A november 27-i gyakorlat témája
Rövid összefoglaló

Az utolsó két gyakorlaton elsősorban azzal foglalkozunk, hogyan lehet megérteni annak a jegyzetnek néhány eredményét, mely a vizsgan kérdezett anyagot tartalmazza.

Megjegyzések a mértékelméleti előkészítéshez.

Az itt szereplő anyagnak (esetleg kissé más szóhasználattal) szerepelnie kellett a mértékelméleti előadáson. A lényeg az, hogy a σ-algebra illetve σ-additív halmazfüggvény fogalma és a róluk szóló általános eredmények lehetővé teszik a valószínűségszámítás felépítését. Fontos, hogy ezen fogalmak segítségével be lehet vezetni a Riemann integrál megfelelőjét, a Lebesgue integrál fogalmát. A valószínűségi változó nem más mint a valószínűségi mezőn értelmezett mérhető függvény.

Egy az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn értelmezett $f(\omega)$ függvény (Lebesgue) integrálját a P mérték szerint a Riemann integrálhoz hasonlóan egyszeű, természetes integrálközelítések limeszeként definiáljuk. A különbség az, hogy ebben az esetben a felosztást másképp végezzük el. Veszünk egy $\varepsilon > 0$ számot, mellyel majd nullához tartunk, minden $k = 0, \pm 1, \pm 2, \dots$ egész számra definiáljuk az

$$A(k, f, \varepsilon) = \{\omega\colon k\varepsilon \le f(\omega) < (k+1)\varepsilon\}$$

halmazokat. A valószínűségszámítás felépítése alapján lehet az $A(k, f, \varepsilon)$ halmazok $P(A(k, f, \varepsilon)$ valószínűségéről beszélni, ezért definiálhatjuk a $\sum\limits_{k=-\infty}^{\infty} k\varepsilon P(A(k, f\varepsilon)$ összegeket. Ezeket tekintjük az $\int f(\omega) dP(\omega)$ integrálközelítéseinek. Nagyon általános feltételek mellett ezeknek az összegeknek van limeszük $\varepsilon \to 0$ esetén, és ez a limesz lesz az integrál.

A valószínűségi változó fogalma ebben a jegyzetben úgg van tárgyalva, hogy a valószínűségi változó értékeit tetszőleges szeparábilis metrikus térben veheti fel. Ez nem okoz elvi nehézséget. Értsük meg az 1.13 Állítás értelmét. (Ennek van jelentősége később a feltételes várható érték definiciójában.) Ezen állítás megértéséhez először az 1.10 definiciót kell megérteni. Az ott szereplő $\sigma(\xi_\gamma,\ \gamma \in \Gamma)$ σ-algebra szemléletes tartalma a következő. Ha bizonyos valószínűségi változók értékeit meg tudjuk figyelni, akkor ez azt jelenti, hogy bizonyos eseményekről meg tudjuk mondani, hogy bekövetkeztek-e vagy sem. A jobb megértés céljából tekintsünk egy példát: Ledobunk véletlenül az egységintervallumra 20 pontot. Legyen $\xi_j(\omega)$ a j-ik dobás eredménye, $1 \le j \le 20$. Ha a $\sigma(\xi_j(\omega),\ 1 \le j \le 10$ σ-algebrát tekintjük, akkor ez tartalmazza az összes olyan (mérhető) eseményt, mely csak az első 10 dobáseredménytől függ, és nem tartalmazza azokat, melyek más eseménytől is függnek. így például az az esemény, hogy az első 10 dobás összegének az eredménye kisebb mint 6 benne van ebben a σ-algebrában, de az az esemény, hogy az első 11 dobás eredménye kisebb mint 6 már nincsen.

Ha egy ξ valószínűségi változó által generált σ- algebrát tekintünk, akkor a ξ valószínűségi változó (mérhető) függvényei benne vannak ebben a ξ valószínűségi változó által megfigyelhető események által generált σ-algebrában, azaz ha meg tudjuk állapítani, hogy a $\xi \in B$ esemény bekövetkezett-e minden B (Borel-mérhető) halmazra, és $g(\cdot)$ (Borel mérhető) függvény, akkor a $g(\xi) \in B$ esemény bekövetkeztét vagy be nem következtét is meg tudjuk állapítani minden B Borel-mérhető halmazra. Sőt igaz ennek az állításnak a megfordítása is, és ez az állítás nehezebb fele. Nevezetesen, ha egy $\eta(\omega)$ valószínűségi változó olyan, hogy mindegyik $\{\omega \colon \eta(\omega) \in B\}$ eseményről meg tudjuk állapítani, hogy bekövetkezett-e vagy sem, a $\xi(\omega)$ ismeretében, akkor létezik olyan $g(x)$ (Borel-mérhető) függvény a számegyenesen, melyre igaz, hogy $\eta(\omega) = g(\xi(\omega))$.

Végül az (1.15) állításról néhány szó. Ez egy valószínűségi vektor eloszlásfüggvényének a jellemzéséről szól. Az állítás nehéz fele az, hogy egy az adott tulajdonságokkal rendelkező függvény eloszlásfüggvény, azaz létezik olyan valószínűségi mező és rajta olyan valószínűségi változó, melynek ez az eloszlásfüggvénye.

Érdemes lehet megjegyezni, hogy van egy természetes konstrukció, de ezen konstrukció helyességének igazolásához bizonyos mértékelméleti eredmények szükségesek. Nevezetesen definiáljuk az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt és rajta a $\xi(\omega)$ valószínűségi változót a következő módon: Legyen $\Omega = R^k$, a k dimenziós Euklideszi tér, egy elemi esemény $\omega = (x_1, \dots, x_k)$-re legyen $\xi(\omega) = \xi(x_1, \dots, x_k) = (x_1, \dots, x_k)$, $\mathcal{A}$ a Borel σ-algebra. Legyen P az F függvény által definiált Lebesgue–Stieltjes mérték. Nem nehéz belátni, hogy az így definiált ξ valószínűsségi változó F eloszlású lesz. Egy nem triviális lépés van ezen konstrukció helyességének bizonyításában. Az, hogy a Lebesgue–Stieltjes mérték valóban valószínűségi mérték. Annak indoklása követel külön érvelést, hogy a Lebesgue–Stieltjes mérték σ-additív.

Megjegyzések a várható értékről írottakról.

Ez a fejezet tulajdonképpen felsorolja azt, amit mértékelméletből a Lebesgue integrálról kellett tanulni. Az itt szereplő eredmények tartalmazzák pontosan azt, hogy milyen limeszelést lehet végrhajtani az integrálás során, ha egyszerű függvényekkel közelítjük az integrálandó függvényt. Ezek a dominált konvergencia tétel (xv) (szokták hívni Lebesgue tételnek is), a monoton konvergencia tétel (xiii) (szokták hívni Beppo–Levy tételnek is), és a Fatou lemma. A későbbiekben (a feltételes várható érték fogalmának bevezetésében) fontos szerepet játszik a (3.7) pontban megfogalmazott Radon–Nykodim tétel. Ennek megfogalmazásához szükséges a (3.6) definició. Jegyezzük meg, hogy a Radon–Nykodim tétel existencia tétel. Ez azt jelenti, hogy nem ismeretes olyan módszer, mely általános esetben eljárást ad egy mértéknek másik mérték szerinti Radon–Nykodim deriváltjának a kiszámítására. Ez teszi elég kényelmetlenné a feltételes valószínűséggel és várható értékkel való számolást az általános esetben. Ugyanis a feltételes valószínűség és várható érték definixiója a Radon–Nyko-dim tételen alapul. Viszont gyakorlati esetekben, amikor ismert a tekintett valószínűségi változók sűrűségfüggvénye, akkor a vizsgálandó feltételes valószínűségek és várható értékek explicit módon számolhatóak néhány alapvető analízisbeli eredmény alapján.

Megjegyzések a karakterisztikus függvényekről írottakról.

Egy fontos megjegyzés, hogy a karakterisztikus függvény komplex értékű függvény, egy komplex szám értékű valószínűségi változónak a várható értéke. Ezért ellenőrizni kell, hogy a valós értékű valószínűségi változókról tanultak ebben az esetben is érvényben maradnak. Különösen fontos, hogy komplex szám értékű valószínűségi változókra is igaz, hogy független valószínűségi változók szorzatának a várható értéke megegyezik ezen valószínűségi változók várható értékeinek szorzatával.

A karakterisztikus függvény módszer tulajdonképpen nem más mint a Fourier analízis alkalmazása a valószínűségszámításban. A jegyzetben felsorolt eredmények közül a legfontosabb a (4.6) folytonossági tétel. A (4.5) állításban szerepelnek a karakterisztikus függvény legfontosabb tulajdonságai. Bár a későbbiekben nincs rá szükség, mégis értsük meg mit allít a (iv) pontban megfogalmazott Bochner tétel.

Tetszőleges integrálható $f(\cdot)$ függvénynek illetve véges (korlátos változású) előjeles mértéknek lehet definiálni az $inte^{itu}f(u)\,du$ $inte^{itu}F(\,du)$ Fourier transzformáltját. A karakterisztikus függvényt az tünteti ki ezek közül, hogy ebben az esetben pozitív függvénynek illetve pozitív (azaz nem előjeles mértéknek) definiáljuk a Fourier transzformáltját. Annak, hogy az eredeti függvény (mérték) pozitív, (nem-negatív) a Fourier transzformáltak terében az felel meg, hogy a függvény illetve mérték Fourier transzformáltja pozitív (szemi)definit. Ez a Bochner tétel tartalma. A kimondott állításban hiányzik az az észrevétel, hogy mivel a valószínűségi mérték 1-re normált, ezért $\varphi(0)=1$.

A (xi) inverziós formulának a bizonyítását nem láttam a jegyzetben. Ennek az eredménynek lenne némi jelentősége a valószínűségszámításban, ha a centrális határeloszlástételt nemcsak eredeti formájában akarnánk bizonyítani, hanem azt is, hogy alkalmas feltevések mellett független valószínűségi változók normalizált összegének a sűrűségfüggvénye tart a normális sűrűségfüggvényhez. Az inverziós formula egy természetes és egyszerű bizonyítása lenne a Fourier analízisben fontos szerepet játszó Parseval formula alkalmazása. Természetesen előbb a Parseval formulát is bizonyítani kellene. A részletek tárgyaláság elhagyom.

Magjegyzések a centrális határeloszlástételhez.

Érdemes bevezetni és megérteni a szériasorozat fogalmát. Bár ez a fogalom a jegyzetben formálisan nincs bevezetve, valójában ott a centrális határeloszlástételt szériasorozatokra kimondott alakja szerepel.

Szériasorozatok definiciója. *A*

$$\begin{array}{lll} \xi_{1,1}, & \dots, & \xi_{1,n_1} \\ \vdots & & \vdots \\ \xi_{k,1}, & \dots, & \xi_{k,n_k} \\ \vdots & & \vdots \end{array}$$

valószínűségi változók rendszere, $k\to\infty$, szériasorozat, ha az egy sorban levő $\xi_{k,1}$, … , ξ_{k,n_k} valószínűségi változók függetlenek. (A különböző sorokban levő valószínűségi változók kapcsolatáról nem tételezünk fel semmit.)

A centrális határeloszlástétel szériasorozatokra megfogalmazott alakjában az egy sorban szereplő valószínűségi változók $S_k = \sum_{j=1}^{n_k} \xi_{n,k}$ összegét tekintjük, és azt állítjuk, hogy alkalmas feltevések mellett az S_k valószínűségi változók konvergálnak eloszlásban (ezt jelenti a gyenge konvergencia) a normális eloszláshoz. Feltehetjük, hogy $\operatorname{Var} S_k \to 1$, ha $k \to \infty$. Valóban, elosztva az egy sorban levő valószíbűségi vátozókat ugyazzal az alkalmas konstanssal ezt feltehetjük. A bizonyításában nincs jelentősége annak, hogy a különböző sorokban szereplő valószínűségi változók között mi a kapcsolat. Az ilyen alakú határeloszlástétel a legáltalánosabb alakú. Ha adva van független, nulla várható értékű valószínűségi változók $\xi_1, \xi_2, \dots$, végtelen sorozata, akkor elkészítve az $n_k = k$, $\xi_{k,j} = \dfrac{xi_j}{B_k}$, $B_k = \sum_{j=1} \operatorname{Var} \xi_j$, $1 \le j \le n_k = k$, sorozatot, egy szériasorozatot definiáltunk, és az erre a szériasorozatra megfogalmazott centrális határeloszlástétel megegyezik az eredeti centrális határeloszlástétellel.

Ahhoz, hogy a centrális határeloszlástétel érvényes legyen az egyik természetes feltétel az, hogy az egyes valószínűségi változók szórásnégyzete az összeg szórásnégyzetéhez képest kicsi legyen, azaz $r_k = \sup\limits_{1 \le j \le n_k} \operatorname{Var} \xi_{k,j}$ tartson nullához. Egy másik feltétel, amelyet feltesszünk az úgynevezett Lindeberg feltétel.

Lindeberg feltétel definiciója: *Legyen $\xi_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, olyan szériasorozat, melyre $E\xi_{k,j} = 0$, $k = 1, 2, \dots$, $1 \le j \le n_k$, és $\lim\limits_{k \to \infty} \sum_{j=1}^{n_k} E\xi_{k,j}^2 = 1$. Ez a szériasorozat akkor és csak akkor teljesíti a Lindeberg feltételt, ha tetszőleges $\varepsilon > 0$ számra*

$$\lim_{k \to \infty} L_n(\varepsilon) = \lim_{k \to \infty} \sum_{j=1}^{n_k} E\xi_{k,j}^2 I\left(\{|\xi_{k,j}| > \varepsilon\}\right) = 0,$$

ahol $I(A)$ egy A halmaz indikátor függvénye.

A 8. fejezet fő eredménye, a Lindeberg tételnek nevezett eredmény (8.1) tétel arra ad becslést, hogy egy megfelelően síma és a végtelenben nem túl gyorsan nővő $g(u)$ függvényre mennyire kicsi az $Eg(S_k) - Eg(\eta_k)$ különbség, ahol S_k a szériasorozat k-ik sorában szereplő összeg, η pedig standard normális eloszlású valószínűségi változó. Emlékezzünk arra, hogy a centrális határeloszlástétel azt jelenti, hogy amennyiben a $g(u) = g_t(u) = e^{itu}$ választást tesszük, akkor az $Eg(S_k) - Eg(\eta_k) \to 0$, ha $k \to \infty$ reláció jelenti a centrális határeloszlástétel érvényességét.

A több-dimenziós normális eloszlásról.

Az első kérdés, ami e fogalom kapcsán felmerül az, hogy miért fontos ez a fogalom. A természetes válasz az, hogy a centrális határeloszlástétel, illetve annak több-dimenziós változata indokolja e fogalom bevezetését. A centrális határeloszlástétel azt mondja, hogy független valószínűségi változók normalizált összegeinek a határeloszlása nagyon általános feltételek mellett mindig ugyanaz a normális eloszlás. Felmerül az a kérdés, hogy érvényes-e ennek az eredménynek a megfelelője, melyek vektor értékű független valószínűségi változók normalizált összegeiről szólnak.

Erre a kérdésre pozitív választ lehet adni, és ez fontos szerepet játszik néhány olyan fontos kérdés vizsgálatában mint a chi-négyzet próba. Az eredmény illetve a több-dimenziós normális eloszlásfüggvény fogalmának a megértéséhez szükséges először megérteni a várható érték illetve a szórásnégyzet fogalmának több-dimenziós megfelelőjét. A nehezebb és tárgyalandó feladat a szórásnégyzet több-dimenziós megfelelőjének, a kovarianciamátrixnak a definiciója.

A kovarianciamátrix definiciója. *Legyen $\xi=(\xi_1,\dots,\xi_k)$ egy k-dimenziós véletlen vektor, melynek minden koordinátájára teljesül az $E\xi_j^2<\infty$, $1\le j\le k$, feltétel. Ekkor a ξ k-dimenziós valószínűségi változó D kovarianciamátrixa az a $k\times k$ mátrix, melynek j-ik sorában és l-ik oszlopában a $\operatorname{Cov}(\xi_j,\xi_l)=E\xi_j\xi_l-E\xi_jE\xi_l$, $1\le j,l\le k$, szám áll.*

A kovarianciamátrix hasonlóan viselkedik mint annak egy dimenziós megfelelője, a szórásnégyzet. Így például független vektorok összegének a kovarianciamátrixa megegyezik az összeadandók kovarianciamátrixának az összegével, ha egy vektort megszorzunk egy c számmal, akkor annak kovarianciamátrixa a c^2 számmal szorzódik meg. Ezenkívül fontos a következő eredmény.

Tétel a kovarianciamátrix jellemzéséről. *Egy k-dimenziós véletlen vektor kovarianciamátrixa $k\times k$ méretű szimmetrikus pozitív szemidefinit mátrix. Egy szimmetrikus $D=(d_{j,l})$, $1\le j,l\le k$ $k\times k$ méretű mátrixot pozitív szemidefinitnek nevezünk, ha minden $x=(x_1,\dots,x_k)$ k-dimeziós vektorra $xDx^*=\sum\limits_{j=1}^k\sum\limits_{l=1}^k x_jx_ld_{j,l}\ge 0$. Megfordítva, minden D $k\times k$ méretű pozitív definit mátrixhoz és m k–dimenziós vektorhoz létezik olyan k–dimenziós normális eloszlású véletlen vektor, melynek a várható értéke m, kovarianciamátrixa pedig D. Sőt az is igaz, hogy egy több-dimenziós normális eloszlású véletlen vektor eloszlását egyértelműen meghatározza annak várható értéke és kovarianciamátrixa.*

Annak érdekében, hogy ezt az eredményt megértsük fel kell idéznünk a több-dimenziós normális eloszlású valószínűségi változó fogalmát.

A több-dimenziós normális eloszlású valószínűségi változó definiciója. *Ha ξ_1, ... , ξ_k független, standard normális eloszlású valószínűségi változók, akkor a segítségükkel definiált $\xi=(\xi_1,\dots,\xi_k)$ véletlen vektort k-dimenziós standard normális eloszlású valószínűségi változónak nevezzük. Ha ξ k-dimenziós standard normális eloszlású valószínűségi változó, A tetszőleges $k\times k$ méretű mátrix m k-dimenziós véletlen vektor, akkor $\eta=\xi A+m$ k-dimenziós normális eloszlású vektor. Akkor nevezünk egy véletlen vektort k-dimenziós normális eloszlású véletlen vektornak, ha annak eloszlása megegyezik egy előbb definiált $\eta=\xi A+m$ alakú k-dimenziós normális eloszlású vektor eloszlásával.*

Be lehet látni, hogy igaz a következő eredmény.

Tétel. *Az előbb definiált $\eta=\xi A+m$ alakú k-dimenziós normális eloszlású vektor várható értéke m kovarianciamátrixa pedig a $D=AA^*$ mátrix. Karakterisztikus* *függvénye a*

$$\varphi(t) = \varphi(t_1, \dots, t_k) = Ee^{i(t,\eta)} = Ee^{i(t_1\eta_1+\cdots+t_k\eta_k)} = e^{i(t,m)-tDt^*/2}$$

függvény, ahol $t = (t_1, \dots, t_k)$, $(\cdot, \cdot)$ *skalárszorzatot* * *pedig transzponáltat jelöl. Ez utobbi formula azért hasznos, mert egy (több-dimenziós) eloszlást meghatároz annak karakterisztikus függvénye.*

Azt a tényt, hogy tetszőleges D pozitív szermidefinit mátrix megjelenik mint alkalmas normális eloszlású valószínűségi vektor kovarianciamátrixa következik egyrészt a fenti állításokból, továbbá a következő lineáris algebrai eredményből.

Tétel. *Egy szimmetrikus* D *pozitív szemidefinit mátrixból lehet "négyzetgyököt vonni", azaz a* $D = AA^*$ *egyenletnek van megoldása. Ez a megoldás nem egyértelmű. Valóban, ha* U *unitér mátrix, azaz* $UU^* = U^*U = I$, *akkor az* $\bar{A} = AU$ *mátrixra is igaz, hogy* $\bar{A}\bar{A}^* = D$, *ha* $D = AA^*$.

Ezek az eredmények azt is jelentik, hogy bár egy előírt eloszlású normális eloszlású véletlen vektor $\eta = \xi A + m$ reprezentációjában az A mátrix megadása nem egyértelmű, mégis a D kovarianciamátrix és m várható érték meghatározza a véletlen vektor eloszlását. Ennek az alábbi fontos következménye is van.

Tétel *Legyen* $\xi = (\xi_1, \dots, \xi_k)$ *normális eloszlású véletlen vektor, melynek első* l *és utolsó* $k - l$ *koordinátája korrelálatlan valamely* $1 \le l < k$ *számra, azaz* $E\xi_p\xi_q = E\xi_p E\xi_q$, *ha* $1 \le p \le l < q \le k$. *Ekkor a* $(\xi_1, \dots, \xi_l)$ *illetve* $(\xi_{l+1}, \dots, \xi_k)$ *vektorok függetlenek.*

Ez utóbbi eredmény kiolvasható abból a tényből, hogy az adott feltétel mellett a ξ véletlen vektor faktorizálódik. Jegyezzük meg, hogy független valószínűségi változók mindig korrelálatlanok, de az állítás megfordítása nem igaz. Léteznek olyan valószínűségi változók, melyek korrelálatlanok, de nem függetlenek. Ez a tény rámutat a normális eloszlású valószínűségi vektorok egyik különleges tulajdonságára, arra, hogy ebben a speciális esetben a koordináták korrelálatlanságából következik azok függetlensége is.