A Valószínűségszámítás II. előadássorozat ötödik témája.

A TÖBBVÁLTOZÓS CENTRÁLIS HATÁRELOSZLÁSTÉTEL

A centrális határeloszlás sok független valószínűségi változó összegének az aszimptotikus eloszlását írja le. Bizonyos kérdések vizsgálatában szükségünk van ennek az eredménynek egy olyan általánosabb változatára, amely független, vektor értékű valószínűségi változók összegének az aszimptotikus eloszlását adja meg. A centrális határeloszlástételnek létezik ilyen általánosítása, és ezt hívják többváltozós centrális határeloszlástételnek. Következő témánk ennek az eredménynek a tárgyalása.

A többváltozós centrális határeloszlástétel megértése érdekében először meg kell ismernünk a több-dimenziós normális eloszlás definicióját és annak néhány fontos tulajdonságát. Ugyanis a többváltozós centrális határeloszlástételekben a több-dimenziós normális eloszlások jelennek meg, mint határeloszlások. Annak érdekében, hogy ezeket az eloszlásokat jól megértsük, be kell vezetni a valószínűségi változók várható értékének és szórásnégyzetének természetes megfelelőjét vektor értékű valószínűségi változók esetére. Ez a a vektor értékű valószínűségi változók várható értékének és kovariancia mátrixának a bevezetését jelenti. Ezenkívül fel kell eleveníteni a lineáris algebra néhány eredményének az ismeretét.

Mielőtt a többváltozós centrális határeloszlástétel tárgyalását elkezdenénk, lássunk két olyan problémát, amelyek vizsgálatában ez az eredmény hasznosnak bizonyult.

a.) Tekintsünk egy dobókockát. Feldobjuk sokszor, felírjuk a dobások eredményét, és ennek alapján akarjuk eldönteni, hogy a dobókocka szabályos-e. Természetes azt várni, hogy a dobókocka akkor szabályos, ha mindegyik dobáseredmény előfordulásának a száma a dobásszámok egyhatoda plusz egy kis eltérés. De mekkora eltéréseket tekinthetünk kicsinek? Ha csak azt nézzük, hogy mennyi annak a valószínűsége, hogy például a hatos dobások számának eltérése a dobások számának egyhatodától kisebb mint egy adott szám, akkor a centrális határeloszlástétel pontos leírást ad erre a problémára. De ha a különböző dobáseredmények együttes viselkedésére vagyunk kiváncsiak, akkor új eredményre van szükségünk.

b.) Legyenek $\xi_1,\dots,\xi_n$ független, a $\left[-\frac12,\frac12\right]$ intervallumban egyenletes eloszlású valószínűségi változók. Ekkor a $\sum\limits_{j=1}^n \xi_j$ összeg normalizáltjának az eloszlására jó leírást ad a centrális határeloszlástétel. Hasonló állítást mondhatunk a $\sum\limits_{j=1}^n \xi_j^2$ összeg normalizáltjának az eloszlására. De tudunk-e hasonló eredményt adni a $\sum\limits_{j=1}^n \xi_j$ és $\sum\limits_{j=1}^n \xi_j^2$ összegek normalizáltjának az együttes eloszlására?

Annak érdekében, hogy lássuk, hogyan lehet ezeket a kérdéseket vizsgálni a centrális határeloszlástétel alkalmas több-dimenziós megfelelőjének, azaz a többváltozós centrális határeloszlástételnek a segítségével bevezetek be néhány jelölést.

Legyenek $\xi^{(j)}=\left(\xi_1^{(j)},\dots,\xi_k^{(j)}\right)$, $j=1,2,\dots$, független, egyforma eloszlású k-dimenziós valószínűségi változók, (véletlen vektorok), ahol rögzített j indexre semmilyen (függetlenség jellegű) feltételt nem teszünk fel az $\xi_1^{(j)}, \dots, \xi_k^{(j)}$ valószínűségi változók együttes eloszlására. Tegyük fel továbbá, hogy $E{\xi_s^{(j)}}^2 < \infty$ minden $1 \le s \le k$ indexre. Tekintsük az $S_n = (S_{n,1}, \dots, S_{n,k}) = \sum\limits_{j=1}^n \xi^{(j)} = \left(\sum\limits_{j=1}^n \xi_1^{(j)}, \dots, \sum\limits_{j=1}^n \xi_k^{(j)} \right)$, $n = 1, 2, \dots$, véletlen összegeket. Be akarjuk látni, hogy az S_n véletlen vektorok alkalmas normalizáltjának létezik határeloszlása, és a határeloszlást pontosan le akarjuk írni. Látni fogjuk, hogy ez lehetséges. A határeloszlástételben megjelenő határeloszlásokat fogjuk több-dimenziós normális eloszlásoknak nevezni.

A teljesség kedvéért fölidézem, hogy itt és a továbbiakban vektor értékű valószínűségi változók függetlenségének alábbi, a bevezető valószínűségszámítás előadásban bevezetett definicióját használjuk.

Vektorértékű valószínűségi változók függetlenségének a definiciója. *Legyenek* $\xi^{(1)} = \left(\xi_1^{(1)}, \dots, \xi_k^{(1)}\right), \dots, \xi^{(n)} = \left(\xi_1^{(n)}, \dots, \xi_k^{(n)}\right)$, *értékeiket az R^k k-dimenziós Euklideszi térben felvevő valószínűségi változók (vektorok) egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi vektorok függetlenek, ha minden* $x^{(1)} = (x_1^{(1)}, \dots, x_k^{(1)}), \dots, x^{(n)} = (x_1^{(n)}, \dots, x_k^{(n)})$ *k-dimenziós vektorra*

$$\begin{aligned} P&\left(\xi_1^{(1)} < x_1^{(1)}, \dots, \xi_k^{(1)} < x_k^{(1)}, \dots, \xi_1^{(n)} < x_1^{(n)}, \dots, \xi_k^{(n)} < x_k^{(n)}\right) \\ &= P\left(\xi_1^{(1)} < x_1^{(1)}, \dots, \xi_k^{(1)} < x_k^{(1)}\right) \cdots P\left(\xi_1^{(n)} < x_1^{(n)}, \dots, \xi_k^{(n)} < x_k^{(n)}\right). \end{aligned}$$

Lássuk, hogyan lehet tárgyalni az előbb megfogalmazott két problémát egy független, vektor értékű valószínűségi változók normalizált összegeinek határeloszlását leíró eredmény segítségével. Az a) feladat vizsgálatának érdekében vezessük be a következő vektor értékű $\xi^{(j)}$ valószínűségi változókat: Ha a dobókockát n alkalommal dobjuk fel, akkor legyen $\xi^{(j)}$, $1 \le j \le n$, a következő 6-dimenziós valószínűségi változó: A $\xi^{(j)}$ valószínűségi változó l-ik koordinátája 1, ha a j-ik dobás értéke l, $1 \le l \le 6$, és legyen a $\xi^{(j)}$ véletlen vektor összes többi koordinátája nulla. Ekkor a $\xi^{(j)}$ valószínűségi változók függetlenek. (Az egyes $\xi^{(j)}$ valószínűségi változók koordinátái ebben a példában nem függetlenek.) Továbbá az $S_n = \sum\limits_{j=1}^n \xi^{(j)}$ összeg l-ik koordinátája egyenlő az l értékű dobások számával minden $1 \le l \le 6$ index esetén. Ezért egy az S_n véletlen összeg aszimptotikus viselkedését nagy n számokra leíró határeloszlástétel hasznos lehet a számunkra. Egy ilyen eredményből következik, hogy az S_n véletlen összegek alkalmas normalizáltjai eloszlásban konvergálnak egy explicit módon megadott eloszlásfüggvényhez. Megjegyzem, hogy a többváltozós centrális határeloszlástétel tárgyalása érdekében az eloszlásban való konvergenciának egy a többdimenziós valószínűségi változók esetében is érvényes definicóját vezettük be.

Bár a többváltozós határeloszlás fogalmát tanultuk, mégis érdemes néhány e fogalommal kapcsolatos eddig nem tárgyalt kérdést megvizsgálni. Természetes az a) feladat megoldásában a véletlen S_n vektorból az $(\frac{n}{6}, \dots, \frac{n}{6})$ vektort kivonni, és ennek a 6-változós különbségvektornak a hosszát tekinteni (az Euklideszi távolságfogalom szerint). Ha a dobókocka szabályos volt, akkor azt várjuk, hogy e véletlen vektor hossza viszonylag kicsi. Felmerül a kérdés, hogy van-e e véletlen vektor alkalmas normalizáltjának határeloszlása, és következik-e egy ilyen eredmény a több-dimenziós centrális határeloszlástételből. Be fogjuk látni, hogy erre a kérdésre igenlő a válasz. De ez az állítás indoklásra szorul. Ugyanis a több-dimenziós eloszlások konvergenciájanak a definiciója annak eredeti alakjában csak nagyon speciális halmazok valószínűségének a konvergenciáját írja elő.

A b) példában megfogalmazott kérdést hasonlóan tárgyalhatjuk az a) példához. Itt olyan $\eta_j = (\xi_j, \xi_j^2)$, $1 \le j \le n$, két-dimenziós független vektorok $S_n = \sum\limits_{j=1}^{n} \eta_j$ összegét vizsgáljuk, amelyekre $\xi_1, \dots, \xi_n$ független, a $\left[-\frac{1}{2}, \frac{1}{2}\right]$ intervallumban egyenletes eloszlású valószínűségi változók. A többváltozós centrális határeloszlástétel segítségével leírható az ilyen véletlen vektorok normalizáltjainak a határeloszlása.

Ismertetni fogom a centrális határeloszlástétel természetes megfelelőjét abban az esetben, ha *független és egyforma eloszlású* véletlen vektorok alkalmasan normalizált összegének a viselkedését vizsgáljuk. Ennek érdekében bevezetem a határeloszlásban megjelenő normális eloszlás több-dimenziós megfelelőjét, amit több-dimenziós normális eloszlásnak fogok nevezni. De ehhez előbb definiálni kell a várható érték és szórásnégyzet fogalmának a több-dimenziós megfelelőjét. Ezenkívül meg kell érteni, hogy a várható érték és szórásnégyzet tulajdonságai hogyan öröklődnek a több-dimenziós esetben. Az egy-dimenziós esetben definiált várható értéknek megfelelő több-dimenziós várható érték (vektor) fogalmának és tulajdonságainak a megértése viszonylag egyszerű, de a szórásnégyzetnek megfelelő kovariancia mátrix tulajdonságainak jó megértése szükségessé teszi néhány a lineáris algebrában tanult fogalom és eredmény felelevenítését. Megjegyzem, hogy most és a továbbiakban is a több-dimenziós vektorokat mint sorvektorokat tekintem. Valójában mind a sor mind az oszlopvektor jelölés lehetséges. Egyik jelölés sem jobb a másiknál, és az irodalomban nem egységes a jelölés. A fontos az, hogy következetes jelölést alkalmazzunk.

Több-dimenziós valószínűségi változó várható értékének és kovariancia mátrixának a definiciója. *Legyen $Z = (Z_1, \dots, Z_k)$ k-dimenziós véletlen vektor, amelynek minden koordinátája teljesíti az $EZ_j^2 < \infty$, $1 \le j \le k$, feltételt. E véletlen vektor várható értéke az $EZ = (EZ_1, \dots, EZ_k)$ k-dimenziós vektor, kovariancia mátrixa pedig az a $D = (d_{j,l})$, $1 \le j, l \le k$, $k \times k$ méretű mátrix, mely mátrix j-ik sorában és l-ik oszlopában lévő elem a $d_{j,l} = \operatorname{Cov}(Z_j, Z_l) = EZ_jZ_l - EZ_jEZ_l$ szám.*

Megfogalmazom a vektorértékű valószínűségi változók várható értékének és kovariancia mátrixának néhány fontos tulajdonságát. Ezek egyszerű következményei a valós szám értékű valószínűségi változók már tárgyalt tulajdonságainak, ezért bizonyításukat elhagyom.

Tétel vektor értékű valószínűségi változók várható értékének és kovariancia mátrixának tulajdonságairól. *Legyenek $Z^{(j)} = \left(Z_1^{(j)}, \dots, Z_k^{(j)}\right)$, $1 \le j \le n$, véletlen k-dimenziós vektorok ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ekkor a $Z^{(1)} + \cdots + Z^{(n)}$ összeg várható értéke megegyezik a $Z^{(j)}$ vektorok várható értékeinek az összegével,* *azaz*

$$E(Z^{(1)} + \cdots + Z^{(n)}) = EZ^{(1)} + \cdots + EZ^{(n)}.$$

Ha a $Z^{(j)}$, $1 \le j \le n$, véletlen vektorok függetlenek, akkor a kovariancia mátrix is additív, azaz, ha a $Z^{(j)}$ mátrix kovariancia mátrixa a D_j mátrix, $1 \le j \le n$, akkor a $Z^{(1)}+\cdots+Z^{(n)}$ véletlen összeg kovariancia mátrixa a $D_1+\cdots+D_n$ mátrix. Ha egy $Z = (Z_1, \dots, Z_k)$, véletlen vektor várható értéke $M = (M_1, \dots, M_k)$, kovariancia mátrixa a D $k \times k$ méretű mátrix, a tetszőleges valós szám, akkor az $aZ = (aZ_1, \dots, aZ_k)$, véletlen vektor várható értéke aM, kovariancia mátrixa pedig az a^2D kovariancia mátrix. Legyen továbbá $x = (x_1, \dots, x_k)$ tetszőleges k-dimenziós vektor. Ekkor $E(Z + x) = EZ + x$, a $Z + x$ vektor kovariancia mátrixa pedig megegyezik a Z vektor kovariancia mátrixával.

A következő eredmény célja annak jellemzése, hogy milyen mátrix jelenhet meg, mint egy alkalmas véletlen vektor kovariancia mátrixa. Ennek az eredménynek az ismertetésében és bizonyításában fel kell használnunk a lineáris algebra néhány alapvető fogalmát és eredményét. Igyekszem az állításokat önmagukban is érthető módon leírni. Először felidézem a következő lineáris algebrai fogalmat.

Szimmetrikus és pozitív (szemi)definit mátrixok definiciója. *Legyen $D = (d_{j,l})$ egy $k \times k$ méretű mátrix. Azt mondjuk, hogy a D mátrix szimmetrikus, ha minden $1 \le j, l \le k$ indexre $d_{j,l} = d_{l,j}$. Pontosabban azt követeljük meg, (ha nemcsak valós, hanem általános komplex értékű elemekkel rendelkező mátrixokat is tekintünk, hogy $d_{j,l} = \bar{d}_{l,j}$, ahol $\bar{z}$ a z komplex szám konjugáltja, azaz, ha $z = a + ib$, akkor $\bar{z} = a - ib$. (De ebben az előadásban csak valós elemű mátrixokkal fogunk dolgozni.) Egy $k \times k$ méretű szimmetrikus $D = (d_{j,l})$ mátrix pozitív (szemi)definit, ha minden $x = (x_1, \dots, x_k)$ k-dimenziós vektorra $xDx^* = \sum\limits_{j=1}^{k} \sum\limits_{l=1}^{k} x_j d_{j,l} x_l \ge 0$. (Ebben a formulában x^* jelöli az x vektor transzponáltját, azaz azt az oszlopvektort, amelynek fölülről számíva l-ik eleme megegyezik az x vektor balról számított l-ik elemének a komplex konjugáltjával. Ekkor xDx^* a szokásos vektor-mátrix szorzást jelöli.*

A $D = (d_{i,j})$ szimmetrikus mátrixot (szigorúan) pozitív definitnek nevezünk, ha pozitív szemidefinit, és ráadásul az $xDx^* = \sum\limits_{j=1}^{k} \sum\limits_{l=1}^{k} x_j d_{j,l} x_l = 0$ reláció csak abban a triviális esetben teljesül, ha $x_1 = x_2 = \cdots = x_k = 0$.

Az eredmények és fogalmak jobb megértése érdekében érdemes megadni a fenti koordinátarendszerfüggő definiciók koordinátarendszertől független „absztrakt” definicióját is, és megérteni a két definició kapcsolatát. Ennek leírását (a bizonyítások többségének elhagyásával) tartalmazza egy a múlt félévben kiosztott lineáris algebrai összefoglaló. Röviden megfogalmazom a legfontosabb fogalmakat és eredményeket.

A mátrixok úgy jelennek meg, mint lineáris transzformációk megadásai egy lineárisan független vektorokból álló koordinátarendszerben. Ha Euklideszi terekben, tehát olyan lineáris terekben dolgozunk, ahol van skalárszorzat, és ezért beszélhetünk egy vektor hosszáról és két vektor által bezárt szögről, akkor a lineáris transzformációkat egy ortonormált bázisban adjuk meg. Természetes az előbb bevezetett fogalmakat először lineáris transzformációkra definiálni, (ezt nevezem koordinátamentes definiciónak), és utána megvizsgálni azt, hogy ezek a fogalmak mit jelentenek a transzformációt megadó mátrixok nyelvén.

Az első definiálandó fogalom egy Euklideszi térben definiált A transzformáció adjungáltja. Egy valamely Euklideszi térben megadott A lineáris transzformáció adjungáltja az az A^* lineáris transzformáció, amelyre teljesül az $(xA, y) = (x, yA^*)$ azonosság az Euklideszi tér minden x és y vektorára, ahol $(\cdot, \cdot)$ skalárszorzatot jelöl. Be lehet látni, hogy minden A lineáris transzformáció esetén egy és csak egy olyan A^* lineáris transzformáció van, amely teljesíti ezt a feltételt. Ez azt jelenti, hogy a lineáris transzformáció, illetve a neki megfelelő mátrix adjungáltjának a definiciója értelmes. Megjegyzem, hogy egy mátrix adjungáltját csak Euklideszi (tehát nem tetszőleges lineáris) térben definiáltuk, mert e fogalom definiciójában felhasználtuk a skalárszorzat fogalmát. Ezután az önadjungált transzformáció fogalmának a megadása egyszerű. Egy A lineáris transzformáció (vagy a neki megfelelő mátrix) akkor és csak akkor önadjungált, ha $A = A^*$. Egy A önadjungált mátrixot akkor nevezünk pozitív szemidefinitnek, ha $(x, xA) \ge 0$ ez Euklideszi tér minden x vektorára, és akkor nevezünk (szigorúan) pozitív definitnek, ha egyrészt teljesül a fenti egyenlőtlenség, másrészt az $(x, xA) = 0$ azonosság csak a triviális $x = 0$ esetben áll fenn.

Be lehet látni, hogy ha egy A transzformáció mátrixát felírjuk egy *tetszőleges* ortonormált bázisban, akkor a transzformáció adjungáltjának a mátrixát a *Szimmetrikus és pozitív (szemi)definit mátrixok definiciójában* megadott módon számíthatjuk ki. Fontos megérteni, hogy ez az állítás azt is jelenti, hogy ha egy A lineáris transzformáció mátrixának az adjungáltját a lineáris transzformáció mátrixának az adjungáltja segítségével számoljuk ki, akkor az így kapott A^* adjungált transzformáció értéke nem függ attól, hogy melyik ortonormált bázisban dolgozunk. Azt, hogy egy önadjungált A lineáris transzormáció pozitív szemidefinit a transzformáció (egy ortonormált bázisban felírt) A mátrixával úgy jellemezhető, hogy $xAx^* \ge 0$, ahol x tetszőleges szám-n-es, és x^* annak transzponáltja. Felírom egy transzformáció adjungáltjának legfontosabb tulajdonságait: $(A + B)^* = A^* + B^*$, $(cA)^* = \bar{c}A^*$, $(A^*)^* = A$, $(AB)^* = B^*A^*$.

A lineáris transzformációk és mátrixok legfontosabb tulajdonságainak felsorolása után megfogalmazok egy fontos eredményt, amely megadja a kovariancia mátrixok jellemzését. Ennek bizonyításában felhasználok egy nem triviális lineáris algebrai eredményt is.

Tétel kovariancia mátrixok jellemzéséről. *Legyen $Z = (Z_1, \dots, Z_k)$ egy k-dimenziós véletlen vektor. Ekkor a Z vektor kovariancia mátrixa szimmetrikus és pozitív szemidefinit mátrix. Megfordítva, tetszőleges D szimmetrikus, pozitív szemidefinit mátrixhoz létezik olyan $Z = (Z_1, \dots, Z_k)$ véletlen vektor, amelynek ez a D mátrix a kovariancia mátrixa. Sőt igaz a következő tartalmasabb állítás is: Legyen $Y = (Y_1, \dots, Y_k)$ olyan véletlen vektor, amelynek a kovariancia mátrixa az identitás mátrix, azaz $\operatorname{Var} Y_j = 1$, $1 \le j \le k$, $\operatorname{Cov}(Y_j, Y_l) = 0$, ha $1 \le j, l \le k$, és $j \ne l$. (Ez a helyzet például akkor, ha az Y_j, $1 \le j \le k$ valószínűségi változók függetlenek, és $\operatorname{Var} Y_j = 1$.) Ekkor létezik olyan $A = (a_{j,l})$ $k \times k$ méretű mátrix, amelyre igaz, hogy a $Z = (Z_1, \dots, Z_k) = (Y_1, \dots, Y_k)A = \left(\sum\limits_{p=1}^{k} a_{1,p} Y_p, \dots, \sum\limits_{p=1}^{k} a_{k,p} Y_p \right)$ véletlen vektor kovariancia mátrixa a D mátrix.*

Igaz továbbá a következő állítás is. Egy $Z = (Z_1, \dots, Z_k)$ véletlen vektor kovariancia mátrixa akkor és csak akkor (szigorúan) pozitív definit, ha e vektor koordinátái között nincs lineáris összefüggés, azaz ha $x_1, \dots, x_k$ valós számokra $\sum\limits_{j=1}^{k} x_j Z_j = K$ valamilyen K (determinisztikus) valós számra egy valószínűséggel, akkor $x_1 = \cdots = x_k = 0$.

Megjegyzés: Ez az állítás annak a valószínűségi változókról szóló egyszerű eredménynek több-dimenziós megfelelője, amely szerint egy valószínűségi változó szórásnégyzete nem negatív szám. Továbbá egy valószínűségi változó szórásnégyzete akkor és csak akkor szigorúan pozitív, ha ez a valószínűségi változó nem egyenlő egy konstanssal egy valószínűséggel. Ennek a ténynek a megfelelője a tétel végén kimondott állítás, amely szerint egy véletlen vektor kovariancia mátrixa pozitív definit, ha nincs a véletlen vektor koordinátáinak olyan lineáris kombinációja, amelyik egy valószínűséggel megegyezik egy számmal. Ugyanis a nem negatív számoknak a pozitív szemidefinit mátrixok, a pozitív számoknak pedig a pozitív definit mátrixok a természetes megfelelői az Euklideszi terekben. A tétel bizonyítása egy alább megfogalmazandó nem triviális lineáris algebrai eredményen alapul, amely szerint minden pozitív szemidefinit mátrix felírható, mint egy alkalmas mátrix négyzete. Ez az állítás annnak a ténynek a megfelelője Euklideszi terekben, amely szerint pozitív számokból lehet négyzetgyököt vonni. Megjegyzem, hogy a pozitív szemidefinit mátrixokból vont négyzetgyök nem egyértelműen meghatározott, mint ahogy a valós számok között is csak akkor egyértelmű a gyökvonás, ha csak a pozitív gyököt tekintjük.

Tétel a lineáris algebrából. *Legyen D pozitív szemidefinit mátrix. Ekkor létezik olyan A mátrix, amelyre érvényes a $D = A^*A$ azonosság, ahol A^* az A mátrix transzponáltját jelöli. Sőt, olyan A mátrixot is választhatunk, amelyre az A mátrix önadjungált, pozitív szemidefinit, és $D = A^2$. Eme megszorítás esetén az $D = A^*A = A^2$ egyenlet megoldása egyértelmű. (Egy $A = (a_{j,l})$ $k \times k$ méretű mátrix transzponáltja az $A^* = (a_{l,j})$, illetve az általános komplex számokat is tartalmazó mátrixok esetében az $A^* = (\bar{a}_{l,j})$ $k \times k$ méretű mátrix, ahol $\bar{z}$ a z komplex szám konjugáltja.)*

A tétel bizonyítása a lineáris algebráról kimondott tétel segítségével. Tekintsünk először egy $Z = (Z_1, \dots, Z_k)$ egy k-dimenziós véletlen vektort és annak $D = (d_{j,l})$, $d_{j,l} = \operatorname{Cov}(Z_j, Z_l)$, $1 \le j, l \le k$, kovariancia mátrixát. Ekkor D szimmetrikus mátrix, mert $d_{j,l} = d_{l,j}$, azaz $\operatorname{Cov}(Z_j, Z_l) = \operatorname{Cov}(Z_l, Z_j)$. Másrészt tetszőleges $x = (x_1, \dots, x_k)$ k-dimenziós vektorra

$$\begin{aligned}\operatorname{Var}\left(\sum_{j=1}^{k} x_j Z_j\right) &= \sum_{j=1}^{k}\sum_{l=1}^{k} E\left(x_j x_l (Z_j Z_l - EZ_j EZ_l)\right) = \sum_{j=1}^{k}\sum_{l=1}^{k} x_j x_l \operatorname{Cov}(Z_j, Z_l) \\ &= \sum_{j=1}^{k}\sum_{l=1}^{k} x_j x_l d_{j,l} = xDx^*,\end{aligned}$$

és $\operatorname{Var}\left(\sum\limits_{j=1}^{k} x_j Z_j\right) \ge 0$. Innen következik, hogy $xDx^* \ge 0$ tetszőleges $x = (x_1, \dots, x_k)$ k-dimenziós vektorra, azaz D szimmetrikus, pozitív szemidefinit mátrix.

Megfordítva, legyen D pozitív szemidefinit mátrix, és $Y = (Y_1, \dots, Y_k)$ olyan véletlen vektor, amelynek a kovariancia mátrixa az identitás mátrix, azaz $\mathrm{Var}\, Y_j = 1$, $1 \le j \le k$, $\mathrm{Cov}\,(Y_j, Y_l) = 0$, ha $1 \le j, l \le k$, és $j \ne l$. A kimondott lineáris algebrai eredmény szerint létezik olyan $A = (a_{j,l})$, $1 \le j, l \le k$ $k \times k$ méretű mátrix, amelyre $D = A^*A$. Azt állítom, hogy a $Z = (Z_1, \dots, Z_k) = YA$, azaz a $Z = (Z_1, \dots, Z_k)$, $Z_j = \sum\limits_{p=1}^{k} a_{p,j} Y_p$, $1 \le j \le k$, véletlen vektor kovariancia mátrixa a D mátrix. Innen következik a tétel második állítása is.

Valóban,

$$\mathrm{Cov}\,(Z_j, Z_l) = \mathrm{Cov}\,\left(\sum_{p=1}^{k} a_{p,j} Y_p, \sum_{q=1}^{k} a_{q,l} Y_q\right) = \sum_{p=1}^{k}\sum_{q=1}^{k} a_{p,j} a_{q,l} \mathrm{Cov}\,(Y_p, Y_q),$$

ahonnan, mivel $\mathrm{Cov}\,(Y_p, Y_q) = 0$, ha $p \ne q$, és $\mathrm{Cov}\,(Y_p, Y_p) = 1$, $\mathrm{Cov}\,(Z_j, Z_l) = \sum\limits_{p=1}^{k} a_{p,j} a_{p,l} = d_{j,l}$, ahol $d_{j,l}$ a $D = A^*A$ mátrix j-ik sorában és l-ik oszlopában szereplő konstans.

Végül a D kovariancia mátrix akkor és csak akkor (szigorúan) pozitív definit, ha $xDx^* > 0$, azaz $\mathrm{Var}\,\left(\sum\limits_{j=1}^{k} x_j \xi_j\right) > 0$, azaz $\sum\limits_{j=1}^{k} x_j \xi_j \ne K$ 1 valószínűséggel valamilyen K konstanssal minden nem azonosan nulla $(x_1, \dots, x_k) \ne 0$ vektorra.

Ezután be tudjuk vezetni a több-dimenziós normális eloszlásfüggvények fogalmát, és meg tudjuk fogalmazni a több-dimenziós centrális határeloszlástételt.

Több-dimenziós normális eloszlások definiciója. *Definiáljuk először a több-dimenziós standard normális eloszlást. Azt mondjuk, hogy egy $(\xi_1, \dots, \xi_k)$ véletlen vektor eloszlása a k-dimenziós standard normális eloszlás, ha a $\xi_1, \dots, \xi_k$ valószínűségi változók függetlenek, és mindegyik ξ_j valószínűségi változó, $1 \le j \le k$, standard normális eloszlású. Ekvivalens megfogalmazásban azt mondhatjuk, hogy egy $(\xi_1, \dots, \xi_k)$ véletlen vektor eloszlása a k-dimenziós standard normális eloszlás, ha e véletlen vektornak létezik sűrűségfüggvénye, és az az $f(u_1, \dots, u_k) = \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{1}{2}\sum\limits_{j=1}^{k} u_j^2\right\}$ függvény.*

Egy $(\eta_1, \dots, \eta_k)$ k dimenziós véletlen vektor k dimenziós normális eloszlású vektor nulla várható értékkel, ha e véletlen vektor eloszlása megegyezik valamely $(\bar\eta_1, \dots, \bar\eta_k) = (\xi_1, \dots, \xi_k)A$ k-dimenziós vektor eloszlásával, ahol A egy $k \times k$ méretű mátrix, továbbá $(\xi_1, \dots, \xi_k)$ egy k-dimenziós standard normális eloszlású véletlen vektor.

Egy $(\zeta_1, \dots, \zeta_k)$ véletlen vektor k-dimenziós normális eloszlású vektor, ha eloszlása megegyezik egy $(\eta_1, \dots, \eta_k) + (m_1, \dots, m_k)$ véletlen vektor eloszlásával, ahol $(\eta_1, \dots, \eta_k)$ egy k-dimenziós normális eloszlású véletlen vektor, amelynek a várható értéke nulla, és $(m_1, \dots, m_k)$ k-dimenziós determinisztikus vektor.

Megjegyzés: Független (egy-dimenziós) normális eloszlású valószínűségi változók összege szintén normális eloszlású. Innen, és a több-dimenziós normális eloszlás definiciójából következik, hogy egy több-változós normális eloszlású valószínűségi változó minden koordinátája normális eloszlású. Az állítás megfordítása nem igaz. Később látni fogunk példát olyan kétváltozós valószínűségi változóra, amelynek mind a két koordinátája normális eloszlású, ő maga mégsem kétváltozós normális eloszlású véletlen vektor.

Bebizonyítok még egy eredményt, amely szükséges a többváltozós centrális határeloszlástétel kimondásához.

Tétel a több-dimenziós normális eloszlás tulajdonságairól. *Tekintsünk egy k-dimenziós $(\eta_1, \dots, \eta_k) = (\xi_1, \dots, \xi_k)A + (m_1, \dots, m_k)$ normális eloszlású valószínűségi változót, ahol A egy $k \times k$ méretű mátrix, $m = (m_1, \dots, m_k)$ k-dimenziós (véletlentől nem függő) vektor és $(\xi_1, \dots, \xi_k)$ egy k-dimenziós standard normális eloszlású véletlen vektor. Akkor $(\eta_1, \dots, \eta_k)$ $m = (m_1, \dots, m_k)$ várható értékű és $D = A^*A$ kovariancia mátrixú véletlen vektor. Egy normális eloszlású véletlen vektor kovariancia mátrixa pozitív (szemi)definit, és megfordítva, minden $k \times k$ méretű pozitív (szemi)definit mátrixhoz és k dimenziós vektorhoz létezik olyan k-változós normális eloszlású véletlen vektor, amelynek ez a kovariancia mátrixa és várható érték vektora. Továbbá egy k-dimenziós normális eloszlású valószínűségi változó eloszlását meghatározza annak m várható érték vektora és D kovariancia mátrixa.*

Megjegyzem, hogy egy rögzített D (szimmetrikus és pozitív szemidefinit) mátrixra az $A^*A = D$ egyenletnek nem egyértelmű a megoldása. Tekintsünk két különböző A és B mátrixot, amelyre $A^*A = B^*B$. A fenti tétel szerint, ha tekintünk egy k-dimenziós standard normális eloszlású $(\xi_1, \dots, \xi_k)$ vektort, illetve a segítségével definiált $(\xi_1, \dots, \xi_k)A$ és $(\xi_1, \dots, \xi_k)B$ véletlen vektorokat, akkor bár ez az utóbbi két véletlen vektor különböző, eloszlásuk megegyezik. Ugyanis mind a két (normális eloszlású) vektor nulla várható értékű és $A^*A = B^*B$ kovariancia mátrixú. Ez a tulajdonság erősen kihasználja azt, hogy normális eloszlású valószínűségi változókról van szó. Annak, hogy egy több-dimenzós normális eloszlást egyértelműen meghatároz annak várható érték vektora és kovariancia mátrixa fontos következményei vannak. Ehhez a kérdéshez később visszatérek.

A tétel bizonyítása. Az, hogy csak pozitív szemidefinit mátrixok lehetnek egy több-dimenziós normális eloszlású valószínűségi változó kovariancia mátrixai, azok viszont lehetnek ilyen kovariancia mátrixok következik a *kovariancia mátrixok jellemzéséről* szóló tételből. (Az, hogy egy normális eloszlású vektor várható értéke tetszőleges vektor lehet nyilvánvaló.) Az az állítás szorul még indoklásra, hogy egy többdimenziós normális eloszlású valószínűségi változó eloszlását meghatározza annak kovariancia mátrixa és várható érték vektora. Ezt az állítást is redukálni lehet arra az állításra, hogy egy nulla várható értékű normális eloszlású véletlen vektor eloszlását meghatározza a kovariancia mátrixa. Később látni fogjuk, hogy ez az eredmény következik a többváltozós normális eloszlásfüggvények karakterisztikus függvényének az alakjából, de itt egy másik bizonyítást ismertetek, amely a lineáris algebra egy önmagában is érdekes állításából, egy Euklideszi tér lineáris transzformációinak úgynevezett polár koordinátás felbontásából következik.

E tétel megfogalmazása előtt idézzük fel az unitér transzformációk definicióját. Egy Euklideszi tér valamely lineáris transzformációját unitérnek nevezünk, ha teljesíti az

$UU^* = I$ és $U^*U = I$ azonosságot. Valójában elég e két azonosság közül csak az egyiket megkövetelni, mert akkor a másik azonosság is szükségszerűen teljesül. Az unitér transzformációk geometriai tartalma az, hogy ezek és csak ezek az Euklideszi tér távolságtartó transzformációi. A következő lineáris algebrai eredményre lesz szükségünk.

Tétel egy mátrix polár felbontásáról. *Legyen A egy Euklideszi tér lineáris transzformációjának a mátrixa. Létezik az A mátrixnak $A = UK$ (és $A = LV$) alakú polár felbontása, ahol U (illetve V) unitér, K (illetve L) pozitív szemidefinit szimmetrikus mátrix. A K mátrix az A^*A pozitív definit, szimmetrikus mátrix egyértelműen meghatározott pozitív négyzetgyöke. (Az L mátrix az AA^* mátrix pozitív négyzetgyöke.)*

Megjegyzés. Ezen eredmény elnevezésének az az oka, hogy a pozitív szemidefinit mátrixok a nem negatív valós számoknak, az unitér transzformációk pedig az egy abszolut értékű komplex számoknak (a komplex tér forgatásainak) a természetes megfelelői, ha a komplex számok terét egy Euklideszi tér operátorainak a terével helyettesítjük. Ahogy egy tetszőleges z komplex szám felírható $z = Re^{i\varphi}$ 'polárkoordinátás' alakban, úgy tetszőleges A mátrix felírható a tételben megadott $A = UK$ alakban. Mivel a mátrixszorzás nem kommutatív, ezért az $A = UK$ és $A = LV$ előállításokban különböző mátrixok szerepelnek az általános esetben. A K mátrix alakja következik az

$$A^*A = (UK)^*(UK) = K^*(U^*U)K = K^*K = K^2 \quad \text{ezért } K = (A^*A)^{1/2}$$

azonosságból.

A tétel bizonyításának a befejezése. Tekintsük először azt az esetet, amikor egy nulla várható értékű normális eloszlású Z vektor kovariancia mátrixa az I identitás mátrix. Ekkor tudjuk, hogy Z előállítható $Z = XU$ alakban, ahol X k-dimenziós standard normális eloszlású vektor, és $I = U^*U$, azaz U unitér transzformáció. Viszont tudjuk, hogy egy k-dimenziós standard normális eloszlású X vektor sűrűségfüggvénye az

$$f(x_1, \dots, x_k) = \prod_{j=1}^{k} \frac{1}{\sqrt{2\pi}} e^{-x_j^2/2} = (2\pi)^{-k/2} \exp\left\{-\frac{1}{2}\sum_{j=1}^{k} x_j^2\right\}$$

függvény, ahonnan látható, hogy az $f(x_1, \dots, x_k)$ sűrűségfüggvény az $x = (x_1, \dots, x_k)$ pontban csak az $x = (x_1, \dots, x_k)$ vektor hosszától függ. Mivel egy U unitér transzformáció távolságtartó, innen látható, hogy az X és $Z = XU$ véletlen vektorok sűrűségfüggvényei megegyeznek.

Legyen $Y = XA$ egy tetszőleges nulla várható értékű, normális eloszlású véletlen vektor, ahol X standard normális eloszlású véletlen vektor. Véve az A mátrix $A = KU$ polárfelbontását felírhatjuk az $Y = XA = XUK = ZK$ azonosságot, ahol $Z = XU$ standard normális eloszlású véletlen vektor. Másrészt K az A^*A, azaz az Y véletlen vektor kovarianciamátrixának az (egyértelműen meghatározott) pozitív négyzetgyöke. Innen következik, hogy az $Y = ZK$ véletlen vektor eloszlását meghatározza a kovariancia mátrixa.

Ezután megfogalmazom a többváltozós centrális határeloszlástételt.

A többváltozós centrális határeloszlástétel. *Legyenek* $\xi^{(j)} = \left(\xi_1^{(j)}, \dots, \xi_k^{(j)}\right)$*,* $j = 1, 2, \dots$*, független, egyforma eloszlású* k*-dimenziós valószínűségi változók, amelyekre teljesül az* $E{\xi_l^{(1)}}^2 < \infty$*,* $1 \le l \le k$*, feltétel. Legyen a* $\xi^{(1)} = \left(\xi_1^{(1)}, \dots, \xi_k^{(1)}\right)$ *vektor várható értéke* $E\xi^{(1)} = \left(E\xi_1^{(1)}, \dots, E\xi_k^{(1)}\right)$*, kovariancia mátrixa pedig egy* D $k \times k$ *méretű mátrix. Definiáljuk az* $S^{(n)} = \left(S_1^{(n)}, \dots, S_k^{(n)}\right) = \sum\limits_{j=1}^{n} \xi^{(j)} = \left(\sum\limits_{j=1}^{n} \xi_1^{(j)}, \dots, \sum\limits_{j=1}^{n} \xi_k^{(j)}\right)$ *összegeket,* $n = 1, 2, \dots$*. Ekkor minden* $(x_1, \dots, x_k)$ k*-dimenziós vektorra érvényes a*

$$\lim_{n \to \infty} P\left(\frac{1}{\sqrt{n}}\left(S_1^{(n)} - ES_1^{(n)}\right) < x_1, \dots, \frac{1}{\sqrt{n}}\left(S_k^{(n)} - ES_k^{(n)}\right) < x_k\right) = \Phi_D(x_1, \dots, x_k) \tag{1}$$

azonosság, ahol $\Phi_D(x_1, \dots, x_k)$ *a* k*-dimenziós nulla várható értékű* D *kovariancia mátrixú normális eloszlásfüggvény értéke az* $(x_1, \dots, x_k)$ *pontban.*

1. megjegyzés. Ahhoz, hogy lássuk, hogy a fenti többváltozós centrális határeloszlástétel értelmes, tudnunk kell a következő két állítást:

i) Tetszőleges (véges) kovariancia mátrix-szal rendelkező véletlen vektorhoz létezik egy vele megegyező kovariancia mátrixú (és nulla várható értékű) normális eloszlású vektor.

ii) A tételben szereplő határeloszlást egyételműen megadtuk, azaz egy nulla várható érték vektorral rendelkező normális eloszlású véletlen vektor eloszlását egyértelműen meghatározza annak kovariancia mátrixa.

Ez a két tulajdonság azonban következik a több-dimenziós normális eloszlás tulajdonságairól szóló tételből.

2. megjegyzés. Az egy és többváltozós centrális határeloszlástétel megfogalmazásában másfajta normalizálást használtunk. Az egyváltozós esetben az összeg szórásával, azaz $\sqrt{n}\sigma$ osztottunk, ahol $\sigma^2 = \operatorname{Var}\xi$, míg a többváltozós esetben $\sqrt{n}$-nel. Az egyváltozós esetben is oszthattunk volna $\sqrt{n}$-nel, és akkor a határeloszlás egy nulla várható értékű és σ^2 szórásnégyzetű normális eloszlású valószínűségi változó eloszlása lett volna. A többváltozós esetben az egyváltozós esethez hasonló normalizálás az $n^{-1/2}\Sigma^{-1}$ mátrix-szal való szorzás lenne, ahol Σ^2 a tekintett véletlen vektorok kovariancia mátrixa, Σ ennek pozitív négyzetgyöke, azaz az az egyértelműen meghatározott pozitív definit Σ mátrix, amelynek négyzete a Σ^2 kovariancia mátrix, Σ^{-1} pedig ennek a mátrixnak az inverze. Ilyen normalizálást akkor választhatunk, ha a Σ^2 mátrix invertálható. Ez akkor teljesül, ha Σ^2 szigorúan pozitív definit. Ekkor a limesz egy standard normális eloszlású véletlen vektor. De bizonyos fontos esetekben ez a feltétel nem teljesül. Például, ha az előadás elején tekintett (szabályos) kocka véletlen dobásait tekintjük, akkor, mint a gyakorlaton megtárgyaljuk, nem invertálható kovariancia mátrix jelenik meg a határeloszlásban. Ez azzal függ össze, hogy a különböző eredményű dobások számának az összege (nem véletlen) konstans. Ez a konstans egyenlő az összes dobás számával.

3. megjegyzés. A többváltozós centrális határeloszlástétel eredeti formájában valójában egy kissé gyengébb állítást fogunk belátni, mint az itt kimondott eredmény. Azt fogjuk megmutatni, hogy a tételben definiált $\left(\frac{1}{\sqrt{n}}(S_1^{(n)}-ES_1^{(n)}),\dots,\frac{1}{\sqrt{n}}(S_k^{(n)}-ES_k^{(n)})\right)$ normalizált összegek eloszlásban konvergálnak egy nulla várható értékű és D kovariancia mátrixú normális eloszlású véletlen vektorhoz. Ez azt jelenti, hogy az (1) formula igaz minden olyan $x=(x_1,\dots,x_k)$ pontban, amely folytonossági pontja a határeloszlásnak. Ez utóbbi feltétel nem üres megszorítás, mert abban az esetben, ha a normális határeloszlás kovarianciamátrixa elfajuló, (determinánsa nulla) akkor ez az eloszlás elfajuló, és nem minden pontjában folytonos. (Mint látni fogjuk ebben az esetben a határeloszlás által indukált Stieltjes mérték egy altérbe van koncentrálva.) A többváltozós centrális határeloszlástétel eredeti alakja (tehát az a formája, amelyben az (1) relációt csak a határeloszlás folytonossági pontjaiban követeljük meg) elegendő lenne a céljainknak. Annak az oka, hogy mégis ebben a formában fogalmaztuk meg az, hogy így egyrészt egyszerűbb az állítás megfogalmazása, másrészt mint látni fogjuk, a tétel a benne szereplő valószínűségi változók bizonyos tulajdonságai miatt ebben a formában is érvényes.

Megjegyzem továbbá, hogy az egyváltozós esethez hasonlóan a többváltozós centrális határeloszlástételnek is létezik általánosítása független nem feltétlenül egyforma eloszlású véletlen vektorok normalizált összegeinek határeloszlására nagyon általános feltételek mellett. Ezzel a kérdéssel azonban itt nem foglalkozunk.

Lássuk hogyan tudjuk vizsgálni a többváltozós centrális határeloszlástétel segítségével az előadás elején megfogalmazott a) problémát dobókocka szabályosságának a vizsgálatáról.

Tekintsük az ott bevezetetett $\xi^{(j)}$, $1\le j\le n$, valószínűségi változókat, azok $S_n=(S_n^{(1)},\dots,S_n^{(6)})$ összegeit, illetve ezen összegek

$$\bar S_n=\left(\frac{1}{\sqrt{n}}(S_n^{(1)}-ES_n^{(1)}),\dots,\frac{1}{\sqrt{n}}(S_n^{(6)}-ES_n^{(6)})\right)$$

normalizáltjait. Emlékezzünk arra, hogy az $S_n^{(l)}$, $1\le l\le 6$, valószínűségi változó az l eredményű dobások számával egyenlő. Ezenkívül tudjuk, hogy $ES_l=\frac{n}{6}$, $1\le l\le 6$, és ki tudjuk számolni annak a normális eloszlású valószínűségi változók kovariancia mátrixát, amelyhez az $\bar S_n$ normalizált összegek eloszlásban tartanak. Ez a D kovariancia mátrix egyenlő a $\xi^{(j)}$ véletlen vektorok kovariancia mátrixával, és $D=(d_{l,m})$, $1\le l,m\le 6$, $d_{l,m}=E\xi_l^{(j)}\xi_m^{(j)}-E\xi_l^{(j)}E\xi_m^{(j)}=-E\xi_l^{(j)}E\xi_m^{(j)}=-\frac{1}{36}$, ha $l\ne m$, és $d_{l,l}=\frac16-(\frac16)^2=\frac{5}{36}$. Ilyen módon az $\bar S_n$ normalizált összeg eloszlására jogunk van alkalmazni az (1) aszimptotikus azonosságot. De természetesebb az $\bar S_n$ normalizált összeg koordinátáinak a négyzetösszegét, azaz a $T_n=\frac{1}{n}\sum\limits_{l=1}^{n}(S_n^{(l)}-\frac{n}{6})^2$ kifejezést vizsgálni. Felmerül a kérdés, létezik-e a T_n valószínűségi változóknak (ismert) határeloszlása. A következő két eredmény célja annak megmutatása, hogy erre a kérdésre, illetve e kérdés hasonló problémákban megjelenő természetes megfelelőire a válasz igenlő.

1. tétel az eloszlásbeli konvergencia tulajdonságairól. *Legyen adva k-dimenziós véletlen vektorok $\xi^{(n)}=(\xi_1^{(n)},\dots,\xi_k^{(n)})$, $n=1,2,\dots,$ sorozata, amelyek eloszlásban konvergálnak egy $\xi^{(0)}=(\xi_1^{(0)},\dots,\xi_k^{(0)})$ k-dimenziós véletlen vektorhoz. Legyen ezenkívül* *adva egy $u(x) = u(x_1, \dots, x_k)$ a k-dimenziós téren értelmezett folytonos függvény. (Az egyszerűség érdekében tekintsünk valós értékű függvényeket, de valójában tekinthetünk tetszőleges l-dimenziós téren értelmezett függvényt.) Ekkor az $\eta_n = u(\xi^{(n)})$, $n = 1, 2, \dots$, valószínűségi változók eloszlásban konvergálnak az $\eta_0 = u(\xi^{(0)})$ valószínűségi változóhoz.*

A tétel bizonyítása. Használjuk az eloszlásban való konvergencia folytonos függvények integráljainak a segítségével megadott jellemzését. Ekkor azt kell belátni, hogy tetszőleges folytonos és korlátos $g(x)$ függvényre $\lim\limits_{n\to\infty} Eg(\eta_n) = Eg(\eta_0)$. Viszont $g(\eta_n) = g(u(\xi^{(n)}))$, $g(\eta_0) = g(u(\xi^{(0)}))$, és a $v(x) = g(u(x))$ függvény is folytonos és korlátos a tétel feltételei miatt. Ezért, szintén a tétel feltételei alapján

$$\lim_{n\to\infty} Eg(\eta_n) = \lim_{n\to\infty} Ev(\xi^{(n)}) = Ev(\xi^{(0)}) = Eg(\eta_0).$$

A tételt bebizonyítottuk.

2. tétel az eloszlásbeli konvergencia tulajdonságairól. *Legyen F_n, $n = 1, 2, \dots$, k-dimenziós eloszlások sorozata, amelyek eloszlásban konvergálnak egy F_0 k-dimenziós eloszláshoz. Jelölje μ_{F_n} az F_n eloszlás szerint indukált Stieltjes mértéket. Ha A olyan halmaz a k-dimenziós téren, amelynek ∂A határa teljesíti a $\mu_{F_0}(\partial A) = 0$ azonosságot az F_0 határeloszlásmérték szerinti μ_{F_0} Stieltes mérték szerint, akkor $\lim\limits_{n\to\infty} \mu_{F_n}(A) = \mu_{F_0}(A)$.*

Mivel számunkra elegendő az első tétel is, a második tétel bizonyítását csak vázlatosan ismertetem.

A tétel bizonyítása. Be lehet látni, hogy az A halmaz indikátor függvényét jól lehet közelíteni, mind alulról mind felülről alkalmas folytonos függvényekkel. Részletesebben kifejtve, tetszőleges $\varepsilon > 0$ számhoz léteznek olyan $f_\varepsilon(x)$ és $g_\varepsilon(x)$ folytonos függvények, amelyekre $0 \le f_\varepsilon(x), g_\varepsilon(x) \le 1$ minden x vektorra a k-dimenziós térben, $f_\varepsilon(x) = 1$, ha $x \in \bar{A}$, ahol $\bar{A}$ az A halmaz lezártja, $f_\varepsilon(x) = 0$, ha $\rho(x, A) \ge \varepsilon$, ahol $\rho(\cdot,\cdot)$ az Euklideszi metrika a k-dimenziós térben, $g_\varepsilon(x) = 0$, ha $x \notin A$, $g_\varepsilon(x) = 1$, ha $\rho(x, A^c) \ge \varepsilon$, ahol A^c az A halmaz komplementere. Felhasználva a $\lim\limits_{n\to\infty} \int f_\varepsilon(x) F_n(\,dx) = \int f_\varepsilon(x) F_0(\,dx)$ és $\lim\limits_{n\to\infty} \int g_\varepsilon(x) F_n(\,dx) = \int g_\varepsilon(x) F_0(\,dx)$ relációkat, majd elvégezve az $\varepsilon \to 0$ határátmenetet megkapjuk a tétel állítását. A részletek kidolgozását elhagyom.

Térjünk vissza a szabályos dobókockával kapcsolatos feladat tárgyalásához. Alkalmazva az *1. tételt az eloszlásbeli konvergencia tulajdonságairól* a $k = 6$ dimenziós euklideszi térben az $u(x) = \sum\limits_{l=1}^{6} x_l^2$ függvénnyel azt kapjuk, hogy az ott bevezetett T_n valószínűségi változók eloszlásban konvergálnak egy ismert kovarianciájú, nulla várható értékű normális eloszlású vektor koordinátáinak a négyzetösszegéhez. Ilyen eredményre volt szükségünk. Felmerül még a kérdés, hogy nem lehet-e a határeloszlást egyszerűbb módon jellemezni. Ehhez a kérdéshez később még visszatérek.

A következő feladat célja annak megmutatása, hogy az előadás elején említett b) feladatot hogyan tudjuk megoldani a többváltozós centrális határeloszlástétel segítségével.

Feladat:

Legyenek $\xi_1,\dots,\xi_n$ független, a $\left[-\frac{1}{2},\frac{1}{2}\right]$ intervallumban egyenletes eloszlású valószínűségi változók. Mutassuk meg, hogy a $\sum\limits_{j=1}^n \xi_j$ és $\sum\limits_{j=1}^n \xi_j^2$ összegek normalizáltjainak, azaz a $\sqrt{\frac{12}{n}}\sum\limits_{j=1}^n \xi_j$ és $\sqrt{\frac{180}{n}}\sum\limits_{j=1}^n \left(\xi_j^2-\frac{1}{12}\right)$ valószínűségi változóknak az együttes eloszlása a két-dimenziós standard normális eloszláshoz konvergál, ha $n\to\infty$.

Megoldás: $E\xi=0$, $E\xi^2=\frac{1}{12}$, $\operatorname{Var}\xi=\frac{1}{12}$, $\operatorname{Var}\xi^2=E\xi^4-(E\xi^2)^2=\frac{1}{80}-\frac{1}{144}=\frac{1}{180}$. Továbbá $\operatorname{Cov}(\xi,\xi^2)=E\xi^3-E\xi E\xi^2=0$. Ezért a $\left(\sqrt{12}\xi_j,\sqrt{180}\left(\xi_j^2-\frac{1}{12}\right)\right)$, $j=1,2,\dots$, véletlen vektorok függetlenek, nulla várható értékkel és az identitás kovariancia mátrix-szal. Innen, és a több-dimenziós centrális határeloszlástételből következik a feladat állítása.

Be akarjuk látni a több-dimenziós centrális határeloszlástételt. Tudjuk eddigi eredményeink alapján, hogy ehhez elegendő megmutatni azt, hogy a normalizált összegek karakterisztikus függvényei konvergálnak egy megfelelő kovarianciájú és nulla várható értékű normális eloszlású vektor karakterisztikus függvényéhez. Ennek bizonyításához először ki kell számolni a több-dimenziós normális eloszlások karakterisztikus függvényeit. Erről szól a következő tétel.

Tétel a több-dimenziós normális eloszású valószínűségi változók karakterisztikus függvényéről. *Legyen $\eta=(\eta_1,\dots,\eta_k)=(\xi_1,\dots,\xi_k)A+(m_1,\dots,m_k)$ egy k-dimenziós normális eloszlású valószínűségi változó, ahol $m=(m_1,\dots,m_k)$ k-dimenziós (determinisztikus) vektor, $A=(a_{j,l})$, $1\le j,l\le k$, $k\times k$ méretű mátrix, továbbá $\xi=(\xi_1,\dots,\xi_k)$ k-dimenziós standard normális eloszlású véletlen vektor. Ekkor az $(\eta_1,\dots,\eta_k)$ véletlen vektor karakterisztikus függvénye az*

$$Ee^{i(t,\eta)}=Ee^{i(t_1\eta_1+\cdots+t_k\eta_k)}=e^{i(t,m)-tA^*At^*/2}=\exp\left\{i\sum_{j=1}^k t_jm_j-\frac{1}{2}\sum_{j=1}^k\sum_{l=1}^k d_{j,l}t_jt_l\right\}$$

függvény, ahol (x,y) jelöli az $x=(x_1,\dots,x_k)$ és $y=(y_1,\dots,y_k)$ vektorok $(x,y)=\sum\limits_{j=1}^k x_jy_j$ skalárszorzatát, $t=(t_1,\dots,t_k)$, $d_{j,l}$ az A^*A kovariancia mátrixának j-ik sorában, és l-ik oszlopában szereplő konstans. A $D=A^*A$ mátrix megegyezik az $\eta=(\eta_1,\dots,\eta_k)$ véletlen vektor kovariancia mátrixával.

Bizonyítás:

$$Ee^{i(t,\eta)}=Ee^{i(t,\xi A+m)}=e^{i(t,m)}Ee^{i(tA^*,\xi)}=e^{i(t,m)}e^{-(tA^*,tA^*)/2}=e^{i(t,m)-tA^*At^*/2},$$

mert, ha $tA^*=\bar t=(\bar t_1,\dots,\bar t_k)$, akkor $Ee^{i(tA^*,\xi)}=Ee^{i(\bar t_1\xi_1+\cdots+\bar t_k\xi_k)}=\prod\limits_{j=1}^k Ee^{i\bar t_j\xi_j}=\prod\limits_{j=1}^k e^{-\bar t_j^2/2}=e^{-(\bar t,\bar t)/2}=e^{-(tA^*,tA^*)/2}$.

Az η véletlen vektor $D = (d_{j,l})$ kovariancia mátrixában a j-ik sor l-ik eleme $d_{j,l} = \operatorname{Cov}(\eta_j, \eta_l) = \operatorname{Cov}\left(\sum\limits_{p=1}^{k} a_{p,j}\xi_p, \sum\limits_{q=1}^{k} a_{q,l}\xi_q\right) = \sum\limits_{p=1}^{k} a_{p,j}a_{p,l}E\xi_p^2 = \sum\limits_{p=1}^{k} a_{p,j}a_{p,l}$, és ez az A^*A mátrix j-ik sorában és l-ik oszlopában álló elem. Ezért az η véletlen vektor kovariancia mátrixa a $D = A^*A$ mátrix. Az nyílvánvaló, hogy az η véletlen vektor várható értéke m.

Következmény. *Egy több-dimenziós normális eloszlású véletlen vektor eloszlását egyértelműen meghatározza annak várható érték vektora és kovariancia mátrixa.*

Megjegyzés. Ezen állítás bizonyítását már láttuk lineáris algebrai módszerekkel. Most azt mutatjuk meg, hogy az egyszerűen következik a karakterisztikus függvények tulajdonságaiból is.

A következmény bizonyítása. Az előző tétel alapján egy normális eloszlású véletlen vektor karakterisztikus függvényét egyértelműen meghatározza annak várható érték vektora és kovariancia mátrixa. Viszont egy eloszlást meghatároz annak karakterisztikus függvénye.

A többváltozós centrális határeloszlástétel bizonyításában azt kell megmutatni, hogy a tekintett véletlen normalizált összegek karakterisztikus függvényei konvergálnak a határeloszlás karakterisztikus függvényéhez. Ennek bizonyítása nem nehezebb, mint a megfelelő egyváltozós állítás bizonyítása. Sőt, az állítás igazolását vissza lehet vezetni az egyváltozós esetre az alábbi egyszerű, de tanulságos eredmény segítségével.

Lemma többváltozós véletlen vektorok eloszlásának konvergenciájának a jellemzéséről. *Legyen adva független, k-változós $\xi^{(n)} = (\xi_1^{(n)}, \dots, \xi_k^{(n)})$, $n = 1, 2, \dots$, véletlen vektorok egy sorozata. Ezek a véletlen vektorok akkor és csak akkor konvergálnak eloszlásban egy $\xi^{(0)} = (\xi_1^{(0)}, \dots, \xi_k^{(0)})$ véletlen vektorhoz, ha koordinátáik bármely $\sum\limits_{r=1}^{n} a_r\xi_r^{(n)}$ lineáris kombinációi eloszlásban konvergálnak a $\sum\limits_{j=1}^{n} a_r\xi_r^{(0)}$ valószínűségi változóhoz.*

A lemma bizonyítása. Azt kell megérteni, hogy mit jelentenek a lemmában megfogalmazott eloszlásban való konvergenciák a karakterisztikus függvények nyelvén. Jelölje $\varphi_n(t_1, \dots, t_k) = Ee^{i(t_1\xi_1^{(n)} + \dots + t_k\xi_k^{(n)})}$ a $\xi^{(n)} = (\xi_1^{(n)}, \dots, \xi_k^{(n)})$, $n = 0, 1, 2, \dots$, véletlen vektor és $\varphi_n(t|a_1, \dots, a_k) = Ee^{it(a_1\xi_1^{(n)} + \dots + a_k\xi_k^{(n)})}$ a $\sum\limits_{j=1}^{n} a_j\xi_j^{(n)}$ lineáris kombináció karakterisztikus függvényét. A fenti jelölésekkel, — felhasználva azt a tényt, hogy eloszlások konvergenciája ekvivalens azok karakterisztikus függvényeinek konvergenciájával, — a lemma állítását úgy fogalmazhatjuk át, hogy az alábbi a) és b) tulajdonságok ekvivalensek.

a) $\lim\limits_{n\to\infty} \varphi_n(t|a_1, \dots, a_k) = \varphi_0(t|a_1, \dots, a_k)$ minden t valós számra és $a_1, \dots, a_k$ paraméterre.

b) $\lim\limits_{n\to\infty} \varphi_n(t_1, \dots, t_k) = \varphi_0(t_1, \dots, t_k)$ minden $t_1, \dots, t_k$ szám k-asra.

Viszont a b) tulajdonság következik az a) tulajdonságból $t=1$, $a_r=t_r$, $1\le r\le k$, választással, az a) tulajdonság pedig következik a b) tulajdonságból $t_r=ta_r$, $1\le r\le k$, választással. A lemmát bebizonyítottuk.

A többváltozós centrális határeloszlástétel bizonyítása. Az előző lemma alapján elég azt bebizonyítani, hogy tetszőleges $a_1,\dots,a_k$ szám k-asra az $\bar S_n(a_1,\dots,a_k)=\frac{1}{\sqrt{n}}\sum\limits_{r=1}^k a_r(S_n^{(r)}-ES_n^{(r)})$ kifejezések eloszlásban konvergálnak a $\zeta(a_1,\dots,a_k)=\sum\limits_{r=1}^k a_r\zeta_r$ véletlen vektorhoz, ahol $(\zeta_1,\dots,\zeta_k)$ nulla várható értékű és D kovariancia mátrixú normális eloszlású véletlen vektor. Vezessük be az $\eta_j=\sum\limits_{r=1}^k a_r\xi_r^{(j)}$, $1\le j\le n$, valószínűségi változókat. Ekkor $\eta_1,\dots,\eta_n$ független, egyforma eloszlású valószínűségi változók, és $\bar S_n=\sum\limits_{j=1}^n(\eta_j-E\eta_j)$. Ezért a bizonyítandó állítás következik az egyváltozós centrális határeloszlástételből független, egyforma eloszlású valószínűségi változók összegére és abból az észrevételből, hogy $\zeta(a_1,\dots,a_k)$ nulla várható értékű normális eloszlású valószínűségi változó, és $\operatorname{Var}\zeta(a_1,\dots,a_k)=\operatorname{Var}\eta_j$.

Következő témánk a többváltozós normális eloszlások néhány fontos tulajdonsága.

Láttuk, hogy egy több-dimenziós normális eloszlást meghatároz annak várható értéke és kovariancia mátrixa. Ennek egyik következményét annak fontossága miatt tétel formájában mondom ki.

Tétel normális eloszlású véletlen vektorok korrelálatlan koordinátáinak függetlenségéről. *Legyenek egy $(\xi_1,\dots,\xi_k)$ normális eloszlású vektor koordinátái korrelátlanok. Akkor e koordináták független, normális eloszlású valószínűségi változók.*

Bizonyítás. A $(\xi_1-E\xi_1,\dots,\xi_k-E\xi_k)$ normális eloszlású vektor várható értéke nulla, kovariancia mátrixa pedig e véletlen vektor koordinátáinak korrelátlansága miatt diagonális. Legyenek a diagonálisban álló elemek σ_j^2, $1\le j\le k$. Tekintsünk $\eta_1,\dots,\eta_k$ független, standard normális eloszlású valószínűségi változókat, és vezessük be segítségükkel a $(\sigma_1\eta_1,\dots,\sigma_k\eta_k)$ véletlen vektort. Ez több-dimenziós normális eloszlású vektor, amelynek várható értéke és kovariancia mátrixa megegyezik a $(\xi_1-E\xi_1,\dots,\xi_k-E\xi_k)$ normális eloszlású vektor várható értékével és kovariancia mátrixával. Ezért e két véletlen vektor eloszlása megegyezik. Így a $\sigma_1\eta_1,\dots,\sigma_k\eta_k$ valószínűségi változók függetlenségéből következik, hogy a $\xi_1-E\xi_1,\dots,\xi_k-E\xi_k$, illetve a $\xi_1,\dots,\xi_k$ valószínűségi változók is függetlenek és normális eloszlásúak.

Tudjuk, hogy független valószínűségi változók korrelálatlanok. A teljesség kedvéért mutatok egy példát korrelálatlan, de nem független valószínűségi változókra. Ez a példa mutatja, hogy a fenti eredményben nagyon fontos volt, hogy egy normális eloszlású véletlen vektor koordinátáit tekintettük.

Példa korrelálatlan, de nem független valószínűségi változókra. Legyen ξ egyenletes eloszlású valószínűségi változó a $\left[-\frac12,\frac12\right]$ intervallumban, Ekkor a ξ és $\eta=\xi^2$ valószínűségi változók korrelálatlanok, de nem függetlenek. Valóban, $E\xi=0$, $E\eta=E\xi^2=\frac{1}{12}$, $E\xi\eta=E\xi^3=0$, $\operatorname{Cov}(\xi,\eta)=E\xi\eta-E\xi E\eta=0$. Másrészt ξ és η nem függetlenek, sőt az η valószínűségi változó a ξ valószínűségi változó determinisztikus függvénye. Egy lehetséges formális indoklása annak, hogy ξ és η nem független a következő: Legyen $0 < a < 1$ tetszőleges szám. Ekkor $\{\omega\colon\ \eta < a^2\} = \{\omega\colon\ |\xi| < a\}$. Ezért $P(\xi < a, \eta < a^2) = P(\xi < a)$, tehát $P(\xi < a, \eta < a^2) \neq P(\xi < a)P(\eta < a^2)$.

A következő állítást, amely szintén azzal kapcsolatos, hogy normális eloszlású véletlen vektor koordinátáinak függetlensége következik azok korrelálatlanságából feladat formájában fogalmazom meg, és a gyakorlaton fogjuk tárgyalni. Ennek az eredménynek fontos következménye van néhány valószínűségszámítási és statisztikai vizsgálatban.

Feladat:

Legyen (ξ, η) normális eloszlású vektor $m = (m_1, m_2) = (E\xi, E\eta)$ várható értékkel és

$$\begin{pmatrix} \sigma_{1,1}, & \sigma_{1,2} \\ \sigma_{2,1}, & \sigma_{1,1} \end{pmatrix} = \begin{pmatrix} E\xi^2 - (E\xi)^2, & E\xi\eta - E\xi E\eta \\ E\xi\eta - E\xi E\eta, & E\eta^2 - (E\eta)^2 \end{pmatrix}$$

kovarianciamátrix-szal. Ekkor létezik a ξ valószínűségi változónak $\xi = a\eta + \zeta$ alakú előállítása alkalmas a konstanssal, és az η valószínűségi változótól független ζ normális eloszlású valószínűségi változóval. Ez azt jelenti, hogy ha (ξ, η) két-dimenziós normális eloszlású valószínűségi változó, akkor az első koordináta kifejezhető, mint a második kooordináta konstansszorosának és egy a második koordinátától független normális eloszlású valószínűségi változó összege. A kívánt a konstanst explicit módon megadhatjuk az $a = \frac{\sigma_{1,2}}{\sigma_{2,2}}$ képlet segítségével.

Hogy általánosítható a fenti állítás abban az esetben, ha ξ és η vektorváltozók is lehetnek?

Megoldás: A $\zeta = \xi - \frac{\sigma_{1,2}}{\sigma_{1,1}}\eta$ valószínűségi változó független az η valószínűségi változótól. Ehhez a több-dimenziós normális eloszlás tulajdonságai alapján elég azt ellenőrizni, hogy $\operatorname{Cov}(\zeta, \eta) = 0$. Innen következik a feladat állítása.

Az az eset, amikor $\xi = (\xi_1, \dots, \xi_s)$, $\eta = (\eta_1, \dots, \eta_p)$, és $(\xi_1, \dots, \xi_s, \eta_1, \dots, \eta_p)$ egy $s + p$ dimenziós normális eloszlású valószínűségi változó hasonlóan tárgyalható. Be lehet látni, hogy létezik olyan A mátrix, amelyre η és $\xi - \eta A$ függetlenek. Ennek érdekében először azt érdemes megmutatni, hogy létezik olyan U unitér mátrix amelyre $\eta U = (\bar\eta_1, \dots, \bar\eta_p) = \bar\eta$ vektor koordinátái függetlenek. Ez abból látható, hogy ha az η véletlen vektor D kovarianciamátrixát $D = U^*\Lambda U$ alakban írjuk fel, ahol U unitér Λ pedig diagonális mátrix, akkor az $\bar\eta = \eta U$ véletlen normális eloszlású vektor kovarianciamátrixa Λ, ahonnan következik, hogy az $\bar\eta$ mátrix koordinátái függetlenek. Legyen $\bar\xi_r = \xi_r - \sum\limits_{k=1}^{p} \frac{E\xi_r\bar\eta_k}{E\bar\eta_k^2}\bar\eta_k$, $r = 1, \dots, s$. Ezt mátrixjelöléssel $\bar\xi = \xi - \bar\eta B$ formában írhatjuk. Ekkor $(\xi - \bar\eta B, \bar\eta)$ olyan $p + s$ dimenziós, normális eloszlású valószínűségi változó, amelynek első s és utolsó p koordinátája korrelálatlan, ezért független. Mivel a $\xi - \bar\eta B$ és $\bar\eta$ vektorok függetlenek, ezért $\zeta = \xi - \bar\eta B = \xi - \eta U B$ független az $\eta = \bar\eta U^*$ vektortól.

Megjegyzés: A valószínűségszámítás illetve statisztika finomabb kérdéseinek vizsgálatában bevezették a feltételes valószínűség és feltételes eloszlás fogalmát olyan esetekben is, amikor a feltétel nulla valószínűséggel következik be. Bizonyos vizsgálatokban fontos kiszámolni, hogy mi a feltételes eloszlása egy több-dimenziós normális vektor bizonyos koordinátáinak azon feltétel mellett, hogy a többi koordináta értékét rögzítjük. E feladat megoldásának kulcslépése a fent tárgyalt feladat megoldása. E kérdésre visszatérek akkor, amikor a feltételes valószínűségeket fogjuk tanulni.

Láttuk, hogy egy normális eloszlású véletlen vektor koordinátái is normális eloszlásúak. Ez az állítás nem megfordítható. Ennek megértése érdekében példát mutatok olyan két-dimenziós (ξ,η) véletlen vektorra, amelynek mind a két koordinátája, azaz mind a ξ mind az η valószínűségi változó normális eloszlású, viszont a (ξ,η) vektor (mint két-dimenziós véletlen vektor) nem normális eloszlású.

Példa véletlen vektorra, amely nem normális eloszlású, noha koordinátái normális eloszlásúak. *Definiáljuk a következő $(\Omega,\mathcal{B},\mathbf{P})$ valószínűségi mezőt: $\Omega=[0,1]$, $\mathcal{B}$ a Borel σ-algebra $[0,1]$-en, és $\mathbf{P}$ a Lebesgue mérték. Definiáljuk a következő ξ és η valószínűségi változókat ezen a mezőn: $\xi(x)=\Phi^{-1}(x)$,*

$$\eta(x)=\begin{cases}\xi(1-x) & \textit{ha } 0\le x<\frac12\\ \xi\left(x-\frac12\right) & \textit{ha } \frac12\le x\le 1\end{cases}.$$

Az ebben a példában definiált ξ és η valószínűségi változók normális eloszlásúak, de a (ξ,η) véletlen vektor nem normális eloszlású.

Indoklás: A ξ és η valószínűségi változók azonos eloszlásúak. Továbbá,

$$P(\xi>x)=\lambda(\Phi(x),1])=1-\Phi(x)\ .$$

Az, hogy a (ξ,η) véletlen vektor nem normális eloszlású következik például a $P(\xi+\eta=0)=\frac12$ azonosságból. Ugyanis, ha (ξ,η) normális eloszlású lenne, akkor az lenne a $\xi+\eta$ valószínűségi változó is. Ennek viszont ellentmond a $P(\xi+\eta=0)=\frac12$ reláció.

Kiszámolom egy több-dimenziós normális eloszlású vektor sűrűségfüggvényét is, (feltéve, hogy az létezik).

Tétel a több-dimenziós normális eloszlású vektor sűrűségfüggvényének az alakjáról. *Legyen adva egy $\eta=(\eta_1,\dots,\eta_k)$ k-dimenziós normális eloszlású valószínűségi változó, amelynek $m=(m_1,\dots,m_k)=(E\eta_1,\dots,E\eta_k)$ a várható érték vektora és $D=(d_{j,l})$, $d_{j,l}=\mathrm{Cov}\,(\eta_j,\eta_l)$, $1\le j,l\le k$, a kovariancia mátrixa. Az η k-dimenziós valószínűségi változónak akkor és csak akkor van sűrűségfüggvénye, ha a D kovariancia mátrix invertálható. Ha a D kovariancia mátrix invertálható, akkor az $\eta=(\eta_1,\dots,\eta_k)$ véletlen vektor sűrűségfüggvénye a következő alakú:*

$$f(x)=f(x_1,\dots,x_k)=\frac{1}{(2\pi)^{k/2}\det D^{1/2}}\exp\left\{-(x-m)D^{-1}(x-m)^*/2\right\},$$

ahol $x=(x_1,\dots,x_k)\in R^k$ k-dimenziós vektor.

Bizonyítás: Az $\eta=(\eta_1,\dots,\eta_k)$ véletlen vektor eloszlása megegyezik egy olyan $\bar\eta=(\bar\eta_1,\dots,\bar\eta_k)=(\xi_1,\dots,\xi_k)A+(m_1,\dots,m_k)$ véletlen vektornak az eloszlásával, amelyikre

ξ_j, $1 \le j \le k$, független standard normális eloszlású valószínűségi változók, és $D = A^*A$. Jegyezzük meg, hogy a lineáris algebra standard eredményei szerint az A és A^* mátrixok egyszerre invertálhatóak vagy nem invertálhatóak, és a $D = A^*A$ mátrix akkor és csak akkor invertálható, ha az A mátrix invertálható. Ezért, ha a D mátrix nem invertálható, akkor az η vektornak nincs sűrűségfüggvénye, ha pedig a D mátrix invertálható, akkor a következő módon számolhatunk:

Alkalmazva az $x = yA + m$ transzformációt $\xi = (\xi_1, \dots, \xi_k)$, $x = (x_1, \dots, x_k)$, $y = (y_1, \dots, y_k)$ és $\varphi(y) = \varphi(y_1, \dots, y_k) = \frac{1}{(2\pi)^{k/2}} e^{-(y_1^2 + \cdots + y_k^2)/2}$ jelöléssel kapjuk, hogy tetszőleges mérhető $B \subset R^k$ halmazra

$$P(\eta \in B) = P(\bar\eta \in B) = P\left(\xi \in (B-m)A^{-1}\right) = \int_{(y_1,\dots,y_k) \in (B-m)A^{-1}} \varphi(y)\, dy$$
$$= \frac{1}{|\det A|} \int_{(x_1,\dots,x_k) \in B} \varphi\left((x-m)A^{-1}\right)\, dx$$

alakú, ahol $|\det A|$ az $x = yA + m$ leképezés Jacobian-ja.

E formulából kiolvasható, hogy a vizsgált normális eloszlású valószínűségi változó sűrűségfüggvénye a $\frac{1}{|\det A|}\varphi\left((x-m)A^{-1}\right)$ függvény. Annak érdekében, hogy bebizonyítsuk a tételt vegyük észre, hogy mivel $D = A^*A$, ezért $\det D = \det A^* \det A = (\det A)^2$, és $|\det A| = \det D^{1/2}$. Továbbá,

$$\begin{aligned}
\varphi\left((x-m)A^{-1}\right) &= \frac{1}{(2\pi)^{k/2}} \exp\left\{ -\frac{((x-m)A^{-1}, (x-m)A^{-1})}{2} \right\} \\
&= \frac{1}{(2\pi)^{k/2}} \exp\left\{ -\frac{(x-m)A^{-1}\left(A^{-1}\right)^*(x-m)^*}{2} \right\} \\
&= \frac{1}{(2\pi)^{k/2}} \exp\left\{ -\frac{(x-m)(A^*A)^{-1}(x-m)^*}{2} \right\} \\
&= \frac{1}{(2\pi)^{k/2}} \exp\left\{ -\frac{(x-m)D^{-1}(x-m)^*}{2} \right\},
\end{aligned}$$

mert $A^{-1}\left(A^{-1}\right)^* = A^{-1}\left(A^*\right)^{-1} = \left(A^*A\right)^{-1}$. Innen következik a Tétel állítása.

Megjegyzés: Egy több-dimenziós normális eloszlású valószínűségi változó karakterisztikus függvényét megadó képletben a kovariancia mátrix szerepel, míg a sűrűségfüggvényét megadó képletben a kovariancia mátrix inverze. Mivel a karakterisztikus függvény kiszámolásában nem kell invertálni a kovariancia mátrixot, ezért a karakterisztikus függvény segítségével könnyebb vizsgálni a normális eloszlásfüggvények tulajdonságait.

Visszatérek az előadás elején megfogalmazott a) problémához. Megfogalmazom annak egy természetes általánosítását, amely fontos szerepet játszik a matematikai statisztikában. Az itt megfogalmazott feladat megoldására kidolgozott módszert hívják az irodalomban χ-négyzet próbának. Megmutatom, hogy a χ-négyzet próba alapjául szolgáló eredmény a több-dimenziós centrális határeloszlástétel és bizonyos lineáris algebrai eredmények következménye. A következő feladattal fogok foglalkozni.

Legyen adva k urna, és ellenőrizzük azt a feltevést, amely szerint ha egy golyót véletlenül bedobunk ezen urnák valamelyikébe, akkor az p_j, $p_j > 0$, valószínűséggel esik a j-ik urnába, $\sum\limits_{j=1}^{k} p_j = 1$. Ezen feltevés ellenőrzésének az érdekében dobjunk egymástól függetlenül n golyót ezekbe az urnákba, és jelölje $\nu_n(j)$ a j-ik urnába eső golyók számát. Döntsük el a kapott eredmény alapján, hogy feltevésünk helyes volt-es.

Be fogjuk látni, hogy feltevésünk teljesülése esetén a $\sum\limits_{j=1}^{k} \frac{(\nu_n(j)-np_j)^2}{np_j}$ valószínűségi változóknak létezik határeloszlásuk $n \to \infty$ esetén, és ez a határeloszlás az úgynevezett $k-1$ szabadságfokú χ^2 (szóban khi négyzet) eloszlás.

Megfogalmazom ezt az eredményt pontosabban is. Először bevezetem a χ^2 eloszlások definicióját.

A k szabadságfokú χ^2 eloszlás definiciója. *Legyen $\xi_1, \dots, \xi_k$ k darab független standard normális eloszlású valószínűségi változó. Ekkor a $\sum\limits_{j=1}^{k} \xi_j^2$ valószínűségi változó eloszlását nevezzük k szabadságfokú $\chi^2(k)$ eloszlásnak.*

Megjegyzés: Láttuk a gyakorlaton, hogy a 2 szabadságfokú $\chi^2(2)$ eloszlás a $\lambda = \frac{1}{2}$ paraméterű exponenciális eloszlás, azaz az az eloszlás, amelynek sűrűségfüggvénye az $f(x) = \frac{1}{2}e^{-x/2}$, ha $x \ge 0$, és $f(x) = 0$, ha $x < 0$.

Ezután megfogalmazom a χ^2 próba alapjául szolgáló eredményt.

Tétel urnadobás eredményének aszimptotikus viselkedéséről. *Legyen adva k darab urna, amelyekbe bedobunk egymástól függetlenül golyókat úgy, hogy mindegyik golyó p_j valószínűséggel esik a j-ik urnába, $1 \le j \le k$, $\sum\limits_{j=1}^{k} p_j = 1$. Jelölje $\nu_n(j)$ a j-ik urnába eső golyók számát n dobás után. Ekkor a*

$$\sum_{j=1}^{k} \frac{(\nu_n(j) - np_j)^2}{np_j}, \qquad n = 1, 2, \dots$$

valószínűségi változók eloszlásban konvergálnak a $k-1$ szabadságfokú $\chi^2(k-1)$ eloszláshoz, ha $n \to \infty$. (Az urnák k száma rögzített.)

Megjegyzem, hogy a fenti tételben megjelenő határeloszlás csak az urnák k számától függ, de nem függ a p_j, $1 \le j \le k$, valószínűségektől. Ez jelzi azt, hogy természetes statisztikát vezettünk be, olyat amelyben a különböző urnákban levő golyók számának az eltérése annak várható értékétől egyforma fontos szerepet játszik. Az, hogy a határeloszlás a $k-1$ szabadságfogú $\chi^2(k-1)$ eloszlás, azzal függ össze, hogy bár k véletlen szám súlyozott négyzetösszegét tekintettük, (az egyes urnákba eső golyók számának eltérését tekintettük azok várható értékétől), de ezek között van egy determinisztikus összefüggés. Nevezetesen az, hogy az összes urnába eső golyók száma minusz azok várható értéke nullával egyenlő. Ezt informálisan úgy szokták interpretálni, hogy $k-1$ szabadsági fokkal rendelkező véletlen vektorok koordinátáinak a négyzetösszegét tekintettük, illetve azok határeloszlását. Ilyen esetben a határeloszlást olyan véletlen összeg adja meg, amelyben mindegyik szabadsági foknak egy olyan összeadandó felel meg, amely független a többi összeadandótól, és az egy standard normális eloszlású valószínűségi változó négyzete.

A tétel bizonyítását több lépésben adom meg. Az első lépésben bebizonyítom a határeloszlástételt, de a határeloszlás nem a számunkra rokonszenves alakban fog megjelenni.

Lemma a χ-négyzet statiszika határeloszlásának létezéséről. *Tekintsük az urnadobás eredményének aszimptotikus viselkedéséről szóló tételt és az abban definiált $\nu_n(j)$ valószínűségi változókat illetve a segítségükkel bevezetett $T_n = \sum\limits_{j=1}^{k} \frac{(\nu_n(j)-np_j)^2}{np_j}$, $n=1,2,\dots$, kifejezéseket. A T_n valószínűségi változók eloszlásban konvergálnak egy olyan nulla várható értékű $(\xi_1,\dots,\xi_k)$ normális eloszlású véletlen vektor koordinátáinak $\sum\limits_{j=1}^{k} \xi_j^2$ négyzetösszegéhez, amelynek $D=(d_{i,j})$, $1\le i,j\le k$, kovariancia mátrixát a*

$$d_{i,j} = -\sqrt{p_ip_j}, \quad \text{ha } i\neq j,\ \ 1\le i,j\le k, \quad d_{j,j}=1-p_j \quad 1\le j\le k \tag{2}$$

képlet adja meg.

A lemma bizonyítása. Definiáljuk a következő $X^{(l)} = (X_1^{(l)},\dots,X_k^{(l)})$ k-dimenziós véletlen vektorokat. $X_j^{(l)} = \frac{1}{\sqrt{p_j}}$, ha az l-ik dobásban a golyó a j-ik urnába esett, és $X_j^{(l)} = 0$, ha nem a j-ik urnába esett, $1\le l\le n$, $1\le j\le k$. Vezessük be továbbá a $\bar S^{(n)} = (\bar S_1^{(n)},\dots,\bar S_k^{(n)})$ normalizált összeget, ahol $\bar S_j^{(n)} = \frac{1}{\sqrt n}(S_j^{(n)} - ES_j^{(n)})$, és $S_j^{(n)} = \frac{1}{\sqrt n}\sum\limits_{l=1}^{n} X_j^{(l)}$, $1\le j\le k$. Ekkor a $X^{(l)}$, $1\le l\le n$, vektorok függetlenek, kovariancia mátrixuk a (2) formulában definiált $D=(d_{i,j})$ mátrix, $\frac{\nu_n(j)}{\sqrt{p_j}} = S_j^{(n)}$, $1\le j\le k$, és a többváltozós centrális határeloszlástétel alapján az $\bar S^{(n)}$ véletlen vektorok eloszlásban konvergálnak egy nulla várható értékű a (2) formulában megadott kovarianca mátrixú $(\xi_1,\dots,\xi_k)$ normális eloszlású véletlen vektorhoz. Ezenkívül igaz a $T_n = \sum\limits_{j=1}^{k}\left(\bar S_j^{(n)}\right)^2$ azonosság. Ezért a a lemma állítása következik az *1. tétel az eloszlásbeli konvergencia tulajdonságairól* néven megfogalmazott eredményből $u(x_1,\dots,x_k) = \sum\limits_{j=1}^{k} x_j^2$ választással.

Ezután a tétel bizonyításához elegendő a következő lemmát belátni.

Lemma a χ^2 statisztikában megjelenő határeloszlástétel jellemzéséről. *Legyen $(\xi_1,\dots,\xi_k)$ normális eloszlású valószínűségi változó nulla várható értékkel és a (2) for-*

mulában definiált $D = (d_{i,j})$, $1 \le i,j \le k$, kovariancia mátrix-szal. Ekkor $\sum_{j=1}^k \xi_j^2$ $\chi^2(k-1)$ eloszlású valószínűségi változó.

Ennek a lemmának a bizonyításában hasznos a következő önmagában is érdekes lemma.

Lemma normális eloszlású véletlen vektor koordináta négyzetösszeg eloszlásáról. *Legyen $\eta = (\eta_1, \dots, \eta_k)$ k-dimenziós normális eloszlású valószínűségi változó nulla várható értékkel és D kovariancia mátrix-szal. Legyenek a D mátrix sajátértékei a $\lambda_1, \dots, \lambda_k$ számok (multiplicitással). Ekkor a $\sum_{j=1}^k \eta_j^2$ valószínűségi változó eloszlása megegyezik egy $\sum_{j=1}^k \lambda_j \xi_j^2$ valószínűségi változó eloszlásával, ahol $\xi_1, \dots, \xi_k$ független standard normális eloszlású valószínűségi változók.*

Bizonyítás: A D mátrix felírható $D = U\Lambda U^*$ alakban, ahol U unitér, Λ pedig olyan diagonális mátrix, amelynek átlójában a D mátrix λ_j sajátértékei vannak (multiplicitással). (Az U mátrix is felírható explicit módon a D mátrix sajátvektorainak segítségével, de erre a tényre most nincs szükségünk.)

Az $\eta = (\eta_1, \dots, \eta_k)$ véletlen vektor eloszlása megegyezik egy $\bar\eta = (\bar\eta_1, \dots, \bar\eta_k) = \xi\Lambda^{1/2}U^* = (\xi_1, \dots, \xi_k)\Lambda^{1/2}U^*$ véletlen vektor eloszlásával, ahol $\xi = (\xi_1, \dots, \xi_k)$ standard normális eloszlású véletlen vektor. Valóban $\bar\eta$ normális eloszlású véletlen vektor, melynek várható értéke nulla és kovariancia mátrixa a

$$(\Lambda^{1/2}U^*)^*\Lambda^{1/2}U^* = U\Lambda^{1/2}\Lambda^{1/2}U^* = U\Lambda U^* = D$$

mátrix. Innen az is következik, hogy a $\sum_{j=1}^k \eta_j^2$ valószínűségi változó eloszlása megegyezik a $\sum_{j=1}^k \bar\eta_j^2$ valószínűségi változó eloszlásával. Vegyük észre, hogy az $\tilde\eta = (\tilde\eta_1, \dots, \tilde\eta_k) = \bar\eta U = (\bar\eta_1, \dots, \bar\eta_k)U$ vektorra teljesül a $\sum_{j=1}^k \bar\eta_j^2 = \sum_{j=1}^k \tilde\eta_j^2$ azonosság, mert U unitér, tehát távolságtartó transzformáció. Viszont $\tilde\eta = \bar\eta U = \xi\Lambda^{1/2}U^*U = \xi\Lambda^{1/2}$. Ez azt jelenti, hogy a $\sum_{j=1}^k \eta_j^2$ valószínűségi változó eloszlása megegyezik a $\sum_{j=1}^k (\lambda_j^{1/2}\xi_j)^2 = \sum_{j=1}^k \lambda_j\xi_j^2$ valószínűségi változó eloszlásával, és ez a Lemma állítása.

A χ^2 statisztikában megjelenő határeloszlástétel jellemzéséről szóló lemma bizonyítása. Az előző lemma alapján elég azt megmutatni, hogy a (2) képletben definiált D kovariancia mátrixnak az 1 $k-1$ multiplicitású sajátértéke (azaz a D mátrixnak $k-1$ ortonormált 1 sajátértékkel rendelkező sajátvektora van) és ezenkívül még a nulla a sajátértéke 1-szeres multiplicitással.

Írjuk fel a D mátrixot $D = I - B$ alakban, ahol I az identitás mátrix, $B = (b_{i,j})$, $b_{i,j} = \sqrt{p_i}\sqrt{p_j}$, $1 \le i,j \le k$. A B mátrix sajátvektorait és sajátértékeit egyszerűen ki tudjuk számolni. Valóban, az $e_1 = (\sqrt{p_1}, \dots, \sqrt{p_k})$ vektor a B mátrix 1 sajátértékű sajátvektora. Egészítsük ki az e_1 vektort egy tetszőleges $e_1, e_2, \dots, e_k$ ortonormált bázissá az R^k Euklideszi térben. Ekkor az e_j, $2 \le j \le k$, vektorok a B mátrix nulla sajátértékű sajátvektorai, mert az $(e_1, e_j) = 0$ azonosságból és a B mátrix alakjából következik, hogy $e_j B = 0$. Abból, hogy az e_j vektorok a B mátrix sajátvektorai $\lambda_1 = 1$ és $\lambda_j = 0$, ha $j \ge 2$ sajátértékkel következik, hogy e vektorok a $D = I - B$ mátrixnak is sajátvektorai $\bar\lambda_j = 1 - \lambda_j$ sajátértékekkel, azaz $\bar\lambda_1 = 0$, és $\bar\lambda_j = 1$, ha $2 \le j \le k$.

Kiegészítés: Megjegyzés a többváltozós centrális határeloszlástétel alakjáról. A többváltozós centrális határeloszlástételt az (1) formulában fogalmaztam meg. Azt állítottam, hogy az (1) formulában felírt azonosság minden $x = (x_1, \dots, x_k)$ vektorra érvényes. Ugyanakkor az előadásban bebizonyított többváltozós centrális határeloszlástétel azt állítja, hogy a tekintett véletlen vektorok eloszlásban konvergálnak a megfelelő kovariancia mátrixú normális eloszláshoz. Ez azt jelenti, hogy az (1) formulában felírt azonosság érvényes a határeloszlás minden folytonossági pontjában. Abban az esetben, ha a (normális) határeloszlás kovariancia mátrixa elfajuló, akkor ez az eloszlás az Euklideszi tér egy (valódi) alterére van koncentrálva, és a határeloszlásnak vannak nem folytonossági pontjai. Ez azt jelenti, hogy az (1) formulában többet állítottunk, mint amit a többváltozós centrális határeloszlástételben bebizonyítottunk. A többváltozós centrális határeloszlástétel eredeti alakja megfelel a számunkra, és az (1) formulának a tételben megfogalmazott erősebb formájának nincs gyakorlati jelentősége. A teljesség kedvéért mégis belátom, hogy az (1) formula a Tételben megfogalmazott formában is érvényes. Ez azon múlik, hogy olyan véletlen vektorok határeloszlását tekintettük, amelyek ugyanabba az altérbe vannak koncentrálva, mint a (normális) határeloszlás.

Az (1) formula általános alakjának bizonyítása érdekében bebizonyítok egy lemmát. Ennek megfogalmazása előtt felidézem, hogy a lineáris algebrában bevezettük egy mátrix rangját. Ez egyenlő a mátrix sorvektorai (vagy oszlopvektorai) által kifeszített altér dimenziójával. (A sor illetve oszlopvektorok által kifeszített altér dimenziója megegyezik. Ezt és néhány további alapvető lineáris algebrai tényt külön magyarázat nélkül fogok használni a bizonyításban.)

Lemma több-dimenziós véletlen vektorok eloszlásáról. *Legyen egy nulla várható értékű $\xi = (\xi_1, \dots, \xi_k)$ véletlen vektor D kovariancia mátrixának a rangja r, $r \le k$. Ekkor létezik a k-dimenziós Euklideszi térnek olyan csak a D kovariancia mátrixtól függő S r-dimenziós altere, amelyre igaz, hogy a ξ véletlen vektor egy valószínűséggel ebbe az S altérbe van koncentrálva, azaz $P((\xi_1(\omega), \dots, \xi_k(\omega)) \in S) = 1$. Sőt igaz a következő állítás is. Létezik r olyan $\eta_j = \sum\limits_{l=1}^{k} a(j,l)\xi_l$, $1 \le j \le r$, valószínűségi változó csak a D kovariancia mátrixtól függő $a(j,l)$ együtthatókkal, amelyekre $E\eta_j^2 = 1$, $1 \le j \le r$, $E\eta_j\eta_l = 0$, ha $1 \le j, l \le r$, és $j \ne l$. Ezenkívül a ξ_j valószínűségi változók is kifejezhetőek az η_j valószínűségi változók lineáris kombinációjaként $\xi_j = \sum\limits_{l=1}^{r} b(j,l)\eta_l$, $1 \le j \le k$, alakban alkalmas csak a D kovariancia mátrixtól függő $b(j,l)$ együtthatókkal.*

Az (1) formula kiterjesztése az általános esetre azon múlik, hogy a tekintett normalizált részletösszegek ugyanabba az altérbe vannak koncentrálva, mint a határeloszlás. Ezt a lineáris alteret az a képlet határozza meg, amely megmutatja, hogy a ξ véletlen vektort hogyan lehet kifejezni az r η_j valószínűségi változó lineáris kombinációjaként.

A lemma bizonyítása. A D kovariancia mátrixot fel lehet írni $D = U\Lambda^2 U^*$ alakban, ahol U unitér Λ pedig diagonális mátrix, amelynek első r diagonális eleme, $\lambda_1, \dots, \lambda_r$ szigorúan pozitív, a további diagonális elemek pedig nullák. (Itt használtuk ki, hogy a D mátrix rangja r.) Jelölje $u_{j,l}$ a D mátrix fenti reprezentációjában szereplő U unitér mátrix j-ik sorában és l-ik oszlopában levő elemet. Definiáljuk az $\bar\eta = (\bar\eta_1, \dots, \bar\eta_k)$ vektort az $\bar\eta_j = \sum\limits_{l=1}^k u_{j,l}\xi_l$, $1 \le j \le k$, képlettel. (Mátrix jelöléssel $\bar\eta = \xi U^*$). Ekkor $\xi = \bar\eta U$, azaz $\xi_j = \sum\limits_{l=1}^k \bar\eta_l u_{l,j}$, és némi számolással azt kapjuk, hogy a $\bar\eta$ kovariancia mátrixa az $U^* D U = \Lambda^2$ mátrix. Innen speciálisan az is adódik, hogy $E\bar\eta_j^2 = 0$, és $\bar\eta_j = 0$ egy valószínűséggel minden $r + 1 \le j \le k$ indexre. Definiáljuk az $\eta_j = \frac{\bar\eta_j}{\lambda_j}$, $1 \le j \le r$, valószínűségi változókat, és legyen $a(j,l) = \lambda_j u_{j,l}$, $1 \le j \le r$, $1 \le l \le k$. Ekkor $E\eta_j^2 = 1$, $E\eta_j\eta_l = 0$, ha $j \ne l$, $1 \le j,l \le r$, és $\eta_j = \sum\limits_{l=1}^k a(j,l)\xi_l$, $1 \le j \le r$. Továbbá $\xi_j = \sum\limits_{l=1}^r b(j,l)\eta_l$, $1 \le j \le k$, $b(j,l) = u_{l,j}\lambda_l$ együtthatókkal. Végül az S altér, ahová a ξ vektor van koncentrálva megegyezik a $\xi_j = \sum\limits_{l=1}^r b(j,l)y_l$, $1 \le j \le k$, $1 \le l \le r$, alakú vektorokból álló altérrel.

Az (1) formula bizonyítása az általános esetben. Alkalmazzuk az előző lemmát a többváltozós centrális határeloszlástételben bevezetett mennyiségek segítségével a $\xi^{(n)} = (\xi_1^{(n)}, \dots, \xi_k^{(n)})$, $\xi_j^{(n)} = \frac{1}{\sqrt{n}}(S_j^{(n)} - ES_j^{(n)})$, $1 \le j \le k$, véletlen vektorra minden $n = 1, 2, \dots$ indexre. Jegyezzük meg, hogy ezen vektorok D kovariancia mátrixa nem függ az n indextől. A többváltozós centrális határeloszlástételből következik, hogy az $\eta_j^{(n)} = \sum\limits_{l=1}^k a(j,l)\xi_l^{(n)}$, $1 \le j \le r$, véletlen vektorok eloszlásban konvergálnak az r-változós standard normális eloszlásfüggvényhez, amely minden pontban folytonos. Másrészt a minket érdeklő valószínűségeket ki lehet fejezni

$$P(\xi_1^{(n)} < x_1, \dots, \xi_k^{(n)} < x_k) = P\left(\sum_{l=1}^r b(1,l)\eta_l^{(n)} < x_1, \dots, \sum_{l=1}^r b(k,l)\eta_l^{(n)} < x_k\right)$$

alakban. Azt kell belátni, hogy

$$\begin{aligned}\lim_{n\to\infty} P&\left(\sum_{l=1}^r b(1,l)\eta_l^{(n)} < x_1, \dots, \sum_{l=1}^r b(k,l)\eta_l^{(n)} < x_k\right)\\ &= P\left(\sum_{l=1}^r b(1,l)\zeta_l < x_1, \dots, \sum_{l=1}^r b(k,l)\zeta_l < x_k\right),\end{aligned}$$

ahol $\zeta_1, \ldots, \zeta_r$ független standard normális eloszlású valószínűségi változók. Ezt be lehet látni a többváltozós centrális határeloszlástétel és a *2. tétel az eloszlásbeli konvergencia tulajdonságairól* segítségével. A részletek kidolgozását elhagyom.