A Valószínűségszámítás II. előadássorozat második előadása.
2002. szeptember 17.

A valószínűségszámítás megalapozása. *Folytatás.*

Ezem az előadáson először valószínűségi változók várható értékével fogunk foglalkozni. Ismertetjük (kissé vázlatosan) ennek definicióját, és megtárgyaljuk legfontosabb tulajdonságait. Ez a problémakör is nagyon szorosan kapcsolódik a mértékelmélethez, azon belül az integrálok, elsősorban a Lebesgu integrál vizsgálatához. Ez a kapcsolat rögtön látszik a várható érték definiciójából.

Valószínűségi változó várható értékének a definiciója. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. A $\xi(\omega)$ valószínűségi változó várható értéke az*

$$E\xi(\omega) = \int_\Omega \xi(\omega)\, dP(\omega)$$

Lebesgue integrál a P mérték szerint, feltéve, hogy ez az integrál létezik. Ha ez az integrál nem létezik akkor a $\xi(\omega)$ valószínűségi vátozó várható értékét nem definiáljuk.

Maga a Lebesgue integrál fogalma szerepel a mértékelméletben, ezért itt annak definiciójáról csak rövid emlékeztetőt adok. A Lebesgue integrál pontos elméletének ismerete nem feltétele ezen előadás megértésének.

Legyen adva egy $(X, \mathcal{A})$ mérhető tér, azaz egy X halmaz, és annak bizonyos kijelölt részhalmazai, melyek σ-algebrát alkotnak, valamint egy véges μ mérték az $\mathcal{A}$ σ-algebrán, ami azt jelenti, hogy $\mu(X) < \infty$. (Valójában ezt a feltételt lehet és érdemes is gyengíteni, tekinthetünk úgynevezett σ-véges mértékeket is. Egy μ mérték σ-véges az $(X, \mathcal{A})$ téren, ha létezik az X halmaznak olyan megszámlálható sok B_j, $j = 1, 2, \dots$, halmazokra való felbontása, melyre azonkívül, hogy $X = \bigcup\limits_{j=1}^{\infty} B_j$, teljesül még az $\mu(B_j) < \infty$, $j = 1, 2, \dots$ feltétel is. Nem véges, de σ-véges mértékre fontos példa a Lebesgue mérték a számegyenesen.) Jelöljük x betűvel az X tér egy általános pontját. Tekintjük a „szép" (mérhető) $f(x)$ valós értékű függvényeket az $(X, \mathcal{A})$ téren, és az ilyen függvényeknek fogjuk definiálni a (Lebesgue) integrálját a μ mérték szerint. A definició a következő természetes elven alapul. Először a legegyszerűbb, úgynevezett lépcsős függvényekre definiáljuk az integrált, mely függvényekre létezik az X halmaznak olyan véges $A_1, \dots, A_n$ particiója, ahol a függvény konstans. Ha az f függvény olyan lépcsős függvény, melyre az X halmaz valamilyen véges $A_1, \dots, A_n$ particiójára, (az, hogy ezen halmazok particiót alkotnak azt jelenti, hogy $A_1, \dots, A_n$ diszjunkt halmazok, és $\bigcup\limits_{j=1}^{n} A_j = X$), $f(x) = c_j$ minden $x \in A_j$ pontra, $j = 1, \dots, n$, akkor az f függvény integrálját az $\int_X f(x)\mu(\,dx) = \sum\limits_{j=1}^{n} f(x_j)\mu(A_j)$, $x_j \in A_j$, $j = 1, \dots, n$, képlettel definiáljuk. Természetes közelítéssel tetszőleges pozitív (vagy negatív) mérhető f függvényt közelíteni lehet pozitív (vagy negatív) lépcsős függvények szorozatával (például elérhető, hogy $\mu\left(x\colon |f_n(x)-f(x)|>\frac{1}{n}\right)<\frac{1}{n}$), és ezután az f függvény integrálját az

$$\int_X f(x)\mu(\,dx)=\lim_{n\to\infty}\int_X f_n(x)\mu(\,dx)$$

képlet segítségével lehet definiálni. (Ez a limesz lehet végtelen (minusz végtelen) is.) Végül egy általános mérhető f függvényt felírhatunk $f(x)=f_+(x)+f_-(x)$ alakban is, ahol $f_+(x)=\max(0,f(x))$ és $f_-(x)=\min(0,f(x))$ az $f(x)$ függvény pozitív és negatív része, és az $f(x)$ függvény integrálját definiálhatjuk az $\int_X f(x)\mu(\,dx)=\int_X f_+(x)\mu(\,dx)+\int_X f_-(x)\mu(\,dx)$ képlettel, feltéve, hogy a jobboldalon szereplő összeg értelmes. (Azaz nem plusz végtelent és minusz végtelent kell összeadni.) Természetesen a fenti definició jogosságának indoklásához számos kérdést tisztázni kell, (például, honnan tudjuk, hogy minden mérhető függvény közelíthető ilyen módon, azt hogy az integrál értéke nem függ attól, hogy milyen közelítést alkalmazunk?), de mivel ez a problémakör a mértékelmélet része, és számunkra nem lesz fontos e kérdések vizsgálata, ezért ezt elhagyjuk. Megjegyezzük, hogy abban, hogy az integrálnak egy kényelmes jó modelljét tudjuk kidolgozni fontos szerepet játszott az a körülmény, hogy a μ mérték nemcsak additív, hanem σ-additív is.)

Beszéljük meg röviden a Lebesgue és korábban tanult Riemann integrál kapcsolatát.

Első megjegyzés: Láttuk, hogy a Lebesgue integrált úgy definiáltuk, hogy először tekintettük a legegyszerűbb, úgynevezett lépcsős függvényeket, és az ilyen függvények integrálját természetes összegként definiáltuk. Általános függvények integráljának a definicióját visszavezetjük erre az esetre alkalmas határátmenet segítségével. A Riemann integrál esetében is hasonló az integrál definiciója, csak ott azok a speciális függvények, melyek integrálját egy összeg segítségével definiáljuk másképp vannak meghatározva. Ugyanis, ha tekintjük annak az intervallumnak egy felosztását kis intervallumokra, és tekintünk egy olyan függvényt, melynek értéke a felosztásban szereplő mindegyik kis intervallumban egy olyan konstanssal egyezik, melynek értéke megegyezik az eredeti függvény értékével ezen intervallum valamelyik kis pontjában, akkor a Riemann integrál definiciójában szereplő közelítő összegek ilyen függvények integráljával egyeznek meg. Tehát a Riemann integrál definiciója is interpretálható úgy mint az integrálandó függvényt közelítő alkalmas egyszerű függvények integráljainak a limesze.

A fő különbség a Riemann és Lebesgue integrálok között az, hogy a Lebesgue integrál definiciójában a közelítő egyszerű (elemi) függvények családja gazdagabb, mint a Riemann integrálok esetében. Annak érdekében, hogy ezt megtehessük, ki kellett építeni egy σ-additív mértéket egy (sok halmazt tartalmazó) σ-algebrára. A fenti gondolatot kidolgozva meg lehet érteni, hogy egy Riemann integrálható függvény egyidejűleg Lebesgue integrálható függvény is (a Lebesgue mérték szerint), és egy ilyen függvény Riemann és Lebesgue integrálja egyenlő.

Második megjegyzés: Érdemes megfogalmazni a következő egyszerű, de fontos észrevételt. Ha $\xi(\omega)$ olyan valószínűségi változó egy $(\Omega,\mathcal{A},P)$ valószínűségi változó, melyre

$P(a \leq \xi(\omega) \leq b) = 1$ alkalmas $-\infty < a \leq b < \infty$ számokra, akkor

$$a \leq \int \xi(\omega) P(\,d\omega) \leq b.$$

Valóban, egy ilyen függvény hasonló tulajdonságú elemi függvényekkel approximálható, és mivel az ilyen elemi függvények integráljaira teljesül az említett tulajdonság, ezért határátmenettel megkapjuk ezt az állítást. Ennek az egyenlőtlenségnek érdemes megfogalmazni az alábbi következményét is. Ha $\xi(\omega)$ és $\eta(\omega)$ két olyan valószínűségi változó, melyre $P(|\xi(\omega) - \eta(\omega)| \leq a) = 1$ valamilyen $a > 0$ számra, akkor a $\xi(\omega)$ és $\eta(\omega)$ valószínűségi egyszerre integrálható, vagy nem integrálható, és

$$\left| \int \xi(\omega) P(\,d\omega) - \int \eta(\omega) P(\,d\omega) \right| \leq a.$$

Harmadik megjegyzés: A Lebesgue integrál, hasonlóan a Riemann integrálhoz lineáris leképezés, azaz ha $f(\cdot)$ és $g(\cdot)$ két integrálható függvény egy $(X, \mathcal{A}, \mu)$ mértéktéren, a és b valós számok, akkor

$$\int (af(x)\mu(\,dx) + bg(x))\,\mu(\,dx) = a \int f(x)\mu(\,dx) + b \int g(x)\mu(\,dx).$$

A Ennek az állítás következménye a várható érték egyik fontos tulajdonsága, amelyiket külön tétel fogalmazza meg.

Tétel. *Legyenek $\xi_1, \dots, \xi_n$ olyan valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyekre $E|\xi_j| < \infty$, $j = 1, \dots, n$, $c_1, \dots, c_n$ tetszőleges valós számok. Ekkor*

$$E(c_1\xi_1(\omega) + \cdots + c_n\xi_n(\omega)) = c_1 E\xi_1(\omega) + \cdots + c_n E\xi_n(\omega).$$

Az egyik rendkívül fontos kérdés az, hogy hogyan tudjuk kiszámolni egy valószínűségi változó várható értékét. Külön hangsúlyozzuk, hogy a valószínűségi problémákban általában nem adjuk meg effektíve azt a valószínűségi mezőt, ahol a valószínűségi változó létezik, hanem csak a valószínűségi változó eloszlását ismerjük. Elegendő-e egy valószínűségi változó eloszlásának ismerete ahhoz, hogy kiszámoljuk a várható értékét? Erre a kérdésre az alábbi tétel ad pozitív választ.

Tétel *Legyen egy $(\Omega, \mathcal{A}, P)$ téren értelmezett $\xi(\omega)$ valószínűségi változó eloszlásfüggvénye $F(x)$. Ekkor a ξ valószínűségi változó várható értéke teljessíti az*

$$E\xi(\omega) = \int_{-\infty}^{\infty} uF(\,du)$$

azonosságot. Általánosabban, ha $g(x)$ tetszőleges (mérhető) valós függvény, akkor

$$Eg(\xi(\omega)) = \int_{-\infty}^{\infty} g(u)F(\,du).$$

Még általánosabban, tekintsünk n $\xi_1(\omega),\dots,\xi_n(\omega)$ valószínűségi változókat egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, és legyen $g(x_1,\dots,x_n)$ tetszőleges (mérhető) n-változós függvény. Ha $F(x_1,\dots,x_n)=P(\xi_1<x_1,\dots,\xi_n<x_n)$ a $(\xi_1(\omega),\dots,\xi_n(\omega))$ vektor eloszlásfüggvénye, akkor

$$Eg(\xi_1(\omega),\dots,\xi_n(\omega))=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(u_1,\dots,u_n)F(\,du_1,\dots,\,du_n).$$

A Tételben megfogalmazott összes azonosság úgy értendő, hogy amennyiben az azonosság egyik oldalán álló kifejezés értelmes, akkor a másik kifejezés is az, és a két kifejezés egyenlő.

A fenti tételben az F eloszlásfüggvény szerinti integrál azt jelenti, hogy tekintjük az F eloszlásfüggvény által meghatározott μ_F Lebesgue–Stieltjes mértéket (a számegyenes vagy az n-dimenziós tér Borel mérhető halmazainak σ-algebráján), és a μ_F mérték szerinti Lebesgue integrált tekintjük. A fent megfogalmazott tétel a mértékelméleti vizsgálatok természetes anyaga, ezért itt annak csak rövid magyarázatát adjuk a részletek kidolgozása nélkül.

Idézzük fel, hogy mint az előző órán megtárgyaltuk egy $F(x_1,\dots,x_n)$ n-változós eloszlásfüggvényhez a következő módon (is) konstruálhatunk egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőt és rajta olyan $\xi_1,\dots,\xi_n$ valószínűségi változókat, melyek együttes eloszlásfüggvénye ez az F eloszlásfüggvény: $(\Omega,\mathcal{A},P)=(R^n,\mathcal{B}^n,\mu_F)$, ahol R^n az n-dimenziós euklidészi tér, $\mathcal{B}^n$ az n-dimenziós tér Borel σ-alagebra, μ_F az F eloszlásfüggvény által meghatározott Lebesgue–Stieltjes mérték, és a ξ_k, $1\le k\le n$, valószínűségi változókat a $\xi_k(x_1,\dots,x_n)=x_k$, $1\le k\le n$, képlet segítségével definiáljuk. Világos, hogy ebben az esetben a Tételben felírt integrál formulák érvényesek.

Az általános eset visszavezethető erre a speciális esetre a következő észrevétel segítségével: Legyen adva egy $(\Omega,\mathcal{A},P)$ valószínűségi mező és rajta egy $F(x_1,\dots,x_n)$ eloszlású $(\xi_1(\omega),\dots,\xi_n(\omega))$ véletlen vektor. Ekkor a $T\colon(\Omega,\mathcal{A},P)\to(R^n,\mathcal{B}^n,\mu_F)$, $T(\omega)=(\xi_1(\omega),\dots,\xi_n(\omega))$, $\omega\in\Omega$, leképezés a mértékelmélet szóhasználatát alkalmazva mértéktartó, ami azt jelenti, hogy tetszőleges $B\in\mathcal{B}$ Borel mérhető halmazra, $\mu_F(B)=P(T^{-1}(B))$, ahol a $T^{-1}(B)$ halmazt, melyet a B halmaz ősképének szoktak hívni a $T^{-1}(B)=\{\omega\colon(\xi_1(\omega),\dots,\xi_n(\omega))\in B\}$ képlet definiálja. A mértékelmélet egy fontos (bár nem nagyon nehéz) eredménye megadja, hogyan transzformálódnak integrálok mértéktartó leképezések hatására, és a fent kimondott képlet ebből kiolvasható. (A tételben kimondott eredmény egyébként azon alapul, hogy véve a fenti T transzformációt a tételben megfogalmazott azonosságok közvetlenül láthatóak lépcsős függvényekre, és ezek az azonosságok kiterjeszthetőek általános függvényekre alkalmas határátmenettel.)

Ahhoz, hogy a fenti integrálokkal jól tudjunk számolni meg kell értenünk a Lebesgue és Riemann integrál kapcsolatát. Mint láttuk minden várható érték kifejezhető az euklidészi téren értelmezett alkalmas Lebesgue–Stieltjes mérték szerinti Lebesgue integrállal. Az analízisben definiáják egy g függvény $\int g(x)F(\,dx)$ Stieltjes integrálját egy F függvény szerint. A mi számunkra az lesz az érdekes, hogy amennyiben egy függvénynek létezik az $\int g(x)F(\,dx)$ Stieltjes integrálja, akkor létezik az $\int g(x)\mu_F(\,dx)$

Lebesgue-Stieltjes integrálja is, ahol μ_F az F függvény által meghatározott Lebesue–Stieltjes mérték, és a két integrál megegyezik. Megfogalmazható ennek az eredménynek a több-dimenziós változata is, de ezzel nem foglalkozunk.

Sok valószínűségi feleadatban fontos szerepet játszik a sűrűségfüggvény szerepe. Megfogalmazzuk ennek definicióját, és ismertetjük legfontosabb tulajdonságait.

Sűrűségfüggvény definiciója. *Legyen $F(x)$ eloszlásfüggvény. Azt mondjuk, hogy a $f(x)$ az $F(x)$ eloszlásfüggvény sűrűségfüggvénye (sokszor azt a szóhasználatot is alkalmazzák, hogy egy $F(x)$ eloszlású valószínűségi változó sűrűségfüggvénye, ha teljesül az*

$$F(x)=\int_{-\infty}^{x} f(u)\,du$$

azonosság minden $-\infty<x<\infty$ számra.

Azt mondjuk, hogy egy n-dimenziós $F(x_1,\dots,x_n)$ eloszlásfüggvénynek sűrűségfüggvénye az $f(x_1,\dots,x_n)$ függvény, ha

$$F(x_1,\dots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f(u_1,\dots,u_n)\,du_1\dots du_n$$

minden $x_1,\dots,x_n$ valós számra.

A sűrűségfüggvény fogalma azért hasznos, mert ha egy valószínűségi változónak létezik sűrűségfüggvénye akkor annak várható értéke kiszámítható e sűrűségfüggvény segítségével. Igaz ugyanis a következő tétel.

Tétel. *Ha egy $F(x)$ eloszlásfüggvénynek f(x) a sűrűségfüggvénye, akkor teljesül az*

$$\int_{=\infty}^{\infty} g(u)F(\,du)=\int_{-\infty}^{\infty} g(u)f(u)\,du \tag{a1}$$

azonossáag minden $g(u)$ függvényre.

Ha egy n-változós $F(x_1,\dots,x_n)$ eloszlásfüggvénynek létezik $f(x_1,\dots,x_n)$ sűrűségfüggvénye, akkor teljesül az

$$\begin{aligned}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} & g(u_1,\dots,u_n)F(\,du_1,\dots,\,du_n)\\ &=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(u_1,\dots,u_n)f(u_1,\dots,u_n)\,du_1,\dots,\,du_n\end{aligned} \tag{a2}$$

azonosság minden n változós $g(u_1,\dots,u_n)$ függvényre. A fenti azonosságok úgy értendőek, hogy a két oldalon szereplő kifejezés egyszerre létezik vagy nem létezik.

Ahhoz, hogy a fenti eredményt használni tudjuk, el kell tudnunk dönteni, hogy egy eloszlásfüggvénynek mikor létezik eloszlásfüggvénye, és ha létezik, akkor ki kell tudnunk azt számolni. Az analízis egyik fontos eredménye, az úgynevezett Newton–Leibniz formula segít e probléma megoldásában. E szerint a formula szerint, ha tekintjük egy $f(x)$ függvény $F(x) = \int_a^x f(u)\,du$ határozatlan integrálját, akkor teljesül az $f(x) = \frac{dF(x)}{dx}$ egyenlőség (majdnem) minden x pontra. (A majdnem kitétel szükséges, hiszen például, ha az $f(x)$ függvény értékét véges sok pontban megváltoztatjuk, akkor az új függvény $F(x)$ határozatlan integrálja megegyezik az eredeti függvény határozatlan integráljával, azaz az $F(x)$ integrál ismeretében nem lehet eldönteni azt, hogy a két függvény közül melyikkel számoltunk. A majdnem minden jelző pontos jelentését megadják a mértékelméletben, ez a mértékelmélet nagyon fontos fogalma.) Igaz ennek az állításnak részleges megfordítása, mely szerint, ha egy $F(x)$ folytonos függvény, és véges sok pont kivételével differenciálható, akkor annak $f(x) = \frac{dF(x)}{dx}$ differenciálhányadosát integrálva visszakapjuk az eredeti függvényt, azaz $F(x) = \int_a^x f(u)\,du + C$ minden x számra egy alkalmas C konstanssal. Számunkra ez az eredmény elsősorban az alábbi következményei miatt érdekes: Ha az $F(x)$ eloszlásfüggvény folytonos, és véges sok pont kivételével deriválható, akkor az $F(x)$ sűrűségfüggvénye az

$$f(x) = \frac{dF(x)}{dx}, \qquad \text{minden } x \text{ számra} \tag{b1}$$

függvény. (A mértékelméletben egy ennél tartalmasabb, de nehezebben bizonyítható eredményt is tárgyalnak. Nevezetesen azt az eredményt, mely szerint annak szükséges és elégséges feltétele, hogy az $F(x)$ eloszlásfüggvénynek létezzék sűrűségfüggvénye az, hogy az $F(x)$ függvény abszolut folytoos legyen. Ekkor az $f(x)$ függvényt a (b1) képlet adja meg. Természetesen ennek az eredményének a megfogalmazásához hozzátartozik az abszolut folytonos függvény fogalmának tisztázása is.)

Érvényes ennek az eredménynek több-dimenziós változata is. Eszerint, amennyiben egy $F(x_1, \dots, x_n)$ n-változós elosztádfüggvény elég síma, akkor létezik $f(x_1, \dots, x_n)$ sűrűségfüggvénye, és azt a

$$f(x_1, \dots, x_n) = \frac{\partial^n F(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n} \tag{b2}$$

képlet adja meg.

Gyakran előfordul, hogy bizonyos feladatokban a benne szereplő valószínűségi változóknak nem az eloszlás, hanem sűrűségfüggvényét adjuk meg. Az eloszlásfüggvényekhez hasonlóan megadhatjuk a sűrűségfüggvények jellemzését is. Ezt az eredményt fogalmazzuk meg a következő tételben.

Tétel. *Egy $f(x)$ függvény akkor és csak akkor sűrűségfüggvénye egy alkalmas eloszlásfüggvénynek, ha $f(x) \ge 0$ (majdnem) minden x valós számra, és*

$$\int_{-\infty}^{\infty} f(u)\,du = 1.$$

Érvényes ennek az állításnak a következő több-dimenziós változata is. Egy n-változós $f(x_1, \dots, x_n)$ függvény akkor és csak akkor sűrűségfüggvénye egy alkalmas n-változós eloszlásfüggvénynek, ha $f(x_1, \dots, x_n) \geq 0$ az n-dimenziós euklidészi tér (majdnem) minden $(x_1, \dots, x_n)$ pontjában, és

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(u_1, \dots, u_n) \, du_1 \dots \, du_n = 1.$$

Felidézzük azt az egyszerű, de hasznos eredményt, mely szerint, amennyiben egy ξ valószínűségi változó diszkrét eloszlású, azaz létezik véges vagy megszámlálható sok olyan $x_1, x_2, \dots,$ érték, melyekre $P(\xi = x_i) = p_i$, $i = 1, 2, \dots,$ $\sum p_i = 1$, akkor a ξ valószínűségi változó egy $g(\xi)$ függvényének a várható értéke az alábbi módon számolható ki:

$$Eg(\xi) = \int g(u) F(\,du) = \sum g(x_i) P(\xi = x_i),$$

ahol $F(\cdot)$ a ξ valószínűségi változó eloszlásfüggvényét jelöli. Ezzel kapcsolatban megfogalmazzuk a következő feladatot:

Feladat:

Jellemezzük az olyan valószínűségi változók eloszlásfüggvényeit, melyek egy valószínűséggel csak véges sok érték valamelyikét vehetik fel. Mutassunk példát olyan diszkrét eloszlású valószínűségi változóra, amelyik minden intervallumban szigorúan monoton nő.

Érdemes megjegyezni a következő egyszerű, de gyakran hasznos összefüggést. Ha egy $F(x)$ eloszlásfüggvényt felbontunk $F(x) = F_1(x) + F_2(x)$ alakban, akkor érvényes tetszőleges $g(x)$ függvényre az

$$\int_{-\infty}^{\infty} g(u) F(\,du) = \int_{-\infty}^{\infty} g(u) F_1(\,du) + \int_{-\infty}^{\infty} g(u) F_2(\,du)$$

azonosság. Továbbá az eloszlás és sűrűségfüggvények közötti kapcsolat általánosítható (számunkra az lesz a lényeges, hogy az alább megfogalmazott állításban eloszlásfüggvények helyett teinthetünk olyan függvényeket is, melyek integrálja nem feltétlenül 1), nevezetesen amennyiben egy F és f függvényekből álló függvénypárra teljesul a (b1) összefüggés (többváltozós függvények esetén a (b2) összefüggés), akkor ezek a függvények egy tetszőleges $g(\cdot)$ függvénnyel együtt teljesítik az (a1) összefüggést (többváltozós esetben az (a2) összefüggést) is. Előfordulhat, hogy olyan valószínűségi változókkal kell dolgoznunk, melyeknek van egy „nem egyre normált sűrűségfüggvénnyel” és egy „nem egyre normált diszkrét kompponenssel” rendelkező része. Ilyenkor érdemes lehet az előbb említett összefüggéseket felhasználni. A következő feladat mutat erre példát.

Feladat:

Az egységintervallumra egyenletes eloszlással ledobunk egy pontot, azaz annak valószínűsége, hogy a pont egy $[a, b] \subset [0, 1]$ intervallumba esik $b - a$, valamint ettől függetlenül egy szabályos pénzdarabot. Ha fej a pénzdobás eredménye, akkor legyen a nyereményünk a leesési hely értékével egyenlő, ha pedig írást dobunk, akkor legyen a nyereményünk a dobás értékétől függetlenül $\frac{1}{2}$. Számoljuk ki nyereményünk várható értékét és szórásnégyzetét.

Megjegyzés: Csak megemlítjük, de pontosan meg nem fogalmazzuk a mértékelmélet néhány fontos eredményényét a (Lebesgue) integrálról, melyekre legalábbis egyelőre nem lesz szükségünk. Ezek az eredmények olyan jellegű állítást fogalmaznak meg, mely szerint, ha függvények egy sorozata egy határfüggvényhez konvergál, akkor nagyon általános feltételek mellett e függvények (Lebesgue) integráljai konvergálnak a határfüggvény integráljához. Ilyen tipusú állítást fogalmaz mag a Lebesgue féle (dominált) konvergencia tétel, a Beppo–Levy tétel monoton függvénysorozatok konvergenciájáról és (bár ez utóbbi kissé eltérő jellegű) a Fatou lemma. Érdemes megjegyezni, hogy míg ahhoz, hogy függvények konvergenciájából csak nagyon erős megkötések mellett következtethetünk a függvények deriváltjainak a konvergenciájára, addig függvények integráljai nagyon általános feltételek esetén konvergál a határfüggvény integráljához.

Részletesebben fogjuk viszont tárgyalni a Fubini tételnek nevezett eredményt, mivel az mind a független valószínűségi változók tulajdonságainak megértésében mind konkrét számolásokban fontos szerepet játszik. Ezért ennek az eredménynek a tartalmáról részletesebben beszélek. Ez az eredmény azt mondja ki, hogy egy szorzattéren értelmezett függvénynek egy szorzat mérték szerinti integrálját szukcessziv integrálaással is ki lehet számolni. A tétel megfogalmazása előtt megemlítem a következő eredményt:

Tétel mértékek szorzatáról. *Legyen adva véges sok $(X_j, \mathcal{A}_j, \mu_j)$, $j = 1, \dots, n$, mértéktér véges μ_j mértékekkel, azaz legyen $\mathcal{A}_j$ az X_j tér bizonyos részhalmazaiból álló σ-algebra, $X_j \in \mathcal{A}_j$, μ_j mérték a $\mathcal{A}_j$ σ-algebrán, melyre $X_j \in \mathcal{A}_j$. Ekkor létezik az $(X, \mathcal{A}, \mu) = (X_1 \times \cdots \times X_n, \mathcal{A}_1 \times \cdots \times \mathcal{A}_n, \mu_1 \times \cdots \times \mu_n)$ szorzat mértéktér, ahol X az összes $(x_1, \dots, x_n)$, $x_j \in X_j$, $1 \le j \le n$, n hosszú sorozatból áll, $\mathcal{A}$ az $A_1 \times \cdots \times A_n$, $A_j \in \mathcal{A}_j$, $1 \le j \le n$ alakú halmazokat tartalmazó legszűkebb σ-algebra, a μ mérték pedig az a mérték, melyre $\mu(A_1 \times \cdots \times A_n) = \mu_1(A_1) \cdots \mu_n(A_n)$ minden ilyen alakú halmazra. A tétel fő tartalma az, hogy létezik ilyen μ mérték az $(X, \mathcal{A})$ téren, méghozzá egyetlen ilyen mérték.*

1. Megjegyzés: Valójában a tétel kissé általánosabban is érvényes. Elég feltenni, hogy a tekintett μ_j mértékek σ-végesek, azaz létezik az X_j halmazoknak olyan $X_j = \bigcup\limits_{k=1}^{\infty} A_j^{(k)}$ előállítása, melyre $\mu(A_j^{(k)}) < \infty$. Egy ilyen általánosítás hasznos lehet, például, ha a számegyenest tekintjük a Borel σ-algebrával és Lebesgue mértékkel, és e tér önmagával vett direkt szorzatát akarjuk tekinteni.

2. Megjegyzés: A most kimondott tétel speciális esete az előző előadásban megfogalmazott eredménynek mértékterek végtelen szorzatáról. Igaz, hogy abban volt egy itt nem szereplő megszorító feltétel, mely szerint valószínűségi mértékeket szorzunk össze, azaz $\mu(X_j) = 1$ minden j-re. Ez a feltétel azonban csak azért szerepelt abban az eredményben, hogy végtelen sok tér szorzatát is tekinthessük. Ha véges sok tér szorzatával foglalkozunk, akkor ez a feltétel nem szükséges, és könnyű tőle megszabadulni.

Most megfogalmazzuk a Fubini tételt.

Fubini tétel. *Legyen adva véges sok $(X_j, \mathcal{A}_j, \mu_j)$, $j = 1, \dots, n$, mértéktér véges (általánosabban σ-véges) μ_j mértékekkel, és tekintsük ezek $(X, \mathcal{A}, \mu) = (X_1 \times \cdots \times X_n, \mathcal{A}_1 \times \cdots \times \mathcal{A}_n, \mu_1 \times \cdots \times \mu_n)$ szorzatterét, melyet az előző tétel megfogalmazásában megadtunk. Legyen $g(x) = g(x_1, \dots, x_n)$ tetszőleges a szorzattéren értelmezett (mérhető) függvény. Ekkor*

$$\begin{aligned}&\int g(x_1, x_2, \dots, x_n)\, d\mu(x_1, \cdots, x_n)\\ &\qquad = \left(\int_{X_1} \cdots \left(\int_{X_2} \cdots \left(\int_{X_n} g(x_1, \dots, x_n)\mu_1(\,dx_1)\right) \mu_2(\,dx_2)\right) \dots \mu_n(\,dx_n)\right).\end{aligned}$$

Ez az azonosság úgy értendő, hogy az azonosság két oldalán lévő integrál illetve szukcesszív integrál egyszerre létezik.

Külön megfogalmazzuk ezt az eredményt abban a speciális esetben, amikor a számegyenesen vett valószínűségi mértékek szorzata szerinti integrált tekintünk, mivel mi leggyakrabban a Fubini tétel ilyen alakú megfogalmazásával találkozunk.

Fubini tétel speciális esete eloszlásfüggvényekre. *Legyenek $F_j(\cdot)$, $1 \le j \le n$, eloszlásfüggvények a számegyenesen, és legyen $g(x_1, \dots, x_n)$ (mérhető) n-változós függvény. Ekkor*

$$\begin{aligned}&\int g(x_1, x_2, \dots, x_n) F_1(\,dx_1) F(\,dx_2) \dots F_n(\,dx_n)\\ &\qquad = \left(\int \cdots \left(\int \left(\cdots \int g(x_1, \dots, x_n) F_1(\,dx_1)\right) F_2(\,dx_2)\right) \dots F_n(\,dx_n)\right).\end{aligned}$$

Ez az azonosság úgy értendő, hogy az azonosság két oldalán lévő integrál illetve szukcesszív integrál egyszerre létezik.

Érdemes külön megfogalmazni ennek az azonosságnak a következő fontos speciális esetét. Ha $g(x_1, \dots, x_n) = g_1(x_1) \cdots g_n(x_n)$ alakú speciális függvények integrálját tekintjük, akkor a következő azonosságot írhatjuk fel:

$$\begin{aligned}&\int g_1(x_1) g_2(x_2) \cdots g_k(x_n) F_1(\,dx_1) F_2(\,dx_2) \dots F_n(\,dx_n)\\ &\qquad = \int g(x_1) F_1(\,dx_1) \int g_2(x_2) F_2(\,dx_2) \cdots \int g_k(x_n) F_k(\,dx_n) = \prod_{j=1}^{n} \int g_j(x) F_j(\,dx).\end{aligned}$$

Megjegyzés: A fenti tételben szereplő integrálok úgy értendőek, hogy amikor $F_j(\,dx_j)$ szerinti integrált tekintjük akkor az F_j eloszlásfüggvény által meghatározott μ_{F_j} Lebesgue–Stieltjes mérték szerinti Lebesgue integrált tekintjük. Hasonlóan az

$$\int g_1(x_1)g_2(x_2)\cdots g_k(x_n)F_1(\,dx_1)F_2(\,dx_2)\dots F_n(\,dx_n)$$

integrál azt jelenti, hogy tekintjük az $F(x_1,\dots,x_n) = F_1(x_1)F_2(x_2)\dots F_n(x_n)$ n-változós eloszlásfüggvényt, és a g függvényt az általa meghatározott μ_F mérték szerint integráljuk. Természetesen, ha ,,szép" függvényeket integrálunk, akkor az itt szereplő Lebesgue integrálok helyettesíthetőek a velük egyenlő Riemann–Stieltjes integrálokkal.

Annak érdekében, hogy lássuk, hogy a fenti megjegyzés hasznos oldjuk meg a következő feladatot.

Feladat:

Legyenek $F_1(x),\dots,F_n(x)$ egyváltozós eloszlásfüggvények. Ekkor az $F(x_1,\dots,x_n) = F_1(x_1)\cdots F_n(x_n)$ függvény n-változós eloszlásfüggvény.

Most megismételünk néhány a bevezető valószínűségszámítás előadásban szereplő fontos fogalmat és eredményt.

Valószínűségi változók függetlenségének a definiciója. *Legyenek $\xi_1,\dots,\xi_n$ valószínűségi változók egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi változók függetlenek, ha minden $x_1,\dots,x_n$ valós számra*

$$P(\xi_1 < x_1,\dots,\xi_n < x_n) = P(\xi_1 < x_1)\cdots P(\xi_n < x_n).$$

Ennek a fogalomnak a többdimenziós változata a következő:

Többdimenziós valószínűségi változók függetlenségének a definiciója. *Legyenek $\xi^{(1)} = (\xi_{1,1},\dots,\xi_{1,k})$, $\dots$, $\xi^{(n)} = (\xi_{n,1},\dots,\xi_{n,k})$, k-dimenziós valószínűségi változók (vektorok) egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi vektorok függetlenek, ha tetszőleges $x^{(1)} = (x_{1,1},\dots,x_{1,k})$, $\dots$, $x^{(n)} = (x_{n,1},\dots,x_{n,k})$, k-dimenziós vektorokra*

$$\begin{aligned}&P\left(\xi_{1,1} < x_{1,1},\dots,\xi_{1,k} < x_{1,k},\ \dots,\ \xi_{n,1} < x_{n,1},\dots,\xi_{n,k} < x_{n,k}\right)\\&\quad = P\left(\xi_{1,1} < x_{1,1},\dots,\xi_{1,k} < x_{1,k}\right)\cdots P\left(\xi_{n,1} < x_{n,1},\dots,\xi_{n,k} < x_{n,k}\right).\end{aligned}$$

Valószínűségi változók végtelen sorozatának függetlensége. *Legyen $\xi_1,\xi_2,\dots$ (egy vagy több-dimenziós) valószínűségi változók sorozata egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi változók függetlenek, ha minden n egész számra a $\xi_1,\dots,\xi_n$ valószínűségi változók függetlenek.*

Kimondunk néhány eredményt, melyek a Fubini tétel következményei.

Tétel A. *Legyenek $\xi_1, \dots, \xi_n$ független valószínűségi változók, $B_1, \dots, B_n$ a számegyenes Borel mérhető részhalmazai. Ekkor*

$$P(\xi_1 \in B_1, \dots, \xi_n \in B_n) = P(\xi_1 \in B_1) \cdots P(\xi_n \in B_n).$$

Tétel B. *Legyenek $\xi_1, \dots, \xi_n$ független k-dimenziós) valószínűségi változók, (k tetszőleges pozitív egész szám, és $g_1, \dots, g_n$ valós értékű függvények a k-dimenziós euklidészi téren. Ekkor*

$$E\left(g_1(\xi_1(\omega) \cdots g_n(\xi_n(\omega))\right) = Eg_1(\xi_1(\omega)) \cdots Eg_n(\xi_n(\omega)).$$

A fenti azonosság úgy értendő, hogy amennyiben az egyik oldalon szereplő kifejezés értelmes, akkor a másik oldalon szereplő kifejezés is értelmes, és egyenlő vele.

Feladat:

Mutassuk meg, hogy a Tétel A. következik a Tétel B.-ből.

Felidézünk néhány további eredményt és fogalmat a bevezető valószínűségszámítás előadásból.

Tétel. *Legyenek $\xi_1, \dots, \xi_n, \xi_{n+1}, \dots, \xi_{n+k}$ független valószínűségi változók egy $\Omega, \mathcal{A}, P)$ valószínűségi mezőn, $f(x_1, \dots, x_k)$ tetszőleges (mérhető) k-változós függvény. Definiáljuk az $\eta = f(\xi_{n+1}, \dots, \xi_{n+k})$ valószínűségi változót, amelyik nem függ az első n $\xi_1, \dots, \xi_n$ valószínűségi változótól. Ekkor a $\xi_1, \dots, \xi_n, \eta$ valószínűségi változók függetlenek.*

Feladat:

Bizonyítsuk be a fenti tételt a Fubini tétel segítségével.

Felidézzük a következő fogalmakat és eredményeket is.

A szórásnégyzet definiciója: *Legyen ξ olyan valószínűségi változó, melyre $E\xi^2 < \infty$. Ekkor a ξ valószínűségi változó szórásnégyzetét az*

$$\operatorname{Var} \xi = E\left(\xi - E\xi\right)^2$$

képlet definiálja. Ha $E\xi^2 = \infty$ akkor a $\operatorname{Var} \xi$ szórásnégyzetet nem definiáljuk vagy azt mondjuk, hogy $\operatorname{Var} \xi = \infty$.

A kovarianciafüggvény definiciója. *Legyen ξ és η két valószínűségi változó, melyek mindegyikének létezik szórásnégyzete, azaz $E\xi^2 < \infty$ és $E\eta^2 < \infty$. Ekkor a ξ és η kovariancifüggvényét $\operatorname{Cov}(\xi, \eta)$-t a*

$$\operatorname{Cov}(\xi, \eta) = E\left[(\xi - E\xi)(\eta - E\eta)\right]$$

kifejezés definiálja. Ha a $E\xi^2 < \infty$ és $E\eta^2 < \infty$ feltételek valamelyike nem teljesül, akkor nem definiáljuk a

$$\mathrm{Cov}\,(\xi,\eta) = E\xi\eta - E\xi E\eta.$$

$\mathrm{Cov}\,(\xi,\eta)$ *kovariancifüggvényt.*

Lemma.

$$\mathrm{Var}\,\xi = E\xi^2 - (E\xi)^2, \qquad \mathrm{Cov}\,(\xi,\eta) = E\xi\eta - E\xi E\eta.$$

Továbbá, $(E\xi)^2 \le E\xi^2$, *és* $|E\xi E\eta| \le \sqrt{E\xi^2 E\eta^2}$.

Tétel. *Ha $\xi_1, \xi_2, \ldots, \xi_n$ független valószínűségi változók, $c_1, c_2, \ldots, c_n$ valós számok, akkor*

$$\mathrm{Var}\,(c_1\xi_1 + c_2\xi_2 + \cdots + c_n\xi_n) = c_1^2 \mathrm{Var}\,\xi_1 + c_2^2 \mathrm{Var}\,\xi_2 + \cdots + c_n^2 \mathrm{Var}\,\xi_n.$$

Általánosabban, ha $\xi_1, \ldots, \xi_n$ tetszőleges valószínűségi változók ugyanazon a valószínűségi mezőn, melyek mindegyikének létezik szórásnégyzete, $c_1, \ldots, c_n$ valós számok, akkor

$$\begin{aligned}
\mathrm{Var}\left(\sum_{j=1}^n c_j\xi_j\right) &= \sum_{j=1}^n c_j^2 \mathrm{Var}\,\xi_j + 2\sum_{1\le j<k\le n} c_jc_k(E\xi_j\xi_k - E\xi_j E\xi_k) \\
&= \sum_{j=1}^n c_j^2 \mathrm{Var}\,\xi_j + 2\sum_{1\le j<k\le n} c_jc_k\,\mathrm{Cov}\,(\xi_j,\xi_k) \\
&= \sum_{j=1}^n c_j^2 \mathrm{Var}\,\xi_j + \sum_{1\le j,k\le n,\ j\ne k} c_jc_k\,\mathrm{Cov}\,(\xi_j,\xi_k).
\end{aligned}$$

Feladat:

Bizonyítsuk be a lemmában szereplő $(E\xi)^2 \le E\xi^2$ és $|E\xi E\eta| \le \sqrt{E\xi^2 E\eta^2}$ egyenlőtlenségeket, melyeket egyébként Schwarz vagy Cauchy–Schwartz egyenlőtlenségnek is szoktak nevezni.

Feladat:

Mutassunk példát olyan $\xi_1, \ldots, \xi_n$ valószínűségi változókra, melyek nem függetlenek, viszont teljesítik a $\mathrm{Cov}\,(\xi_j,\xi_k) = 0$ azonosságot minden $1 \le j < k \le n$ számpárra, ezért

$$\mathrm{Var}\,(c_1\xi_1 + c_2\xi_2 + \cdots + c_n\xi_n) = c_1^2 \mathrm{Var}\,\xi_1 + c_2^2 \mathrm{Var}\,\xi_2 + \cdots + c_n^2 \mathrm{Var}\,\xi_n$$

minden valós $c_1, \ldots, c_n$ számsorozatra.

Feladat:

Mutassunk példát három olyan ξ_1, ξ_2, ξ_3 (nem független) valószínűségi változóra, melyekre $E\xi_1\xi_2 = E\xi_1E\xi_2$, $E\xi_1\xi_3 = E\xi_1E\xi_3$, $E\xi_2\xi_3 = E\xi_2E\xi_3$, de $E\xi_1\xi_2\xi_3 \ne E\xi_1E\xi_2E\xi_3$.

A Markov és Csebisev egyenlőtlenség. *Legyen ξ nem negatív valószínűségi változó, azaz olyan valószínűségi váltoó, melyre $P(\xi \ge 0) = 1$. Ekkor*

$$P(\xi \ge x) \le \frac{E\xi}{x}$$

minden $X > 0$ számra. (Ez a Markov egyenlőtlenség.)

Legyen ξ olyan valószínúségi változó, melyre $E\xi^2 < \infty$. Ekkor

$$P(|\xi - E\xi| \ge x) \le \frac{\operatorname{Var}\xi}{x^2}$$

minden $x > 0$ számra. Ez a Csebisev egyenlőtlenség.

A Csebisev egyenlőtlenség következménye. *Legyenek $\xi_1, \dots, \xi_n$ független valószínűségi változók, melyek mindegyike teljesíti az $E\xi_j < \infty$, $1 \le j \le n$, feltételt. E valószínűségi változók átlaga tetszőleges $\varepsilon > 0$ számra teljesíti az alábbi egyenlőtlenséget:*

$$P\left(\left|\frac{1}{n}\sum_{j=1}^{n}(\xi_j - E\xi_j)\right| > \varepsilon\right) \le \frac{\operatorname{Var}\left(\sum\limits_{j=1}^{n}\xi_j\right)}{n^2\varepsilon^2} = \frac{\sum\limits_{j=1}^{n}\operatorname{Var}\xi_j}{n^2\varepsilon^2}.$$

Felidézzük két sűrűségfüggvény konvoluciójának a fogalmát, illetve azt az eredményt, mely megadja, hogyan lehet két független (sűrűségfüggvénnyel rendelkező) valószínűségi változó összegének a sűrűségfüggvényét kiszámolni konvolució segítségével. Megjegyzem, hogy ez az eredmény is (melynek bizonyítását elhagyom), a Fubini tételen alapul. Ugyanis a Fubini tétel lehetővé teszi, hogy két független valószínűségi változó összegének az eloszlásffüggvényét felírjuk az eloszlásfüggvények szorzata szerinti integrál segítségével, majd ezt differeniálva megkapjuk a sűrűségfüggvényt. Ezen eljárást végrehajtva, alkalmas helyettesítések segítségével némi számolással megkapjuk az alább megfogalmazott eredményt.

(Sűrűség)függvények konvoluciójának a definiciója. *Legyen $f(\cdot)$ és $g(\cdot)$ két sűrűségfüggvény a számegyenesen, általánosabban integrálható függvények, azaz tegyük fel, hogy $\int_{-\infty}^{\infty}|f(u)|\,du < \infty$ és $\int_{-\infty}^{\infty}|g(u)|\,du < \infty$. Az $f(\cdot)$ és $g(\cdot)$ függvények $f * g(\cdot)$ konvoluciója az*

$$f * g(x) = \int_{-\infty}^{\infty} f(u)g(x-u)\,du, \quad -\infty < x < \infty,$$

függvény.

Megjegyzés: Egyszerű (lineáris) transzformációval kapjuk, hogy a konvoluciót másképp is kiszámolhatjuk. Ez mutatja, hogy a konvolucióban résztvevő függvények szimmetrikus szerepet játszanak.

$$\begin{aligned} f * g(x) &= \int_{-\infty}^{\infty} f(u)g(x-u)\,du = \int_{-\infty}^{\infty} f(x-u)g(u)\,du \\ &= \int_{-\infty}^{\infty} f\left(\frac{x}{2}-u\right) g\left(\frac{x}{2}+u\right)\,du, \quad -\infty < x < \infty. \end{aligned}$$

Tétel független valószínűségi változók összegének a sűrűségfüggvényéről. *Legyen ξ és η két független valószínűségi változó $f(\cdot)$ és $g(\cdot)$ sűrűségfüggvénnyel. Ekkor a $\xi + \eta$ összegnek is létezik sűrűségfüggvénye, és az az*

$$f * g(x) = \int_{-\infty}^{\infty} f(u)g(x-u)\,du, \quad -\infty < x < \infty$$

függvény.