A Valószínűségszámítás I. előadássorozat tizenkettedik előadása.
2007. május 8.

A többváltozós centrális határeloszlástétel

Láttuk, hogy a centrális határeloszlástétel értékes információt ad arról, hogy hogyan viselkedik (nagyszámú) független, egyforma eloszlású valószínűségi változó összegének az eloszlása. Bizonyos kérdések vizsgálatában felmerül ennek a problémának olyan általánosabb változata, ahol nem valós, hanem vektorértékű független, egyforma eloszlású valószínűségi változók összegeinek a viselkedése érdekel minket. Lássunk néhány olyan problémát, ahol ilyen kérdések jelennek meg.

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 a dobásszámok egyhatoda plusz egy kis eltérés. De mely eltéréseket tekinthetünk kicsinek? Ha csak azt nézzük, mennyi annak a valószínűsége annak, 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?

Létezik a centrális határeloszlástételnek olyan több-dimenziós változata, amelyet az irodalomban többváltozós centrális határeloszlástételnek neveznek. A fenti kérdések vizsgálatában ez az eredmény bizonyult hasznosnak. Ezt hasonlóan lehet bizonyítani a már tárgyalt egyváltozós centrális határeloszlástételhez. Jelen előadás fő célja nem a bizonyítás, hanem az eredmény megértése. Elsősorban azt kell megértenünk, hogy mi a centrális határeloszlástétel megfogalmazásában megjelenő normális eloszlás többváltozós megfelelője. Ennek keretében tisztáznunk kell azt is, hogy mi a valós értékű valószínűségi változók esetében bevezetett várható érték és szórásnégyzet megfelelője vektor értékű valószínűségi változók esetében. Ugyancsak fontos megérteni vektorértékű valószínűségi változók eloszlásban való konvergenciájának a definicióját, illetve annak legfontosabb tulajdonságait.

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 között. (Több-dimenziós valószínűségi változók függetlenségének a fogalmát ismertettem a 9. előadás elején, és az megtalálható az előadás ismertetésének első és második oldalán.) 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. Célunk annak bizonyítása, 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 eloszlásnak nevezni.

Lássuk, hogyan lehet tárgyalni az előbb említett két problémát egy ilyen jellegű eredmény segítségével. Az a) feladat vizsgálatának érdekében vezessük be a következő valószínűségi változókat: Ha a dobókockát n alkalommal dobjuk fel, akkor legyen ξ_j, $1 \le j \le n$, a következő 6-dimenziós valószínűségi változó: A ξ_j valószínűségi változó l-ik koordinátája 1, ha a j-ik dobás értéke l, $1 \le l \le 6$, és ξ_j összes többi koordinátája nulla. Ekkor a ξ_j valószínűségi változók függetlenek, (az egyes valószínűségi változók koordinátái nem feltétlenül függetlenek), és 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. Hasonló a helyzet a b) példában, ha 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 $\left[-\frac12, \frac12\right]$ intervallumban egyenletes eloszlású valószínűségi változók.

Ismertetni fogom a centrális határoszlástétel természetes megfelelőjét abban az esetben, ha független és egyforma eloszlású véletlen vektorok alkalmasan normált összegének a viselkedését vizsgáljuk. Annak érdekében, hogy ezt megtehessem, be kell vezetnem a határeloszlásban megjelenő normális eloszlás több-dimenziós megfelelőjét, amit több-dimenziós normális eloszlásnak fogunk nevezni. De ehhez előbb be kell vezetnem az egy-dimenziós esetben definiált várható érték és szórásnégyzet több-dimenziós megfelelőjét. Ezenkívül meg kell érteni, hogy a várható érték és szórásnégyzet tulajdonságai hogyan öröklődnek, ha azokanak a több-dimenziós esetben megjelenő megfelelőit vizsgáljuk. 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 a bizonyítást elhagyom.

Tétel vektor értékű valószínűségi változók várható értékének és kovariancia mátrixának néhány tulajdonságáról. *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*

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

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 mátrixok jelenhetnek meg, mint 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 úgy leírni, hogy önmagukban érthetőek legyenek, Egy kiegészítésben leírom a felhasznált eredmények ismertetését és bizonyítását.

Először idézzük fel 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, ami ebben az előadásban nem fog előfordulni), hogy* $d_{j,l} = \bar{d}_{l,j}$, *ahol $\bar{z}$ a z komplex szám konjugáltja, azaz, ha* $z = a + ib$, *akkor* $\bar{z} = a - ib$. *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 elemével. Ekkor xDx^* a szokásos vektorszorzást 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 fent leírt koordinátarendszerben leírt definiciók koordinátarendszertől független ,,absztrakt" definicióját is és megérteni a két definició kapcsolatát. Ennek leírását (a bizonyítások többségének elhagyásával) tartalmazza a lineáris algebrai összefoglaló. A következő eredményben megfogalmazom azt a fontos eredményt, amely megadja a kovariancia mátrixok jellemzését. A bizonyítás felhasznál bizonyos nem triviális lineáris algebrai eredményeket is.

Tétel a 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 azt is tudjuk, hogy egy valószínűségi változó szórásnégyzete 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 felelnek meg magasabb dimenzióban. Magának a tételnek a bizonyítása egy alább megfogalmazandó nem triviális lineáris algebrai eredményen alapul, amelyik azt mondja ki, hogy egy pozitív szemidefinit mátrix felírható, mint egy alkalmas mátrix nényzete. Ez az állítás annnak a ténynek a több-dimenziós megfelelője, amely szerint pozitív számokból lehet négyzetgyököt vonni. Jegyezzük meg, 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 és szemidefinit, és $D = A^2$. Eme megszorítás esetén az $D = A^*A = AA$ 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 $D = (d_{j,l})$, $d_{j,l} = \mathrm{Cov}\,(Z_j, Z_l)$, $1 \le j, l \le k$, kovariancia mátrixát. Ekkor D szimmetrikus mátrix, mert $d_{j,l} = d_{l,j}$, azaz $\mathrm{Cov}\,(Z_j, Z_l) = \mathrm{Cov}\,(Z_l, Z_j)$. Másrészt tetszőleges $x = (x_1, \dots, x_k)$ k-dimenziós vektorra

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

és $\mathrm{Var}\left(\sum\limits_{j=1}^k x_j Z_j\right) \ge 0$. 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 feladat második állítása is. Viszont $\mathrm{Cov}\,(Z_j, Z_l) = \mathrm{Cov}\left(\sum\limits_{p=1}^k a_{p,j} Y_p, \sum\limits_{q=1}^k a_{q,l} Y_q\right)$, ezért $\mathrm{Cov}\,(Z_j, Z_l) = \sum\limits_{p=1}^k \sum\limits_{q=1}^k a_{p,j} a_{q,l} \mathrm{Cov}\,(Y_p, Y_q)$, 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 $\sum\limits_{j=1}^k x_j \xi_j \ne K$ valamilyen 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 $f(u_1,\dots,u_k)=\frac{1}{(2\pi)^{k/2}}\exp\left\{-\frac12\sum\limits_{j=1}^k u_j^2\right\}$ 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: A több-dimenziós normális eloszlás definiciójából, illetve a 10. előadás elején tárgyalt feladat azon következményéből, hogy független normális eloszlású valószínűségi változók lineáris kombinációi is normális eloszlásúak 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.

Megfogalmazok még egy később bizonyítandó eredményt, amely szükséges a több-dimenziós centrális eloszlástétel 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.*

Jegyezzük meg, hogy 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.

Tanulságos lehet a következő példa, amely segíthet az előző tétel jobb megértésében. Legyen $\xi = (\xi_1,\dots,\xi_k)$ k-dimenziós standard normális eloszlású véletlen vektor és U egy unitér transzformáció a k-dimenziós térben, azaz legyen $UU^* = U^*U = I$, ahol I az identitás mátrix a k-dimenziós térben. Az előző tétel szerint a $\xi = \xi I$ és $\eta = \xi U$ véletlen vektorok eloszlása megegyezik, mert mind a két vektor normális eloszlású, (minden koordinátában) nulla várható értékkel és $UU^* = I$ kovariancia mátrix-szal. Ez annak ellenére így van, hogy az U unitér mátrix lényegesen különbözhet az I identitás mátrixtól.

Viszont a fenti példa állításának okát meg tudjuk érteni, ha felidézzük az unitér transzformáció szemléletes jelentését. Egy unitér transzformáció távolságtartó lineáris leképezés, ezért a tér egy forgatása (és utána esetleg a tükrözése egy $k-1$ dimenziós altérre). Másrészt láttuk, hogy egy standard normális eloszlású véletlen vektor sűrűségfüggvénye egy x pontban csak e pontnak az origótól mért $|x|$ távolságától függ. Ezért egy ilyen valószínűségi változó eloszlása nem változik, ha elforgatjuk, (és akkor sem, ha utána esetleg tükrözzük egy $k-1$-dimenziós altérre). A fenti példa ezt a tényt fejezi ki.

Megfogalmazom ezen előadás fő eredményét.

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*

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

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.

A fenti két tulajdonság következik a több-dimenziós normális eloszlás tulajdonságairól szóló tételből. E tétel legtöbb állítását már igazoltuk. De még meg kell indokolni, hogy egy több-dimenziós normális eloszlást miért határoz meg annak várható érték vektora és kovariancia mátrixa.

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 következő feladat mutatja, nem invertálható kovariancia mátrix jelenik meg a határeloszlásban. Ez azzal függ össze, hogy a különböző eredményű dobások számának az összege (nem véletlen) konstans, mert egyenlő az összes dobás számával.

Feladat:

Legyenek $\xi^{(j)} = (\xi_1^{(j)}, \dots, \xi_6^{(j)})$ független, $1 \le j \le n$, egyforma eloszlású véletlen vektorok $P(\xi_k^{(j)} = 1) = \frac{1}{6}$, minden $1 \le k \le 6$, $1 \le j \le n$ indexre, és $\xi_{k'}^{(j)} = 0$, ha $k' \ne k$, és $\xi_k^{(j)} = 1$, minden $1 \le k, k' \le 6$ indexre. Legyen $S = \frac{1}{\sqrt{n}} \sum\limits_{j=1}^{n} \xi_j$ e véletlen vektorok normalizált összege. Ekkor a $\xi^{(j)}$ és S vektorok kovariancia mátrixa az a $D = (d_{i,k})$, $1 \le i, k \le 6$, mátrix, amelyre $d_{i,k} = -\frac{1}{36}$, ha $i \ne k$, $d_{k,k} = \frac{5}{36}$. A D mátrix nem invertálható.

Megoldás: A $\xi^{(j)}$ vektor D mátrixának elemei $d_{i,k} = E\xi_i^{(j)}\xi_k^{(j)} - E\xi_i^{(j)}E\xi_k^{(j)} = -E\xi_i^{(j)}E\xi_k^{(j)} = -\frac{1}{36}$,ha $i \ne k$, és $d_{k,k} = E\left(\xi_k^{(j)}\right)^2 - (E\xi_k^{(j)})^2 = \frac{1}{6} - \left(\frac{1}{6}\right)^2 = \frac{5}{36}$. Az S véletlen vektor kovariancia mátrixa ugyanez a D mátrix. A D mátrix nem invertálható, mert a sorösszegei nullával egyenlőek.

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.

A többváltozós centrális határeloszlástétel bizonyítása hasonló a klasszikus egyváltozós esethez. Az ott bevezetett fogalmaknak és eredményeknek megadhatóak a több-dimenziós általánosításaik, és a megfelelő eredmények hasonlóan bizonyíthatóak. Ezért csak a fogalmakat és az eredményeket fogom ismertetni. A bizonyítás a karakterisztikus függvény fogalmának több-dimenziós általánosításán alapul, illetve azon, hogy az egydimenziós esetben bizonyítható eredmények természetes módon általánosíthatóak a több-dimenziós esetre is. Ennek a kérdésnek rövid tárgyalását a kiegészítésben írom le, ami nem része a fő anyagnak. Viszont az előadás fő részében ismertetem (bizonyítás nélkül) azokat az eredményeket, amelyek segítenek a többváltozós centrális határeloszlástétellel kapcsolatos feladatok megoldásában.

Több-dimenziós eloszlásfüggvények eloszlásban való konvergenciájának a definiciója. *Legyen $F_n(x_1,\dots,x_k)$, $n=1,2,\dots$, k-dimenziós eloszlásfüggvények sorozata. Azt mondjuk, hogy az $F_n(x_1,\dots,x_k)$ eloszlásfüggvények eloszlásban konvergálnak egy k-dimenziós $F(x_1,\dots,x_k)$ eloszlásfüggvényhez, ha $\lim\limits_{n\to\infty} F_n(x_1,\dots,x_k)=F(x_1,\dots,x_k)$ az $F(\cdot)$ (határ)eloszlásfüggvény minden $(x_1,\dots,x_k)$ folytonossági pontjában.*

Tétel az eloszlásban való konvergencia jellemzéséről folytonos függvények segítségével. *$F_n(u_1,\dots,u_k)$, $n=1,2,\dots$ k-dimenziós eloszlásfüggvények sorozata akkor és csak akkor konvergál eloszlásban egy $F(u_1,\dots,u_k)$ k-dimenziós eloszlásfüggvényhez, ha minden a k-dimenziós téren értelmezett folytonos, korlátos $g(u_1,\dots,u_k)$ függvényre teljesül a*

$$\lim_{n\to\infty}\int g(u_1,\dots,u_k)\,F_n(\,du_1,\dots,\,du_k)=\int g(u_1,\dots,u_k)\,F(\,du_1,\dots,\,du_k)$$

azonosság.

Megjegyzés: A többváltozós centrális határeloszlástétel valójában azt mondja ki, hogy a tekintett normalizált összegeknek van határeloszlása, és az a tételben definiált több-dimenziós normális eloszlás. Ez a tétel megfelel céljainknak, de ha precízek vagyunk, akkor az általunk kimondott tétel némi magyarázatra szorul. Ugyanis az ott kimondott limesz reláció nem következik a normalizált összegek eloszlásban való konvergenciájából abban az esetben, ha a határeloszlásként megjelenő több-dimenziós normális eloszlás kovariancia mátrixa nem invertálható. Ekkor ugyanis ez az eloszlás egy altérre van koncentrálva, és a határeloszlás nem folytonos minden pontban. A tételben kimondott állítás ekkor is érvényes, de erre a tényre nincs szükségünk. Ennek az állításnak az igazolásához azt kell kihasználni, hogy a tételben szereplő összeadandók eloszlása ugyanabba az altérbe van koncentrálva, mint a határeloszlás.

Az egydimenziós esetben ki tudtuk számolni a centrális határeloszlástétel alapján annak (közelítő) valószínűségét, hogy sok független valószínűségi változó normalizált összege egy intervallumba esik. A több-dimenziós esetben is szeretnénk hasonló eredményeket alkalmazni. De ebben az esetben egy általánosabb probléma jelenik meg. Olyan eredményre van szükségünk, amely segít kiszámítani annak valószínűségét, hogy független véletlen vektorok normalizált összegei valamilyen általános geometriai alakzatba, például egy gömbbe vagy ellipszoidba esnek. Olyan eredmények érdekelnek minket, amelyek azt mondják ki, hogy ilyen valószínűségek határértéke is létezik, és ez a határérték egyenlő azzal a számmal, amit a többváltozós centrális határeloszlástétel sugall. A következő két tétel valamelyikéből, amelyek az eloszlásban való konvergencia tulajdonságairól szólnak, levezethetjük a centrális határeloszlás ilyen irányú következményeit.

1. Tétel az eloszlásbeli konvergencia tulajdonságairól. *Legyen ξ_n, $n = 1, 2, \dots$, k-dimenziós véletlen vektorok sorozata, amelyek eloszlásban konvergálnak egy ξ_0 k-dimenziós véletlen vektorhoz, és legyen F 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 $\eta_k = F(\xi_k)$, $k = 1, 2, \dots$, valószínűségi változók eloszlásban konvergálnak az $\eta_0 = F(\xi_0)$ valószínűségi változóhoz.*

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 $\lim\limits_{n\to\infty} \mu_{F_n}(A) = \mu_{F_0}(A)$.*

A fenti két tétel bizonyítását elhagyom, noha azok nem nehezek. Valójában e tételek sokkal általánosabb változata is érvényes és viszonylag egyszerűen bizonyítható szeparábilis metrikus tereken definiált valószínűségi mértékekre. (De ehhez egy általánosabb fogalomrendszer bevezetésére lenne szükségünk, amivel itt nem fogok foglalkozni.) Viszont megmutatom, hogyan lehet megoldani a fenti két eredmény valamelyikének és a többváltozós centrális határeloszlástétel segítségével az előadás elején megfogalmazott feladatot kockák szabályosságának az ellenőrzéséről.

Megmutatom, hogy hogyan lehet az eloszlásbeli konvergencia tulajdonságairól szóló 1. tétel segítségével segítségével megoldani a bevezetés elején megfogalmazott feladatot kockák szabályosságának ellenőrzését alkalmas dobássorozat megfigyelésének segítségével. (A 2. tételt is alkalmazhatnánk, de az első tétel alkalmazása kissé kényelmesebbnek bizonyult.) Valójában egy általánosabb eredményt fogok tárgyalni, amely elvezet a matematikai statisztika egyik fontos módszeréhez, az úgynevezett χ-négyzet próbához.

Tekintsük a következő feladatot:

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. 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 eloszlás.

Megfogalmazom ezt az eredményt pontosabban is. Előtte 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 9. előadás 5. feladatában, hogy a 2 szabadságfokú $\chi^2(2)$ eloszlás a $\lambda=\frac12$ paraméterű exponenciális eloszlás, azaz az az eloszlás, amelynek sűrűségfüggvénye az $f(x)=\frac12 e^{-x/2}$, ha $x\ge 0$, és $f(x)=0$, ha $x<0$.

Ezután megfogalmazom a χ^2 próba alapjául szolgáló eredményt. A bizonyítás ismerete nem kötelező a vizsgán.

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 az n-ik 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 összeadandó felel meg, amelyik független a többi összeadandótól, és az egy standard normális eloszlású valószínűségi változó négyzete. Az eredmény bizonyítását egy nem kötelező kiegészítésben írom le.

Lássuk, hogy az előadás elején említett második feladatot hogy tudjuk megoldani a több-változós 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.

A fenti eredmények bizonyításában kihasználtuk azt a tényt, hogy egy normális eloszlású vektor eloszlását egyértelműen meghatározza annak várható érték vektora és kovariancia mátrixa. Ezt a kiegészítésben bebizonyítom e véletlen vektor karakterisztikus függvényének a segítségével. De ez a bizonyítás mély, eddig nem tanult matematikai ismereteket is felhasznál. Itt kiszámolom egy normális eloszlású véletlen vektor sűrűségfüggvényét. Megmutatom, hogy ez a sűrűségfüggvény akkor és csak akkor létezik, ha a normális eloszlású véletlen vektor kovariancia mátrixa invertálható. Ebben az esetben a sűrűségfüggvény kifejezhető a várható érték és kovariancia mátrix segítségével. Innen következik, hogy a normális eloszlást meghatározza a kovariancia mátrixa és várható érték mátrixa abban az esetben, ha a kovariancia mátrix invertálható. A bizonyítandó állítás igazolása az általános esetben visszavezethető erre az eredményre, de ezt az érvelést nem dolgozom ki. Viszont adok egy másik, lineáris algebrai eredményeken alapuló bizonyítást arra, hogy egy több-dimenziós normális eloszlást meghatároz annak várható érték vektora és kovariancia mátrixa. A következő lineáris algebrai eredményt fogom használni, amelyet mátrixok polár felbontásának is neveznek.

Tétel mátrixok polár felbontásáról. *Egy A $k \times k$ (négyzetes) mátrix felírható $A = UK$ alakban, ahol U unitér és K szimmetrikus, pozitív definit mátrix. (Hasonlóan felbontható $A = K_1 U_1$ alakban egy szimmetrikus, pozitív definit és unitér mátrix szorzatára.)*

Megjegyzés. Azért nevezik ezt a felbontást polár felbontásnak, mert ez annak a megfelelője, hogy egy z komplex szám felírható $z = Re^{i\varphi}$ alakban, ahol $R = |z|$ a z szám abszolut értéke, az $e^{i\varphi}$ tényezővel való szorzás pedig φ szöggel való forgatást jelent. Mivel két mátrix szorzata nem feltétlenül kommutativ az $A = UK$ és $A = K_1 U_1$ felbontásban különböző mátrixok szerepelhetnek.

A normális eloszlás egyértelműségéről szóló tétel bizonyítása a fenti lineáris algebrai összefüggés segítségével. Elég belátni, hogy egy $\zeta = \xi A$ nulla várható értékű normális eloszlású vektor eloszlását, ahol ξ k-dimenziós standard normális eloszlású vektor, A pedig $k \times k$ méretű mátrix meghatározza annak $D = A^* A$ kovariancia mátrixa. Ennek érdekében írjuk fel az A mátrix $A = UK$ polár felbontását, és vegyük észre, hogy $D =$

$A^*A = KU^*UK = K^2$, és mivel egy pozitív (szemi)definit mátrix pozitív szemidefinit négyzetgyöke egyértelműen meghatározott, ezért $K = D^{1/2}$. Innen $\zeta = \xi A = \xi U D^{1/2} = \eta D^{1/2}$, ahol $\eta = \xi U$. Viszont láttuk, hogy egy U unitér transzformációra az $\eta = \xi U$ vektor is standard normális eloszlású. Innen következik, hogy a $\zeta = \eta D^{1/2}$ véletlen vektor eloszlása csak a $D^{1/2}$ ezért csak a D mátrixtól függ.

Most megadom egy több-dimenziós normális eloszlású vektor sűrűségfüggvényét.

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 $\varphi(y) = \varphi(y_1, \dots, y_k) = \frac{1}{(2\pi)^{k/2}} e^{-(y_1^2 + \dots + y_k^2)/2}$ jelöléssel kapjuk, hogy tetszőleges mérhető $B \subset R^k$ halmazra

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

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 $|\det A| = \det D^{1/2}$, ahonnan

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

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.

Annak, hogy egy normális eloszlású véletlen vektor eloszlását egyértelműen meghatározza annak várható érték vektora és kovariancia mátrixa meglepő és fontos következményei vannak. Lássuk be a következő (fontos) állítást.

Feladat:

Legyen $(\xi_1,\dots,\xi_k)$ k-dimenziós normális eloszlású véletlen vektor, amelynek koordinátái korrelálatlanok, azaz $\mathrm{Cov}\,(\xi_i,\xi_j)=0$, ha $i\neq j$. Ekkor $\xi_1,\dots,\xi_k$ független normális eloszlású valószínűségi változók.
Megoldás: Feltehetjük az általánosság megszorítása nélkül, hogy $E\xi_j=0$, $1\le j\le k$. Legyen $E\xi_j^2=\sigma_j^2$. Tekintsünk független standard normális eloszlású $\eta_1,\dots,\eta_k$ valószínűségi változókat, és vezessük be a $\tilde\eta=(\sigma_1\eta_1,\sigma_2\eta_2,\dots,\sigma_k\eta_k)$ véletlen vektort. Ekkor $\tilde\eta$ nulla várható értékű, normális eloszlású véletlen vektor, amelynek kovariancia mátrixa megegyezik a $\xi=(\xi_1,\dots,\xi_k)$ kovariancia mátrixával. Ezért e két vektor eloszlása megegyezik. Innen következik, hogy a $\xi_1, \dots, \xi_k$ valószínűségi változók függetlenek és normális eloszlásúak.

Annak érdekében, hogy a fenti feladat eredményét jobban tudjuk értékelni oldjuk meg a következő feladatot.

Feladat:

Mutassunk példát két korrelálatlan ξ és η valószínűségi változóra, amelyek nem függetlenek.
Megoldás: Sok egyszerű példát adhatunk. Tekintsük például a következő példát: 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, $E\xi = 0$, $E\eta = E\xi^2 = \frac{1}{12}$, $E\xi\eta = E\xi^3 = 0$, $\mathrm{Cov}\,(\xi, \eta) = E\xi\eta - E\xi E\eta = 0$. 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 $P(\xi < a, \eta < a^2) = P(\xi < a)$, azaz $P(\xi < a, \eta < a^2) \neq P(\xi < a)P(\eta < a^2)$.

Végül lássunk példát olyan két-dimenziós (ξ, η) 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.

Feladat:

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: $\xi(x) = \Phi^{-1}(x)$,

$$\eta(x) = \begin{cases} \xi(1-x) & \text{ha } 0 \le x < \frac{1}{2} \\ \xi\left(x - \frac{1}{2}\right) & \text{ha } \frac{1}{2} \le x \le 1 \end{cases} .$$

Lássuk be, hogy ξ és η normális eloszlású valószínűségi változók, de a (ξ, η) vektor nem normális eloszlású valószínűségi vektor.
Megoldás: A ξ és η valószínűségi változók azonos eloszlásúak. Továbbá,

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

Az, hogy (ξ, η) vektor nem normális eloszlású következik például a $P(\xi + \eta = 0) = \frac{1}{2}$ azonosságból. Ugyanis, ha (ξ, η) normális eloszlású lenne, akkor az lenne a $\xi + \eta$ valószínűségi változó is. Ennek viszont ellentmond a $P(\xi + \eta = 0) = \frac{1}{2}$ reláció.

Nem kötelező feladat.

Legyen ξ egy l változós normális eloszlású valószínűségi változó, A egy $k \times l$ méretű mátrix. Mutassuk meg (alapvető lineáris algebrai ismeretek felhasználásával), hogy $\eta = \xi A$ egy k változós normális eloszlású valószínűségi változó.

Kiegészítés: A több-dimenziós centrális határeloszlástétel bizonyításáról.
A bizonyításban az egyváltozós esetben ismertetett eredmények több-dimenziós változataira van szükségünk.

Az egydimenziós esethez hasonlóan a fenti tétel és Weierstrass második approximációs tétele (pontosabban annak több-dimenziós általánosítása) segítségével eloszlásfüggvények konvergenciáját jól lehet vizsgálni a karakterisztikus függvények alább bevezetett több-dimenziós általánosításának a segítségével.

Több-dimenziós valószínűségi változók karakterisztikus függvényének a definiciója. *Legyen $\xi = (\xi_1, \dots, \xi_k)$ k-dimenziós valószínűségi változó valamely $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, és jelölje $F(u_1, \dots, u_k) = P(\xi_1 < u_1, \dots, \xi_k < u_k)$, $-\infty < u_j < \infty$, $1 \le j \le k$, a $\xi = (\xi_1, \dots, \xi_k)$ véletlen vektor eloszlásfüggvényét. A $\xi = (\xi_1, \dots, \xi_k)$ k-dimenziós ξ valószínűségi változónak a karakterisztikus függvénye a*

$$\varphi(t_1, \dots, t_k) = Ee^{i(t_1\xi_1 + \dots + t_k\xi_k)} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{i(t_1u_1 + \dots + t_ku_k)} F(\,du_1, \dots, du_k),$$
$$-\infty < t_j < \infty, \quad 1 \le j \le k,$$

függvény. Adva egy $F(u_1, \dots, u_k)$ k-dimenziós eloszlásfüggvény az F eloszlásfüggvény karakterisztikus függvényét is definiálni fogjuk mint egy F eloszlású $\xi = (\xi_1, \dots, \xi_k)$ k-dimenziós valószínűségi változó karakterisztikus függvényét. (Mivel a karakterisztikus függvényt a $(\xi_1, \dots, \xi_k)$ valószínűségi változó eloszlásfüggvényének a segítségével ki lehet számolni, ezért jogunk van egy eloszlásfüggvény karakterisztikus függvényéről beszélni.)

Tétel. *Legyen $F(x_1, \dots, x_k)$ és $G(x_1, \dots, x_k)$ két eloszlásfüggvéy, amelyek karakterisztikus függvényei megegyeznek. Ekkor $F(x_1, \dots, x_k) = G(x_1, \dots, x_k)$ minden k-dimenziós $(x_1, \dots, x_k)$ vektorra.*

Megfogalmazom az eloszlásfüggvények és azok karakterisztikus függvényei konvergenciája közötti kapcsolatot leíró Alaptételnek nevezett állítás több-dimenziós változatát.

Az eloszlások konvergenciájáról szóló Alaptétel több-dimenziós változata. *Legyen $F_n(u_1, \dots, u_k)$, $-\infty < u_j < \infty$, $1 \le j \le k$, k-dimenziós eloszlásfüggvények egy sorozata $\varphi_n(t_1, \dots, t_k) = \int e^{i(t_1u_1 + \dots + t_ku_k)} F_n(\,du_1, \dots, \,du_k)$ karakterisztikus függvényekkel, $n = 1, 2, \dots$. Ha a $\varphi_0(t_1, \dots, t_k) = \lim\limits_{n\to\infty} \varphi_n(t_1, \dots, t_k)$ határérték létezik minden $-\infty < t_j < \infty$, $1 \le j \le k$, számra, és a $\varphi_0(t_1, \dots, t_k)$ limeszfüggvény folytonos az origóban, akkor létezik olyan $F_0(u_1, \dots, u_k)$ eloszlásfüggvény a k-dimenziós térben, amelynek a $\varphi_0(t_1, \dots, t_k)$ függvény a karakterisztikus függvénye. Továbbá e feltétel teljesülése esetén az $F_n(u_1, \dots, u_k)$ eloszlásfüggvények eloszlásban konvergálnak ehhez az $F_0(u_1, \dots, u_k)$ eloszlásfüggvényhez.*

Megfordítva, ha $F_n(u_1, \dots, u_k)$, $n = 1, 2, \dots$, eloszlásfüggvényeknek egy olyan sorozata, amely egy $F_0(u_1, \dots, u_k)$ eloszlásfüggvényhez konvergál eloszlásban, $\varphi_n(t_1, \dots, t_k)$, $n = 1, 2, \dots$, jelöli az $F_n(u_1, \dots, u_k)$, $\varphi_0(t_1, \dots, t_k)$ pedig az $F_0(u_1, \dots, u_k)$ eloszlásfüggvény karakterisztikus függvényét, akkor $\varphi_0(t_1, \dots, t_k) = \lim\limits_{n\to\infty} \varphi_n(t_1, \dots, t_k)$ minden

$-\infty < t_j < \infty$, $1 \le j \le k$, számra. Továbbá, ez a konvergencia egyenletes minden véges intervallumban.

Ahhoz, hogy a fenti eredményeket használni tudjuk a több-dimenziós centrális határeloszlástétel bizonyításában ismernünk kell a több-dimenziós normális eloszlások karakterisztikus függvényeit. Erről szól a következő eredmény.

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 a*

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

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:

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

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

Az η véletlen vektor $D = (d_{j,l})$ kovariancia mátrixában a j-ik sor l-ik eleme $d_{j,l} = \mathrm{Cov}\,(\eta_j, \eta_l) = \mathrm{Cov}\left(\sum\limits_{p=1}^k a_{p,j}\xi_p, \sum\limits_{q=1}^k a_{q,l}\xi_q\right) = \sum\limits_{p=1}^k a_{p,j}a_{p,l}E\xi_p^2 = \sum\limits_{p=1}^k a_{p,j}a_{p,l}$, 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.*

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 karekterisztikus függvénye.

2. Kiegészítés. A χ-négyzet próbáról.

Az urnadobás eredményének aszimptotikus viselkedéséről szóló határeloszlástétel bizonyítása azon alapul, hogy egyrészt meg tudjuk adni a

$$\left\{\frac{\nu_n(j)-np_j}{\sqrt{np_j}},\ j=1,\dots,k\right\}$$

véletlen vektorok határeloszlását a több-dimenziós centrális határeloszlástétel segítségével. Másrészt alkalmazhatjuk az eloszlásbeli konvergencia tulajdonságairól kimondott 1. tételt. Ennek alapján tudjuk azt, hogy ha $\left(\eta_1^{(n)},\dots,\eta_k^{(n)}\right)$ véletlen vektorok eloszlásban konvergálnak egy $(\eta_1,\dots,\eta_k)$ véletlen vektorhoz $n\to\infty$ esetén, és ha tekintünk egy $f(x_1,\dots,x_k)$ folytonos k-változós függvényt, akkor az $f\left(\eta_1^{(n)},\dots,\eta_k^{(n)}\right)$, $n=1,2,\dots$, valószínűségi változók eloszlásban konvergálnak az $f(\eta_1,\dots,\eta_k)$ valószínűségi változóhoz. Ezen eredmény az f függvény alkalmas (természetes) megválasztásával lehetővé teszi a minket érdeklő statisztikák határeloszlásának megadását. De ezzel még nem fejezzük be a tétel bizonyítását. A kapott határeloszlást alkalmas, jól áttekinthető alakban akarjuk megadni. Ez, mint látni fogjuk, bizonyos lineáris algebrai meggondolások segítségével lehetséges.

Ezt azért tudjuk megtenni, mert mint azt később megmutatom, ha $(\eta_1,\dots,\eta_k)$ k-dimenziós nulla várható értékű D kovarianciamátrixú normális eloszlású véletlen vektor valamilyen ismert D kovariancia mátrix-szal, akkor bizonyos alapvető lineáris algebrai eredmények felhasználásával egyszerű formában megadhatjuk a $\sum\limits_{j=1}^k \eta_j^2$ valószínűségi változó eloszlását.

Annak érdekében, hogy az urnadobás eredményének aszimptotikus viselkedéséről szóló tételt bebizonyítsuk, először határozzuk meg a tételben szereplő $\frac{(\nu_n(j)-np_j)}{\sqrt{np_j}}$, $1\le j\le k$, véletlen vektorok határeloszlását $n\to\infty$ esetén a több-dimenziós centrális határeloszlástétel segítségével. Ennek érdekében vezessük be a következő

$$Z_m=(Z_m^{(1)},\dots,Z_m^{(k)}),\quad m=1,2,\dots,n,$$

véletlen vektorokat. A Z_m vektor $Z_m^{(j)}$ koordinátája 1, és az összes többi koordinátája nulla ha az m-ik dobásban az j-ik urnába esik a golyó. Vezessük be e véletlen vektorok következő transzformáltjait is: $X_m=(X_m^{(1)},\dots,X_m^{(k)})$, $m=1,2,\dots,n$, $X_m^{(j)}=\frac{Z_m^{(j)}-p_j}{\sqrt{p_j}}$, $1\le m\le n$, $1\le j\le k$. Vegyük észre, hogy a vizsgált $\nu_n(j)$ mennyiségek (a j-ik urnába eső golyók száma az n-ik dobás után) és a fenti X_m véletlen vektorok között a következő kapcsolat érvényes.

$$\frac{1}{\sqrt{n}}\sum_{m=1}^n X_m=\left(\frac{\nu_n(1)-np_1}{\sqrt{np_1}},\dots,\frac{\nu_n(k)-np_k}{\sqrt{np_k}}\right).$$

Továbbá az X_m vektorok függetlenek és egyforma eloszlásúak, várható érték vektoruk nulla, és kovariancia mátrixukat a $\operatorname{Cov}(X_m^{(j)},X_m^{(l)})=-\sqrt{p_jp_l}$, $1\le j,l\le k$, $j\ne l$, és

$\mathrm{Var}\, X_m^{(j)} = (1-p_j)$, $1 \le j,l \le k$, $1 \le m \le n$, képlet adja meg. Ugyanis egyszerű számolás segítségével kapjuk, hogy $EZ_m = (p_1,\dots,p_k)$, $\mathrm{Cov}\,(Z_m^{(j)}, Z_m^{(l)}) = -p_j p_l$, $1 \le j,l \le k$, $j \ne l$, és $\mathrm{Var}\, Z_m^{(j)} = p_j(1-p_j)$, $1 \le j,l \le k$, $1 \le m \le n$. Innen látszik, hogy $EX_m = 0$, és $\mathrm{Cov}\,(X_m^{(j)}, X_m^{(l)}) = \frac{\mathrm{Cov}\,(Z_m^{(j)}, Z_m^{(l)})}{\sqrt{p_j p_l}} = -\sqrt{p_j p_l}$, $\mathrm{Var}\, X_m^{(j)} = \frac{\mathrm{Var}\, Z_m^{(j)}}{p_j} = 1-p_j$. A több-dimenziós centrális határeloszlástétel szerint a $\left(\frac{\nu_n(1)-np_1}{\sqrt{np_1}},\dots,\frac{\nu_n(k)-np_k}{\sqrt{np_k}}\right)$ eloszlású véletlen vektorok eloszlásban konvergálnak egy nulla várható értékű és $D = (d_{i,j})$, $d_{i,j} = -\sqrt{p_i p_j}$, ha $i \ne j$, $d_{j,j} = 1-p_j$ kovariancia mátrixú $(\eta_1,\dots,\eta_k)$ normális vektorhoz. (A fenti érvelésből speciálisan az is következik, hogy létezik olyan normális eloszlású véletlen vektor, amelynek D a kovariancia mátrixa.)

Ezért alkalmazva az eloszlásbeli konvergencia tulajdonságairól szóló 1. tételt az $f(x_1,\dots,x_k) = \sum\limits_{j=1}^k x_j^2$ folytonos függvénnyel azt kapjuk, hogy a $\sum\limits_{j=1}^k \frac{(\nu_n(j)-np_j)^2}{np_j}$ valószínűségi változóknak van határeloszlásuk, és az megegyezik a $\sum\limits_{j=1}^k \eta_j^2$ valószínűségi változó eloszlásával, ahol $(\eta_1,\dots,\eta_k)$ normális eloszlású véletlen vektor nulla várható értékkel és D kovariancia mátrix-szal. Ezt a kifejezést felírhatjuk független standard normális eloszlású valószínűségi változók kvadratikus alakjában is. Ezt a kifejezést szeretnénk minél egyszerűbb, áttekinthetőbb alakban megadni. E célból hasznos belátni az alábbi lemmát.

Lemma. *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 $\sum\limits_{j=1}^k \eta_j^2$ valószínűségi változó eloszlása megegyezik egy $\sum\limits_{j=1}^k \lambda_j \xi_j^2$ 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, melynek 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. Ezért az η és $\bar\eta$ véletlen vektorok eloszlása megegyezik. Innen az is következik, hogy a $\sum\limits_{j=1}^k \eta_j^2$ valószínűségi változó eloszlása megegyezik a $\sum\limits_{j=1}^k \bar\eta_j^2$ 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 $\sum\limits_{j=1}^{k}\bar{\eta}_j^2=\sum\limits_{j=1}^{k}\tilde{\eta}_j^2$ 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 $\sum\limits_{j=1}^{k}\eta_j^2$ valószínűségi változó eloszlása megegyezik a $\sum\limits_{j=1}^{k}(\lambda_j^{1/2}\xi_j)^2=\sum\limits_{j=1}^{k}\lambda_j\xi_j^2$ valószínűségi változó eloszlásával, és ez a Lemma állítása.

Az urnadobás eredményének aszimptotikus viselkedéséről szóló tétel bizonyításának a befejezése. Tekintsük a $\sum\limits_{j=1}^{k}\eta_j^2$ valószínűségi változót, ahol $(\eta_1,\dots,\eta_k)$ k-dimenziós normális eloszlású valószínűségi változó, amelynek várható értéke nulla, és $D=d_{j,l}$, $d_{j,l}=\operatorname{Cov}(\eta_j,\eta_l)$, $1\le j,l\le k$, kovariancia mátrixát a $\operatorname{Var}\eta_j=(1-p_j)$, $1\le j,l\le k$, és $\operatorname{Cov}(\eta_j,\eta_l)=-\sqrt{p_jp_l}$, ha $1\le j,l\le k$, és $j\neq l$ képletek határozzák meg. Az előző lemma alapján azt kell megmutatni, hogy a fenti D kovariancia mátrixnak az 1 $k-1$ multiplicitású sajátértéke (azaz $k-1$ 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, $B=(b_{i,j})$, $b_{i,j}=\sqrt{p_i}\sqrt{p_j}$, $1\le i,j\le k$, és vegyük észre, hogy amennyiben egy B mátrixnak a sajátvektorai $e_1,\dots,e_k$ vektorok, $\lambda_1,\dots,\lambda_k$ sajátértékkel, akkor tetszőleges c számra az $I+cB$ mátrix sajátvektorai ugyanazok az $e_1,\dots,e_k$ vektorok $1+c\lambda_1,\dots,1+c\lambda_k$ sajátértékkel. Ezt az eredményt alkalmazva $c=-1$ választással elegendő megtalálni a B mátrix sajátértékeit. Viszont egy $B=(b_ib_j)$, $1\le i,j\le k$ alakú mátrix egyik sajátvektora a $b=(b_1,\dots,b_k)$ vektor $\sum\limits_{j=1}^{k}b_j^2$ sajátértékkel, és a b vektort ortogonálisan kiegészítő altér a B mátrix $k-1$-dimenziós saját altere nulla sajátértékkel. Ez azt jelenti, hogy a nulla a B mátrix $k-1$ multiplicitású sajátértéke, mert B $k-1$-dimenziós nulla sajátértékű sajátalterének egy ortonormált bázisa $k-1$ ortonormált sajátvektort biztosít nulla sajátértékkel.

Az előbb megfogalmazott állítások egyszerűen ellenőrizhetőek. Valóban egyszerű számolás mutatja, hogy a $b=(b_1,\dots,b_k)$ vektorra $bB=\left(\sum\limits_{j=1}^{k}b_j^2\right)b$, és ha valamely $c=(c_1,\dots,c_k)$ vektorra $\sum\limits_{j=1}^{k}c_jb_j=0$, akkor $cB=0$, mert ebben az esetben a cB vektor l-ik koordinátája $b_l\sum\limits_{j=1}^{k}c_jb_j=0$. Ez azt jelenti a b vektor ortogonális kiegészítő alterének elemei a B mátrix sajátvekorai nulla sajátértékkel.

Jelen esetben $D=I-B$ alakú mátrixot tekintettünk, ahol a B mátrix a fent vizsgált alakú a $b_j=\sqrt{p_j}$ számokkal. Ezért $\sum\limits_{j=1}^{k}b_j^2=\sum\limits_{j=1}^{k}p_j=1$, és a B vektornak a nulla $k-1$-szeres az 1 pedig egyszeres multiplicitású sajátértéke. Ez azt jelenti, hogy a $D=I-B$ mátrixnak a nulla egyszeres az 1 pedig $k-1$-szeres multiplicitású gyöke, mint állítottuk.