ) esetében mutatja az ellipszoidot (az ábrán ellipszist). Legyen  $\mathbf{U}$  az az unitér mátrix, amely a  $\mathbf{B}_{\bar{\xi}}$  mátrixot főtengetyre transzformálja:

$$\text{diag}(\sigma^2) = \mathbf{B}_{\bar{\eta}} = \mathbf{U}^T \mathbf{B}_{\bar{\xi}} \mathbf{U}, \quad \mathbf{B}_{\bar{\xi}} = \mathbf{U} \mathbf{B}_{\bar{\eta}} \mathbf{U}^T. \quad (3.45)$$

A  $\mathbf{B}_{\bar{\eta}}$  diagonális mátrix az  $\bar{\boldsymbol{\eta}} = \mathbf{U}^T(\bar{\boldsymbol{\xi}} - \mathbf{a})$  vektor kovarianciamátrixa. A főátlóban levő

$$\sigma_i^2 = D^2(\eta_i), \quad i = 1, 2, \dots, n$$

elemek a  $\mathbf{B}_{\bar{\xi}}$  mátrix sajátértékei. A transzformált koordináta-rendszerben tehát a (3.43) szerinti felület egyenlete

$$\mathbf{y}^T \mathbf{B}_{\bar{\eta}}^{-1} \mathbf{y} = \sum_{i=1}^n \frac{y_i^2}{\sigma_i^2} = 1, \quad (3.46)$$

ahol  $\mathbf{y}$  az eredeti  $\mathbf{x}$  vektor transzformáltja:  $\mathbf{y} = \mathbf{U}^T(\mathbf{x} - \mathbf{a})$ .

Az  $\mathbf{U}$  mátrix oszlopai  $\mathbf{B}_{\bar{\xi}}$  sajátvektorai, tehát  $\mathbf{y}$  komponensei az  $(\mathbf{x} - \mathbf{a})$  vektornak a sajátvektorokra vett vetületei. A 3.4 ábrán  $\theta$ -val jelöljük az  $\mathbf{u}_1$  sajátvektornak az eredeti  $x_1$  koordinátatengellyel bezárt szögét. Ezzel a jelöléssel az  $\mathbf{U}$  mátrix elemei a következők:

$$\mathbf{U} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

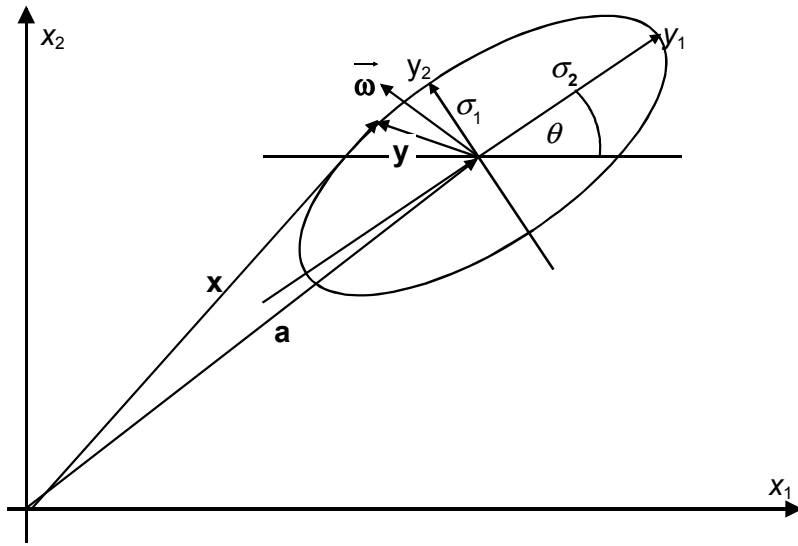
Ha az  $(\mathbf{x} - \mathbf{a})$  vektor polárszöge  $\varphi$  (az ábrán nincs jelölve), akkor

$$\mathbf{x} - \mathbf{a} = r \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix},$$

amivel

$$\mathbf{y} = \mathbf{U}^T(\mathbf{x} - \mathbf{a}) = r \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} = r \begin{bmatrix} \cos(\varphi - \theta) \\ \sin(\varphi - \theta) \end{bmatrix}.$$

Az  $\mathbf{u}_1$  és  $\mathbf{u}_2$  vektorok által kifeszített koordináta-rendszer az eredetihez képest  $\theta$  szöggel van elforgatva az óramutató járásával ellentétes irányban. Utóbbi képletünkből látszik, hogy  $(\mathbf{x} - \mathbf{a})$  és  $\mathbf{y}$  ugyanazt a vektort jelöli, de komponenseik egymáshoz képest elforgatott koordináta-rendszerekben vannak kifejezve. Ugyanez az állítás érvényes az  $n$ -dimenziós térben is.



3.4. ábra. Korrelációs ellipszis

A (3.46) egyenlet szerint az  $y_i/\sigma_i$  hányadosok egy egységvektor komponensei. Jelöljük ezt  $\vec{\omega}$ -val:  $\omega_i = y_i/\sigma_i$ ,  $i = 1, 2, \dots, n$ . Az  $\mathbf{y}$  vektor hosszúságának

$$\mathbf{y}^T \mathbf{y} = \sum_{i=1}^n \sigma_i^2 \omega_i^2 = \vec{\omega}^T \mathbf{B}_{\vec{\eta}} \vec{\omega}$$

négyzete – (3.44) analógiájára – az  $\vec{\omega}^T \vec{\eta}$  valószínűségi változó szórásnégyzete. Legyen  $\vec{\Omega} = \mathbf{U} \vec{\omega}$ , illetve  $\vec{\omega} = \mathbf{U}^T \vec{\Omega}$ . A fentiek analógiájára  $\vec{\Omega}$  és  $\vec{\omega}$  ugyanazt a vektort jelölik, de komponenseik egymáshoz képest elfogartott koordinátarendszerekben vannak kifejezve. A korrelációs ellipszoid  $\mathbf{x}$  pontjához rendeljük az  $\vec{\Omega}$  egységvektort. (3.45) egyenlet alapján

$$r^2 = (\mathbf{x} - \mathbf{a})^T (\mathbf{x} - \mathbf{a}) = \mathbf{y}^T \mathbf{y} = \vec{\omega}^T \mathbf{B}_{\vec{\eta}} \vec{\omega} = \vec{\Omega}^T \mathbf{U} \mathbf{B}_{\vec{\eta}} \mathbf{U}^T \vec{\Omega} = \vec{\Omega}^T \mathbf{B}_{\vec{\xi}} \vec{\Omega}.$$

Ezzel igazoltuk fenti állításunkat: az ellipszoid  $\mathbf{x}$  pontjához vezető sugár  $r$  hossza megegyezik az  $\vec{\Omega}^T \vec{\xi}$  valószínűségi változó szórásával.

$\vec{\omega}$  (és így  $\vec{\Omega}$ ) iránya – mint a 3.4. ábrán látszik –  $\mathbf{y}$ -hoz képest el van forgatva, ha a  $\sigma_i$  sajátértékek nem mind azonosak. Kivételek a sajátvektorokkal párhuzamos  $\mathbf{y}$  vektorok, mert esetükben  $\vec{\omega} \parallel \mathbf{y}$ . Most bizonyított állításunk alapján megérthetjük, mit jelentenek a  $\mathbf{B}_{\vec{\xi}}$  kovarianciamátrix sajátértékei:  $\sigma_i^2$  az  $\mathbf{u}_i^T \vec{\xi}$  valószínűségi változó szórásnégyzete.

Tekintünk két azonos dimenziójú véletlen vektort, mondjuk  $\vec{\alpha}$ -t és  $\vec{\beta}$ -t, amelyek várható értéke azonos, és megszerkesztjük korrelációs ellipszoidjukat. Ha azt találjuk, hogy  $\vec{\alpha}$ -é teljes egészében benne van  $\vec{\beta}$ -ében, akkor ez azt je-

lenti, hogy tetszőleges  $\bar{\Omega}$  egységvektorra vonatkozóan  $\bar{\Omega}^T \bar{\alpha}$  szórása kisebb, mint  $\bar{\Omega}^T \bar{\beta}$ -é. Köznyelven szólva ezt úgy is kifejezhetjük, hogy  $\bar{\alpha}$  mérése minden tekintetben pontosabb, mint  $\bar{\beta}$ -é. Amikor a korrelációs ellipszoidok átmetszik egymást, különböző  $\bar{\Omega}$ -kra  $\bar{\Omega}^T \bar{\alpha}$  és  $\bar{\Omega}^T \bar{\beta}$  szórása közötti relációk egymástól eltérőek lehetnek. A korrelációs ellipszoid további alkalmazását láthatjuk az 4.4. alfejezetben.