A Valószínűségszámítás I. előadássorozat tizenharmadik előadása.
2003. május 6.

A több-dimenziós centrális határeloszlástétel

Lássunk előbb néhány problémát, melyek megmutatják milyen kérdéseket kívánunk vizsgálni.

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 $\left[-\frac{1}{2},\frac{1}{2}\right]$ 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? Erre a kérdésre a több-dimenziós centrális határeloszlástétel ad pozitív választ.

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. (Többdimenziós valószínűségi változók függetlenségének a fogalmát ismertettem a 9. előadás elején, és az megtalálható az előadás ismertetésének első 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. Be fogjuk látni, hogy ez lehetséges. A határeloszlástételben megjelenő határeloszlásokat fogjuk több-dimenziós normális eloszlásnak nevezni.

Lássuk, hogyan lehet tárgyalni a fent 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,

$1 \leq j \leq n$, a következő valószínűségi 6-dimenziós változó: A ξ_j valószínűségi változó l-ik koordinátája 1, ha a j-ik dobás értéke l, $1 \leq l \leq 6$, és ξ_j összes többi koordinátája nulla. Ekkor a ξ_j valószínűségi változók függetlenek, (az egyes valószínűségi változók koordinátái nem azok), és az $S_n = \sum\limits_{j=1}^{n} \xi_j$ összeg l-ik koordinátája egyenlő az l értékű dobások számával minden $1 \leq l \leq 6$ 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 $\eta_j = (\xi_j, \xi_j^2)$, $1 \leq j \leq n$, 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.

Fő célunk annak az eredménynek az ismertetése és megértése, amelyik a centrális határoszlástétel természetes megfelelője abban az esetben, ha független (és mondjuk egyforma eloszlású véletlen vektorok alkalmasan normált összegének a viselkedését vizsgáljuk. Annak érdekében, hogy ezt megtehessük meg kell találnunk a határeloszlásban megjelenő normális eloszlás több-dimenziós megfelelőjét, amit több-dimenziós normális eloszlásnak fogunk nevezni. De mielőtt ezt megtesszük szükségünk van van az egy-dimenziós esetben szereplő várható érték és szórásnégyzet több-dimenziós megfelelőjét. Ezenkívül meg akarjuk érteni, hogy a várható érték és szórásnégyzet tulajdonságai, hogyan öröklődnek, ha azokanak a több-dimenziós esetben megjelenő megfelelőit vizsgáljuk. Az egy-dimenziós esetben megjelenő 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 $EZ_j^2 < \infty$, $1 \leq j \leq k$, feltételt. E véletlen vektor várható értéke az $EZ = (EZ_1, \dots, EZ_k)$ k-dimenziós vektor, kovariancia mátrixa pedig az a $D = (d_{j,l})$, $1 \leq j, l \leq k$, $k \times k$ 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.*

Fogalmazzuk meg a vektorértékű valószínűségi változók várható értékének és kovariancia mátrixának néhány fontos tulajdonságát. Az alábbi tétel formájában megfogalmazott állítás bizonyítása a gyakorlaton tárgyalandó feladat lesz.

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

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

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 $a^2 D$ 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.

Feladat: Bizonyítsuk be a fent megfogalmazott tételt.

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, Viszont 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 a $xDx^* = \sum\limits_{j=1}^{k} \sum\limits_{l=1}^{k} x_j d_{j,l} x_l = 0$ 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) megadom a kiegészítésben. 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 $\mathrm{Var}\, Y_j = 1$*,* $1 \le j \le k$*,* $\mathrm{Cov}\,(Y_j, Y_l) = 0$*, ha* $1 \le j, l \le k$*, és* $j \ne l$*. (Ez a helyzet például akkor, ha az* Y_j*,* $1 \le j \le k$ *valószínűségi változók függetlenek, és* $\mathrm{Var}\, Y_j = 1$*.) 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* $x_1 = \dots = x_k = 0$*.*

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 pozitúv 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 lineűáaris 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 négyzetgyók nem egyértelműen meghatározott, mint ahogy a valós számok között is csak akkor egyértelmű a gyökvonás, ha csak a pozitív gyököt tekintjük. leírni.

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

$$\mathrm{Var}\left(\sum_{j=1}^{k} x_j Z_j \right) = \sum_{j=1}^{k} \sum_{l=1}^{k} E\left(x_j x_l (Z_j Z_l - E Z_j E Z_l) \right) = \sum_{j=1}^{k} \sum_{l=1}^{k} x_j x_l \mathrm{Cov}\,(Z_j, Z_l)$$

$$= \sum_{j=1}^{k} \sum_{l=1}^{k} x_j x_l d_{j,l} = xDx^*,$$

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 $f(u_1, \dots, u_k) = \dfrac{1}{(2\pi)^{k/2}} \exp\left\{-\dfrac{1}{2} \sum\limits_{j=1}^{k} u_j^2\right\}$ 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.

Később meg fogjuk adni a több-dimenziós normális eloszlás más ekvivalens jellemzését. Először azonban megfogalmazunk egy a több-dimenziós centrális eloszlástétel kimondásához szükséges később bizonyítandó eredményt.

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. Továbbá egy k-dimenziós normális eloszlású valószínűségi változó eloszlását meghatározza annak m várható értéke 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. Ennek további fontos következményei vannak, amelyeket később tárgyalni fogunk.

A több-dimenziós normális eloszlás tulajdonságairól szóló valamint a kovariancia mátrixok jellemzéséről szóló tételekből következik, hogy bármely $(\xi_1,\dots,\xi_k)$ k-dimenziós valószínűségi változó esetén létezik olyan k-dimenziós normális eloszlású valószínűségi változó, amelynek kovariancia mátrixa és várható értéke megegyezik ennek a $(\xi_1,\dots,\xi_k)$ valószínűségi változónak a kovariancia mátrixával és várható értékével. Továbbá ennek a k-dimenziós normális eloszlású valószínűségi változónak az eloszlását meghatározza a $(\xi_1,\dots,\xi_k)$ valószínűségi változó kovariancia mátrixa és várható értéke. Erre az észrevételre szükségünk van ahhoz, hogy lássuk: Az alább megfogalmazott több-dimenziós centrális határeloszlás értelmes állítás.

A több-dimenziós 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.

Megjegyezzük, hogy az egydimenziós esethez hasonlóan, a több-dimenziós 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 több-dimenziós centrális határeloszlástétel bizonyítása hasonló a klasszikus egydimenziós 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. Külön érdemes hangsúlyozni, hogy a vizsgálatok során bevezetjük a több-dimenziós karakterisztikus függvény fogalmát, és a több-dimenziós karakterisztikus függvények vizsgálata nagy segítséget jelent a több-dimenziós normális eloszlások viselkedésének megértésében is.

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: Mivel minden több-dimenziós normális eloszlás eloszlásfüggvénye folytonos, (bár nem feltétlenül van sűrűségfüggvénye) ezért a több-dimenziós centrális határeloszlástétel azt mondja ki, hogy az ott tekintett normalizált összegek eloszlásban konvergálnak egy nulla várható értékű és megfelelő kovariancia mátrixú normális eloszláshoz. Jegyezzük meg, hogy a tekintett normalizált összegek várható értéke és kovarianciamátrixa megegyezik egy a határeloszlással rendelkező véletlen vektor várható értékével és kovariancia mátrixával.

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*

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

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.*

Megfogalmazzuk 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* $\varphi_n(t_1,\dots,t_k)=\int e^{i(t_1u_1+\cdots+t_ku_k)}F_n(\,du_1,\dots,\,du_k)$ *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 $d_{j,l} = \operatorname{Cov}(\eta_j, \eta_l) = \operatorname{Cov}\left(\sum\limits_{p=1}^k a_{p,j}\xi_p, \sum\limits_{q=1}^k a_{q,l}\xi_q\right) = \sum\limits_{p=1}^k a_{p,j}a_{p,l}E\xi_p^2 = \sum\limits_{p=1}^k a_{p,j}a_{p,l}$, 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.

Annak, hogy a több-dimenziós normális eloszlásfüggvény karakterisztikus függvényének ilyen az alakja van néhány mind a valószínűségszámításban mind a matematikai statisztikában nagyon fontos következménye. Ennek kifejtése az utolsó előadás témája lesz.

Végül leírjuk egy több-dimenziós normális eloszlású valószínűségi változó sűrűségfüggvényét, illetve megadjuk, hogy mikor létezik ilyen sűrűségfüggvény.

Tétel a több-dimenziós normális eloszlásfüggvények 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 értéke és* $D = (d_{j,l})$*,* $d_{j,l} = \operatorname{Cov}(\eta_j, \eta_l)$*,* $1 \le j, l \le k$*, 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 az 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 $\varphi(y) = \varphi(y_1, \dots, y_k) = \dfrac{1}{(2\pi)^{k/2}} e^{-(y_1^2 + \cdots + y_k^2)/2}$ 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 $\dfrac{1}{|\det A|}\varphi\left((x - m)A^{-1}\right)$ függvény. Annak érdekében, hogy bebizonyítsuk a tételt vegyük észre, hogy mivel $D = A^*A$, ezért $\det D = \det A^* \det A = \det A^2$, ezért $|\det A| = \det D^{1/2}$, és

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

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.