A Valószínűségszámítás I. előadássorozat tizennegyedik előadása.
2003. május 13.

A több-dimenziós centrális határeloszlástétel.

Részben felidézzük a múlt előadás bizonyos fogalmait és eredményeit részben megfogalmazzuk ezek néhány fontos következményét. Megjegyzem, hogy ezek a következmányek elsősorban a később tanulandó matematikai statisztikában játszanak fontos szerepet.

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 $f(u_1,\dots,u_k)=\dfrac{1}{(2\pi)^{k/2}}\exp\left\{-\dfrac12\sum\limits_{j=1}^k u_j^2\right\}$ 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.*

A teljesség kedvéért felidézem az eloszlásban való konvergenxia definicióját a több-dimenziós esetében.

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.*

Mint azt egy az előző előadásban is megfogalmazott eredmény állítja az egy-dimenziós esethez hasonlóan több-dimenziós eloszlások esetében is jellemezhető eloszlások konvergenciája úgy, hogy folytonos és korlátos függvények ezen eloszlások szerinti integráljai konvergálnak. Viszont a több-dimenziós eloszlások konvergenciájának az esetében felmerül egy olyan kérdés, amelyik az egy-dimenziós esetben nem játszik olyan fontos szerepet. Ugyanis a magasabb dimenziós esetben a tér halmazainak a geometriai struktúrája gazdagabb. Szeretnénk például tudni, hogy a centrális határeloszlástétel teljesüléséből következik-e az, hoyg annak valószínűsége, hogy az egyes véletlen vektorok egy adott szép halmazba, például egy gömbbe esik megegyezik-e ennek a halmaznak a határmérték szerinti mértéke szerint. Erre a kérdésre alkalmas feltételek mellett pozitív választ lehet adni, de ez külön bizonyításra szorul. Alább megfogalmazzuk a pontos eredményt, illetve egy másik szintén hasznos eredményt, de ezen eredmények bizonyítását csak a kiegészítésben fogom ismertetni.

Tétel A. *Konvergáljanak k-dimenziós $F_n(x_1,\dots,x_k)$, $n=1,2,\dots,$ eloszlásfüggvények valamely $F_0(x_1,\dots,x_k)$ eloszláshoz, ha $n\to\infty$, és jelölje μ_{F_n}, $n=0,1,2,\dots,$ ezen eloszlások által meghatározott Stieltjes mértéket. Ekkor minden olyan Borel-mérhető A halmazra, amelynek ∂A határa teljesíti a $\mu_{F_0}(\partial A)=0$, teljesül a $\lim\limits_{n\to\infty}\mu_{F_n}(A)=\mu_{F_0}(A)$ azonosság.*

Tétel B. *Legyen $(S_{1,n},\dots,S_{k,n})$, $n=1,2\dots,$ k-dimenziós valószínűségi vektorok sorozata, amelyik eloszlásban konvergál egy $(S_1,\dots,S_k)$ k-dimenziós véletlen vektorhoz $n\to\infty$ esetén, és legyen $f(x_1,\dots,x_k)$ egy k-változós folytonos függvény. Ekkor a $T_n=f(S_{1,n},\dots,S_{k,n})$, $n=1,2,\dots,$ valószínűségi változók eloszlásban konvergálnak a $T=f(S_1,\dots,S_k)$ valószínűségi változóhoz $n\to\infty$ esetén.*

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.

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.*

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.*

Megfogalmazzuk a fenti eredmények néhány fontos következményét. Először jellemezzük a több-dimenziós normális eloszlások karakterisztikus függvényeit.

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*

$$Ee^{i(t,\eta)}=Ee^{i(t_1\eta_1+\cdots+t_k\eta_k)}=e^{i(t,m)-tA^*At^*/2}=\exp\left\{i\sum_{j=1}^k t_jm_j-\frac12\sum_{j=1}^k\sum_{l=1}^k d_{j,l}t_jt_l\right\}$$

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 $Ee^{i(tA^*,\xi)} = Ee^{i(\bar{t}_1\xi_1+\cdots+\bar{t}_k\xi_k)} = \prod_{j=1}^k Ee^{i\bar{t}_j\xi_j} = \prod_{j=1}^k e^{-\bar{t}_j^2/2} = e^{-(\bar{t},\bar{t})/2} = e^{-tA^*At^*/2}$.

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

Az előző tételnek van néhány fontos következménye.

1. következmény. *Egy k-dimenziós normális eloszlást meghatároz várható érték vektora és kovariancia mátrixa.*

2. következmény. *Legyen (ξ, η) két-dimenziós normális eloszlású véletlen vektor. Ha a ξ és η valószínűségi változók korrelálatlanok, azaz* $\mathrm{Cov}\,(\xi, \eta) = 0$, *akkor ξ és η független valószínűségi változó. Általánosabban, legyen $(\xi_1, \dots, \xi_k)$ k-dimenziós normális eloszlású valószínűségi változó, amelyre igaz, hogy valamilyen j, indexre, $1 \le j < k$, az első j és utolsó $k - j$ koordináták korrelálatlanok, azaz* $\mathrm{Cov}\,(\xi_p, \xi_q) = 0$, *ha $1 \le p \le j < q \le k$. Ekkor a $(\xi_1, \dots, \xi_j)$ és $\xi_{j+1}, \dots, \xi_k)$ véletlen vektorok független j és $k - j$-dimenziós normális eloszlású valószínűségi változók.*

Fontos megjegyzés: Láttuk korábban, hogy független valószínűségi változók korrelálatlanok, viszont ennek az állításnak a megfordítása nem igaz, azaz léteznek olyan korrelálatlan valószínűségi változók, amelyek nem függetlenek. A fent megfogalmazott 2. következmény viszont azt állítja, hogy ha olyan valószínűségi változókat tekintünk, amelyek egy több-dimenziós normális eloszlású véletlen vektor koordinátáinak is felfoghatók, akkor e valószínűségi változók korrelálatlanságából következik azok függetlensége. Az a megszorítás, hogy a tekintett valószínűségi változók egy normális eloszlású véletlen vektor koordinátái nagyon szigorú megkötés. Viszont ez nagyon fontos speciális eset, amelyik fontos szerepet játszik mind a valószínűségszámításban mind a matematikai statisztikában.

3. következmény. *Egy $\xi = (\xi_1, \dots, \xi_k)$ k-dimenziós valószinűségi változó akkor és csak akkor normális eloszlású, ha karakterisztikus függvénye*

$$Ee^{i(t,\xi)} = Ee^{i(t_1\xi_1+\cdots+t_k\xi_k)} = e^{i(t,m)-tDt^*/2} = \exp\left\{ i\sum_{j=1}^k t_j m_j - \frac{1}{2}\sum_{j=1}^k\sum_{l=1}^k d_{j,l}t_jt_l \right\}$$

alakú függvény, ahol $m = (m_1, \dots, m_k)$ k-dimenziós (valós koordinátákból álló) vektor, $D = (d_{j,l})$, $1 \le j, l \le k$, $k \times k$ méretű pozitív szemidefinit mátrix, (x, y) jelöli az $x = (x_1, \dots, x_k)$ és $y = (y_1, \dots, y_k)$ $(x, y) = \sum_{j=1}^k x_jy_j$ skalárszorzatát, $t = (t_1, \dots, t_k)$, *tetszőleges k-dimenziós vektor. A normális eloszlású $\xi = (\xi_1, \dots, \xi_k)$ valószínűségi változó fenti jellemzésében az m vektor a ξ több-dimenziós valószínűségi változó várható értéke, D pedig a kovariancia mátrixa.*

Az első következmény következik abból, hogy egyrészt egy normális eloszlású valószínűségi változó karakterisztikus függvényét fel lehet írni annak várható érték vektora és kovariancia mátrixa ismeretében, másrészt egy eloszlást meghatároz annak karakterisztikus függvénye.

A második következmény indoklása hasonló. Tekintsük az abban megfogalmazott általánosabb állítás bizonyítását, és vezessük be a következő jelöléseket. Jelölje $m = (E\xi_1, \dots, E\xi_k)$ a $(\xi_1, \dots, \xi_k)$ véletlen vektor várható értékét, $D = (d_{p,q})$, $d_{p,q} = \operatorname{Cov}(\xi_p, \xi_q)$, $1 \le p, q \le k$, ennek a vektornak a kovariancia mátrixát. Legyen $m_1 = (E\xi_1, \dots, E\xi_j)$ a $(\xi_1, \dots, \xi_j)$ vektor várható értéke, $D_1 = (d_{p,q})$, $d_{p,q} = \operatorname{Cov}(\xi_p, \xi_q)$, $1 \le p, q \le j$, ennek a vektornak a kovariancia mátrixa, végül legyen $m_2 = (E\xi_{j+1}, \dots, E\xi_k)$ a $(\xi_{j+1}, \dots, \xi_k)$ vektor várható értéke, $D_2 = (d_{p,q})$, $d_{p,q} = \operatorname{Cov}(\xi_p, \xi_q)$, $j+1 \le p, q \le k$, ennek a vektornak a a kovariancia mátrixa. Ezenkívül tekintsünk egy tetszőleges $\mathbf{t} = (t_1, \dots, t_k)$ k-dimenziós vektort, és legyen $\mathbf{t}_1 = (t_1, \dots, t_j)$, $\mathbf{t}_2 = (t_{j+1}, \dots, t_k)$.

Ekkor a $(\xi_1, \dots, \xi_k)$ véletlen vektor karakterisztikus függvénye a $\varphi(t_1, \dots, t_k) = e^{i(\mathbf{t},m) - \mathbf{t}D\mathbf{t}^*/2}$ függvény. Az exponensben szereplő kifejezés felírható mint $i(\mathbf{t}, m) - \dfrac{\mathbf{t}D\mathbf{t}^*}{2} = i(\mathbf{t}_1, m_1) - \dfrac{\mathbf{t}_1 D_1 \mathbf{t}_1^*}{2} + i(\mathbf{t}_2, m_2) - \dfrac{\mathbf{t}_2 D_2 \mathbf{t}_2^*}{2}$, mert $\operatorname{Cov}(\xi_p, \xi_q) = 0$, ha $1 \le p \le j < q \le k$ vagy $1 \le q \le j < p \le k$. Ezért $\varphi(t_1, \dots, t_k) = \varphi_1(t_1, \dots, t_j)\varphi_2(t_{j+1}, \dots, t_k)$, ahol $\varphi_1(t_1, \dots, t_j) = e^{i(\mathbf{t}_1, m_1) - \mathbf{t}_1 D_1 \mathbf{t}_1^*/2}$, a $(\xi_1, \dots, \xi_j)$, $\varphi_2(t_{j+1}, \dots, t_k) = e^{i(\mathbf{t}_2, m_2) - \mathbf{t}_2 D_2 \mathbf{t}_1^*/2}$ pedig a $(\xi_{j+1}, \dots, \xi_k)$ valószínűségi változó karakterisztikus függvénye. Mivel a karakterisztikus függvény meghatározza az eloszlásfüggvényt ez azt jelenti, hogy a $(\xi_1, \dots, \xi_k)$ véletlen vektor eloszlásfüggvénye megegyezik két olyan független j illetve $k-j$-változós véletlen vektor együttes eloszlásával, amelyek közül az első véletlen vektor karakterisztikus függvénye $\varphi_1(t_1, \dots, t_j)$ a másodiké pedig $\varphi_2(t_{j+1}, \dots, t_k)$. Mivel $(\xi_1, \dots, \xi_j)$ véletlen vektor karakterisztikus függvénye $\varphi_1(t_1, \dots, t_j)$, a $(\xi_{j+1}, \dots, \xi_k)$ véletlen vektor karakterisztikus függvénye pedig $\varphi_2(t_{j+1}, \dots, t_k)$, innen következik a második következmény állítása.

A harmadik következmény azonnal következik a több-dimenziós normális eloszlású valószínűségi változók karakterisztikus függvényének jellemzését leíró tételből, abból a kimondott lineáris algebrai tételből, amely szerint minden pozitív szemidefinit D mátrix felírható $D = A^*A$ alakban, illetve abból a tényből, hogy egy A^*A alakú mátrix mindig pozitív szemidefinit. Ez utóbbi állítás ellenőrzéséhez vegyük észre, hogy tetszőleges $x = (x_1, \dots, x_k)$ vektorra $xA^*Ax^* = xA^*(xA^*)^* = (xA^*, xA^*) \ge 0$ és ezt az egyenlőtlenséget kellett megmutatni.

Az első következményből, illetve normális eloszlású valószínűségi változó kovarianca mátrixának korábban kiszámolt alakjából következik a korábban megfogalmazott tétel a több-dimenziós normális eloszlás tulajdonságairól. Annak érdekében, hogy lássuk fontos volt feltenni a második következményben, hogy normális eloszlású valószínűségi változókat vizsgálunk oldjuk meg a következő feladatot.

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 $\sum_{j=1}^{n} \xi_j$ és $\sum_{j=1}^{n} \xi_j^2$ összegek normalizáltjainak azaz az $\sqrt{\frac{12}{n}} \sum_{j=1}^{n} \xi_j$ és $\sqrt{\frac{180}{n}} \sum_{j=1}^{n} \left(\xi_j^2 - \frac{1}{12}\right)$ 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.

Feladat:

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

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

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

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

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

Az az eset, amikor $\xi = (\xi_1, \dots, \xi_s)$, $\eta = (\eta_1, \dots, \eta_p)$, és $(\xi_1, \dots, \xi_s, \eta_1, \dots, \eta_p)$ egy $s + p$ dimenziós normális eloszlású valószínűségi változó hasonlóan tárgyalható. Lássuk be, hogy létezik olyan $\mathbf{A}$ mátrix, amelyre η és $\xi - \eta\mathbf{A}$ függetlenek. Ennek érdekében lássuk be először, hogy létezik olyan $\mathbf{U}$ unitér mátrix amelyre $\eta\mathbf{U} = (\bar\eta_1, \dots, \bar\eta_p) = \bar\eta$ vektor koordinátái függetlenek. Ugyanis, ha az η véletlen vektor D kovarianciamátrixát $D = \mathbf{U}^*\Lambda\mathbf{U}$ alakban írjuk, ahol $\mathbf{U}$ unitér Λ pedig diagonális mátrix, akkor az $\bar\eta = \eta\mathbf{U}$ véletlen normális eloszlású vektor kovarianciamátrixa

Λ, ahonnan következik, hogy az $\bar{\eta}$ mátrix koordinátái függetlenek. Legyen $\bar{\xi}_r = \xi_r - \sum\limits_{k=1}^{p} \frac{E\xi_r\bar{\eta}_k}{E\bar{\eta}_k^2}\bar{\eta}_k$, $r = 1, \dots, s$. Ezt mátrixjelöléssel írjuk $\bar{\xi} = \xi - \bar{\eta}\mathbf{B}$ formában. Ekkor $(\xi - \bar{\eta}B, \bar{\eta})$ olyan $p + s$ dimenziós normális eloszlású valószínűségi változó, amelynek első s és utolsó p koordinátája korrelálatlan, ezért független. Mivel a $\xi - \bar{\eta}\mathbf{B}$ és $\bar{\eta}$ vektorok függetlenek, ezért $\zeta = \xi - \bar{\eta}\mathbf{B} = \xi - \eta\mathbf{UB}$ független az $\eta = \bar{\eta}\mathbf{U}^*$ vektortól.

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

A több-dimenziós centrális határeloszlástételt be lehet bizonyítani az egydimenziós valószínűségi centrális határeloszlástételhez hasonlóan az Alaptételnek nevezett állítás több-dimenziós változatának segítségével. Valójában van egyszerűbb megoldás is. Meg lehet mutatni, — szintén az Alaptétel több-dimenziós változata segítségével, — hogy a több-dimenziós határeloszlástételek következnek azok egy-dimenziós változatából. Ennek részleteit, amelynek tárgyalását bizonyos részletességgel tartalmazza az előző előadáshoz írt, de valójában nem elmondott ismertetés, itt elhagyjuk.

Egy több-dimenziós normális eloszlású valószínűségi változó koordinátái normális eloszlásúak. A következő feladatban példát mutatunk arra, hogy ennek az állításnak a megfordítása nem igaz.

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. Miért?

A χ^2 eloszlásokról és χ^2 próbáról.

Tekintsük a több-dimenziós normális eloszlások bevezetőjében tekintett első problémát. Hogyan tudjuk ellenőrizni, hogy egy dobókocka szabályos-e?

Dobjuk fel a dobókockát n alkalommal, és legyen

$$(\nu(1),\dots,\nu(6))=(\nu_n(1),\dots,\nu_n(6))$$

az $1,\dots,6$ dobáseredmények száma. Megmutatjuk a több-dimenziós centrális határeloszlás segítségével, hogy az $\frac{1}{n}\sum\limits_{j=1}^{6}\left(\nu_n(j)-\frac{n}{6}\right)^2$ valószínűségi változóknak van határeloszlásuk, ha $n\to\infty$, (feltéve, hogy egy szabályos dobókocka egymástól független dobásainak az eredményeit tekintjük), és ezt a határeloszlást explicit módon meg tudjuk adni. Ez lehetőséget ad egy kocka szabályosságának ellenőrzésére. Ennek a feladatnak a következő természetes általánosítását fogjuk tekinteni:

Legyen adva k urna, és ellenőrizni akarjuk azt a feltételezést, amely szerint ha egy golyót véletlenül bedobunk ezen urnák valamelyikébe, akkor az p_j, $p_j>0$, valószínűséggel esik a j-ik urnába, $\sum\limits_{j=1}^{k}p_j=1$. Ennek a feltételezésnek az ellenőrzése 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 feltételezésünk teljesülése esetén a $\sum\limits_{j=1}^{k}\frac{(\nu_n(j)-np_j)^2}{np_j}$ valószínűségi változóknak létezik határeloszlása $n\to\infty$ esetén, és ez a határeloszlás az úgynevezett $k-1$ szabadságfokú χ^2 eloszlás. Megfogalmazzuk ezt az eredményt pontosabban, előtte azonban megadjuk 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 10. előadáson tárgyalt feladatok egyikében, hogy a 2 szabadságfokú $\chi^2(2)$ eloszlás a $\lambda=\frac{1}{2}$ paraméterű exponenciális eloszlás, azaz az az eloszlás, amelynek sűrűségfüggvénye az $f(x)=\frac{1}{2}e^{-x/2}$, ha $x\ge0$, és $f(x)=0$, ha $x<0$.

Ezután megfogalmazzuk a χ^2 próba alapjául szolgáló eredményt. A bizonítást viszont csak a kiegész'ítésben adom meg, ami azt jelenti, hogy ez nem kötelező része a vizsgának.

Tétel. *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.)

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 vettünk, 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.

Kiegészítés.

A 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 belátjuk, hogy ha $\left(\eta_1^{(n)}, \dots, \eta_k^{(n)}\right)$ véletlen vektorok eloszlásban konvergálnak egy $(\eta_1, \dots, \eta_k)$ véletlen vektorhoz $n \to \infty$ esetében, és $f(x_1, \dots, x_k)$ folytonos k-változós függvény, 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. Ez lehetővé teszi a minket érdeklő statisztikák határeloszlásának megadását. Ezt a határeloszlást egyszerű, jól áttekinthető alakban akarjuk megadni. Ezt azért tudjuk megtenni, mert mint azt egy megoldásával együtt ismertetett feladatban megfogalmazom, 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.

Először megfogalmazzuk és bebizonyítjuk a fent említett állítást. Ez a fő részben is megfogalmazott Tétel B.

Tétel B. *Legyen* $(S_{1,n},\dots,S_{k,n})$, $n=1,2\dots$, *k-dimenziós valószínűségi vektorok sorozata, amelyik eloszlásban konvergál egy* $(S_1,\dots,S_k)$ *k-dimenziós véletlen vektorhoz* $n\to\infty$ *esetén, és legyen* $f(x_1,\dots,x_k)$ *egy k-változós folytonos függvény. Ekkor a* $T_n=f(S_{1,n},\dots,S_{k,n})$, $n=1,2,\dots$, *valószínűségi változók eloszlásban konvergálnak a* $T=f(S_1,\dots,S_k)$ *valószínűségi változóhoz* $n\to\infty$ *esetén.*

Bizonyítás: Használjuk a tételt az eloszlásban való konvergencia jellemzéséről folytonos függvények segítségével. Eszerint az $S_{1,n},\dots,S_{k,n}$ véletlen vektorok sorozatának eloszlásban való konvergenciát úgy is megfogalmazhatjuk, hogy tetszőleges folytonos és korlátos $h(x_1,\dots,x_k)$ függvényre teljesül a $\lim\limits_{n\to\infty} Eh(S_{1,n},\dots,S_{k,n})=Eh(S_1,\dots,S_k)$ reláció. Legyen $g(\cdot)$ folytonos és korlátos függvény. Ekkor $g(f(x_1,\dots,x_k))$ folytonos és korlátos függvény, ezért teljesül a

$$\lim_{n\to\infty} Eg(T_n)=\lim_{n\to\infty} Eg\left(f\left(S_{1,n},\dots,S_{k,n}\right)\right)=Eg\left(f\left(S_1,\dots,S_k\right)\right)=Eg(T)$$

reláció. Innen következik a Tétel B állítása.

Most megfogalmazom, hogyan lehet redukálni a a χ^2 próbáról szóló tételt.

A χ^2 próbáról szóló tétel redukciója. *Legyen* $\eta=(\eta_1,\dots,\eta_k)$ *k-dimenziós normális eloszlású véletlen vektor, melynek várható értéke és kovriancia mátrixa teljesíti az* $E\eta_j=0$, $1\le j\le k$, $\operatorname{Cov}(\eta_j,\eta_l)=-\sqrt{p_jp_l}$, $1\le j,l\le k$, $j\ne l$, *és* $\operatorname{Var}\eta_j=(1-p_j)$, $1\le j\le k$ *relációkat. Akkor* $\sum\limits_{j=1}^k \eta_j^2$ *eloszlása a* $k-1$ *szabadságfokú* $\chi^2(k-1)$ *eloszlás. (A most megfogalmazott eredményben speciálisan azt is állítjuk, hogy létezik az adott kovarianciával rendelkező véletlen vektor tetszőleges* p_j, $p_j\ge0$, $1\le j\le k$, $\sum\limits_{j=1}^k p_j=1$ *számokra.)*

A χ^2 próbáról szóló tétel visszavezetése annak redukciójára. Ahogy az előadás elején tetük, vezessük be a $Z_m=(Z_m^{(1)},\dots,Z_m^{(k)})$, $m=1,2,\dots,n$, véletlen vektorokat, melyeket a következő módon definiálunk: 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 ennek a következő transzformáltjait: $X_m=(X_m^{(1)},\dots,X_m^{(k)})$, $m=1,2,\dots,n$, $X_m^{(j)}=\dfrac{Z_m^{(j)}-p_j}{\sqrt{p_j}}$, $1\le m\le n$, $1\le j\le k$. Vegyük észre egyrészt, 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 az alábbi azonosságot teljesítik: $\dfrac1{\sqrt n}\sum\limits_{l=1} X_m=\left(\dfrac{\nu_n(1)-np_1}{\sqrt{np_1}},\dots,\dfrac{\nu_n(k)-np_k}{\sqrt{np_k}}\right)$. Továbbá az X_m vektorok függetlenek és egyforma eloszlásúak, várható érték vektoruk nulla, és kovariancia mátrixuk teljesíti a $\operatorname{Cov}(X_m^{(j)},X_m^{(l)})=-\sqrt{p_jp_l}$, $1\le j,l\le k$, $j\ne l$, és $\operatorname{Var}X_m^{(j)}=(1-p_j)$, $1\le j,l\le k$, $1\le m\le n$ relácókat. Ugyanis egyszerű számolás adja, hogy $EZ_m=(p_1,\dots,p_k)$, $\operatorname{Cov}(Z_m^{(j)},Z_m^{(l)})=-p_jp_l$, $1\le j,l\le k$,

$j \neq l$, és $\operatorname{Var} Z_m^{(j)} = p_j(1-p_j)$, $1 \le j,l \le k$, $1 \le m \le n$. Innen látszik, hogy $EX_m = 0$, és $\operatorname{Cov}(X_m^{(j)}, X_m^{(l)}) = \dfrac{\operatorname{Cov}(Z_m^{(j)}, Z_m^{(l)})}{\sqrt{p_j p_l}} = -\sqrt{p_j p_l}$, $\operatorname{Var} X_m^{(j)} = \dfrac{\operatorname{Var} Z_m^{(j)}}{p_j} = 1-p_j$. Innen következik speciálisan az is, hogy a χ^2 próbáról szóló tétel redukciójában szereplő kovarianca mátrixú (nulla várható értékú) (normális eloszlású) $(\eta_1,\dots,\eta_k)$ véletlen vektor valóban létezik. Továbbá, a több-dimenziós centrális határeloszlástétel szerint a $\left(\dfrac{\nu_n(1)-np_1}{\sqrt{np_1}},\dots,\dfrac{\nu_n(k)-np_k}{\sqrt{np_k}}\right)$ eloszlású véletlen vektorok eloszlásban konvergálnak egy ilyen$(\eta_1,\dots,\eta_k)$ normális vektorhoz. Ezért alkalmazva az előző tételt az $f(x_1,\dots,x_k) = \sum\limits_{j=1}^k x_j^2$ folytonos függvénnyel kapjuk, hogy a $\sum\limits_{j=1}^k \dfrac{(\nu_n(j)-np_j)^2}{np_j}$ valószínűségi változóknak van határeloszlásuk, és az megegyezik a $\sum\limits_{j=1}^k \eta_j^2$ valószínűségi változó eloszlásával. Ez azt jelenti, hogy a χ^2 próbáról szóló tétel bizonyításához elég annak redukált változatát belátni.

A fenti eredmény azt jelenti, hogy egy normális eloszlású véletlen vektor koordinátáiból képezett kvadratikus forma eloszlását kell kiszámolnunk. Emlékezzünk arra, hogy egy normális vektor lineáris transzformáltja is normális eloszlású, és tetszőleges normális eloszlású vektor előáll, mint standard normális eloszlású véletlen vektor lineáris transzformáltja. Ez azt sugallja, hogy próbáljuk meg a vizsgált (vagy egy vele azonos eloszlású) kifejezést előállítani, mint alkalmas független standard normális eloszláaú valószínűségi változók egyszerű kvadratikus formáját. A χ^2 próbáról szóló tétel redukciójában eredetileg megadott kifejezést azért kívánjuk átalakítani, mert igaz ugyan, hogy normális valószínűségi változók nagyon egyszerű kvadratikus formájáról van szó, viszont a kvadratikus formában szereplő változók nem függetlenek. Viszont az alábbi lemma nagyon hasznos lesz a számunkra.

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 nátrix 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. Innek 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.

Most befejezzük a χ^2 próbáról szóló tétel bizonyítását.

A χ^2 próbáról szóló tétel redukciójának a bizonyítása. Tekintsük a $\sum\limits_{j=1}^{k} \eta_j^2$ valószínűségi változót, ahol $(\eta_1, \dots, \eta_k)$ k-dimenziós normális eloszlású valószínűségi változó, melynek 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 $\operatorname{Cov}(\eta_j, \eta_l) = -\sqrt{p_j p_l}$, ha $1 \le j,l \le k$, és $j \ne l$ 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 $B = (b_i b_j)$, $1 \le i,j \le k$ 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 $\sum\limits_{j=1}^{k} c_j b_j = 0$, akkor $cB = 0$, mert e vektor l-ik koordinátája

$b_l \sum_{j=1}^{k} c_j b_j = 0$, és ez jelenti a b vektor ortogonális kiegészítő alterére megfogalmazott állítást.

Jelen esetben a tekintett B mátrix a fent vizsgált alakú a $b_j = \sqrt{p_j}$ számokkal. Ezért $\sum_{j=1}^{k} b_j^2 = \sum_{j=1}^{k} p_j = 1$, 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.

Végül ismertetem a fő részben megfogalmazott Tétel A bizonyítását.

A Tétel A bizonyítása. A bizonyításban az eredmény állítását visszavezetjük a Tétel B állítására. Ebben felhasználjuk az euklideszi tér következő fontos topológiai tulajdonságát. Legyen adva az R^k k-dimenziós euklideszi tér két F_1 és F_2 diszjunkt zárt részhalmaza. Ekkor létezik egyan f folytonos függvény az R^k téren, amelyre $f(x) = 0$, ha $f(x) \in F_1$, $f(x) = 1$, ha $f(x) \in F_2$, és $0 \le f(x) \le 1$ minden $x \in R^k$ szánra. (Ilyen függvény például az $f(x) = \dfrac{\rho(x, F_1)}{\rho(x, F_1) + \rho(x, F_2)}$ függvény, ahaol $\rho(x, F)$ jelöli az x pont és F halmaz távolságát.

Adott $\delta > 0$ számra jelölje $A_\delta = \{x\colon \rho(x, A) < \delta$ az A halmaz δ sugarú külsejét, $A_\delta^0 = \{x\colon \rho(x, R^k \setminus A) le\delta$ az A halmaz δ sugarú belsejét, $\bar{A} = A \cup \partial A$, A lezártját, $A^0 = A \setminus \partial A$, A belsejét. Ekkor $\bar{A} = \bigcap_{\delta>0} A_\delta$, $A^0 = \bigcup_{\delta>0} A_\delta^0$. Innen a valószínűségi mérték folytonosságából, illetve abból a feltételből, hogy $\mu_{F_0}(\partial A) = 0$, következik, hogy $\lim_{\delta\to 0} \mu_{F_0}(A_\delta^0) = \mu_{F_0}(A^0) = \mu_{F_0}(A) = \mu_{F_0}(\bar{A}) = \lim_{\delta\to 0} \mu_{F_0}(A_\delta)$. Ezért minden $\varepsilon > 0$ számhoz, létezik olyan $= \delta(\varepsilon) > 0$ szám, melyre $\mu_{F_0}(A) - \varepsilon \le \mu_{F_0}(A_\delta^0) \le \mu_{F_0}(A) \le \mu_{F_0}(A_\delta) \le \mu_{F_0}(A) + \varepsilon$.

Válasszunk egy ilyen δ számot valamint ezután olyan f folytonos függvényt az R^k téren, amelyre $f(x) = 1$, ha $x \in \bar{A}$, $f(x) = 0$, ha $x \in R^k \setminus A_delta$, és $0 \le f(x) \le 1$ minden $x \in R^k$ halmazra. (Mivel $\bar{A}$ és $R^k \setminus A_\delta$ zárt halmazok, ezért ez lehetséges.) Ekkor a Tétel B alapján

$$\limsup_{n\to\infty} \mu_{F_n}(A) \le \limsup_{n\to\infty} \int f(x)\mu_{F_n}(\,dx) = \int f(x)\mu_{F_0}(\,dx) \le \mu_{F_0}(A_\delta) \le \mu_{F_0}(A) + \varepsilon.$$

Mivel ez az állítás minden $\varepsilon > 0$ számra elmondható, azt kapjuk, hogy $\limsup_{n\to\infty} \mu_{F_n}(A) \le \mu_{F_0}(A)$.

Válasszunk ezután olyan f függvényt, amelyre $f(x) = 1$, ha $x \in A_\delta^0$, $f(x) = 0$, ha $x \in R^k \setminus A_0$, és $0 \le f(x) \le 1$ minden $x \in R^k$ pontra. (Ez lehetséges, mert A_δ^0 és $R^k \setminus A_0$ zárt halmazok.) Ezzel a választással az előzőhöz hasonló érveléssel kapjuk, hogy $\liminf_{n\to\infty} \mu_{F_n}(A) \ge \mu_{F_0}(A)$. A bebizonyított egyenlőtlenségekből következik a Tétel A állítása.

A valószínűségszámítás bevezető előadás rövid áttekintése

Megfogalmaztunk néhány valószínűségi problémát, azután rátértünk a valószínűségszámítás precíz Kolmogorov által megadott ismertetésére. Különösen fontos volt annak megértése, hogyan lehet a valószínűségszámítás problémáit e formális, a halmazelmélet nyelvét használó modellbe átfogalmazni. Ugyancsak fontos volt annak megértése, hogy ez az átfogalmazás hogyan segít konkrét feladatok megoldásában. A Kolmogorov-féle elmélet kidolgozásában, annak érdekében hogy ne csak a véges sok lehetséges eseményt tartalmazó modellt tudjuk vizsgálni, szükség volt bizonyos mély mértékelméleti eredmények felidézésére. Elegendő a fogalmak és eredmények jelentésének megértése, illetve azt látni, hogyan lehet ezt feladatok megoldásában használni (geometriai eloszlások), a bizonyítások részleteinek ismerete nem szükséges.

Tárgyaltuk és bizonyítottuk a Borel–Cantelli lemmát. Itt is elegendő az eredmény értése és az hogy tudjuk használni. A következő fontos fogalom a feltételes valószínűség volt. Nagyon fontos megérteni, hogy hogyan kell ezt használni. Az itt kimondott eredményeket jobban lehet megérteni, ha a konkrét feladatokat vizsgáljuk. Definiáltuk események függetlenségét. Az itt kimondott fogalmak, eredmények ismerete rendkívül fontos. A tárgyalás során szóba kerültek mértékelméleti fogalmak és eredmények. Az egyszerűbb fogalmakat, mint valószínűségi mező, σ-algebra ismerni kell, de már olyan fogalmaknak az ismeretét, mint például a majdnem-mindenütt teljesülő tulajdonságok, Lebesgue integrál, Fubini tétel elegendő körülbelül tudni, azon a szinten, ami ahhoz kell, hogy a leírt valószínűségi eredményeket értsék és tudják használni. (Az, hogy milyen szigorúsággal követelik meg ezen eredmények ismeretét a mértékelmélet vizsgán nem az én problémám.)

Ezután egyrészt konkrét előbb diszkrét, később folytonos eloszlású valószínűségi változókat vizsgáltunk. Definiáltuk valószínűségi változók függetlenségét, várható értékét, szórásnégyzetét és különböző valószínűségi változók kovarianciafüggvényét. Ezeket először diszkrét valószínűségi változókra definiáltuk ezeket a fogalmakat, és csak később tárgyaltuk folytonos esetben. Ez utóbbi definició kidolgozása érdekében vázoltuk azt, hogyan kell definiálni a Lebesgue integrált. Ez utóbbi fogalom pontos ismerete (ezen tantárgy keretein belül) nem szükséges. Viszont tudni kell a várható érték, szórásnégyzet, független valószínűségi változók összegének eloszlásának (a konvoluciónak) kiszámítását, (akkor is, ha nem feltétlenül független valószínűségi változók összegét vizsgáljuk és ennek érdekében tudni kell az analízisben tanult integrál és differenciálszámítást is. A tanult diszkrét és folytonos eloszlásokat, illetve azok valószínűségi tartalmát tudni kell.

Tárgyaltuk a Markov és Csebisev egyenlőtlenséget. Ezután az előadás fontos témája volt az eloszlások fogalma, legfontosabb tulajdonságai, a valószínűségszámításban szereplő különböző konvergenciafogalmak (eloszlásban való konvergencia, sztochasztikus konvergencia, egy valószínűséggel való konvergencia) valamint azok jelentése és kapcsolata. Ugyancsak fontos megérteni a nagy számok gyenge és erős törvényét. A bizonyítások részleteinek ismerete nem szükséges.

Az előadás legfontosabb témája a centrális határeloszlástétel volt. Ennek tárgyalása érdekében bevezettük a karakterisztikus függvény fogalmát. Ennek erészletes tárgyalására azonban nem maradt idő, ezért ezt nem fogom kérni. Viszont nagyon fontos tudni azt, hogyan és mire lehet használni a centrális határeloszlástételt. Ebbe beletartoznak az előadáson tárgyalt statisztikai problémák is.

Végül a többdimenziós centrális határeloszlástételt és a χ-négyzet próbát tárgyaltuk. Ezen belül tárgyaltuk bizonyos fogalmak (eloszlás függvény, sűrűségfüggvény, eloszlásban való konvergencia, várható érték, szórásnégyzet) több-dimenziós megfelelőit. Ezen fogalmakat és a hozzájuk kötödő eredményeket ismerni kell. Viszont a bizonyításokat nem kérem, ezekről elegendő bizonyos durva áttekintését ismerni.

Végül megjegyzem, hogy a vizsga elsősorban feladatok megoldásából fog állni. Ezért a felkészüléshez azt javaslom, hogy ne csak az előadások, hanem a gyakorlatokon tárgyalt feladatokat is tanulmányozzák. Ez utóbbiak is megtalálhatóak (a feladatok megoldásaival együtt) megtalálhatóak a homepage-emen.

Homepage-em címe: http://www.renyi.hu/~major