Alapvető ismeretek valószínűségi problémák számolásánál

Valószínűségi problémák vizsgálata érdekében bevezettük az $(\Omega, \mathcal{A}, P)$ valószínűségi mező fogalmát. Egyik alapvető probléma valószínűségi változók várható értékének a definiciója és kiszámítása volt. Maga a definició nagyon egyszerű, ha ismerjük a Lebesgue integrál fogalmát.

Valószínűségi változó várható értékének a definiciója. *Egy $(\Omega, \mathcal{A}, P)$ téren definiált $\xi(\omega)$ valószínűségi változónak (azaz mérhető függvénynek) a várható értéke az $\int \xi(\omega)P(\,\omega)$ Lebesgue integrállal egyenlő, feltéve, hogy $\int |\xi(\omega)|P(\,\omega) < \infty$. (Ellenkező esetben nem definiáljuk a várható értéket.)*

Ebből a definicióból azonnal kiolvasható a várható érték néhány fontos tulajdonsága. Így például a várható érték additivítása az integrál additivításának egyszerű következménye. De ahhoz, hogy jól tudjunk számolni várható értékkel további ismeretekre van szükségünk. Az esetek többségében a várható értéket közvetlenül nem a definiciója alapján számoljuk ki, hanem bevezetjük egy valószínűségi változó eloszlásfüggvényét, és olyan tételeket bizonyítunk be, amelyek lehetővé teszik a várható érték kiszámolását az eloszlásfüggvény segítségével. Ahhoz, hogy jól tudjunk számolni várható értékekkel meg kell érteni a várható érték eredeti definiciója és a várható érték eloszlásfüggvények segítségével való kiszámolásáról szóló képletek kapcsolatát.

E kapcsolat megértésében a mértékelmélet alább ismertetett mértéktartó transzformációk szerinti integráltranszformációkról szóló tételének van kulcsfontosságú szerepe. E tétel megfogalmazása előtt bevezetem a következő fogalmat.

Mértéktartó transzformáció definiciója. *Legyen adva két $(X, \mathcal{X}, \gamma)$ és $(Y, \mathcal{Y}, \nu)$ mértéktér, azaz egy X illetve Y alaphalmaz, ezeknek bizonyos kijelölt, mérhető részhalmazainak $\mathcal{X}$ illetve $\mathcal{Y}$ rendszere, amelyek σ-algebrát alkotnak, és egy γ illetve ν mérték az $\mathcal{X}$ és $\mathcal{Y}$ σ-algebrán. Egy minden $x \in X$ pontban definiált T, $T(x) = y$, $x \in X$, $y \in Y$ transzformációt mértéktartónak nevezünk, ha minden $B \in \mathcal{Y}$ halmaz $T^{-1}(B) = \{x\colon Tx \in B\}$ ősképe mérhető halmaz, azaz $T^{-1}(B) \in \mathcal{X}$, és $\gamma(T^{-1}(B)) = \nu(B)$ minden $B \in \mathcal{Y}$ halmazra.*

Tétel mértéktartó transzformációk szerinti integráltranszformációkról. *Legyen adva egy T mértéktartó transzformáció egy $(X, \mathcal{X}, \gamma)$ mértéktérről egy $(Y, \mathcal{Y}, \nu)$ mértéktérre, valamint egy $\mathcal{Y}$ mérhető $f(y)$ függvény az $(Y, \mathcal{Y})$ téren. Vezessük be a $g(x) = f(Tx)$ függvényt az $(X, \mathcal{X})$ téren. Ekkor az $f(Tx)$ függvény $\mathcal{X}$ mérhető, és érvényes az*

$$\int_Y f(y)\nu(\,dy) = \int_X f(Tx)\gamma(\,dx)$$

azonosság. Ez úgy értendő, hogy az azonosság két oldalán levő integrálok egyszerre abszolut konvergensek, azaz az $\int_Y |f(y)|\nu(\,dy) < \infty$ reláció akkor és csak akkor teljesül, ha $\int_X |f(Tx)|\gamma(\,dx) < \infty$.

Egy valószínűségi változó vagy véletlen vektor eloszlását és eloszlásfüggvényét úgy definiálják, hogy ez lehetővé tegye a várható érték kiszámítását az előző tétel alapján az eloszlás szerinti integrál segítségével, ha megfelelő módon definiáljuk a tétel alkalmazásában a mértéktereket és T transzformációt.

Eloszlás és eloszlásfüggvény definiciója. *Legyen* $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))$ *egy* k*-dimenziós valószínűségi változó egy* $(\Omega, \mathcal{A}, P)$ *valószínűségi mezőn, azaz az* $(\Omega, \mathcal{A}, P)$ *mértéktér leképezése az* $(R^k, \mathcal{B}^k)$ *térbe, ahol* $\mathcal{B}^k$ *az* R^k k*-dimenziós Euklideszi tér Borel* σ*-algebrája. A* ξ *valószínűségi változó eloszlását úgy definiáljuk, mint a következő* μ *halmazfüggvényt a* $\mathcal{B}^k$ σ*-algebrán:* $\mu(B) = P(\{\omega\colon \xi(\omega) \in B\})$*, minden* $B \in \mathcal{B}$ *halmazra. A* ξ *valószínűségi változó eloszlását az*

$$F(x_1, \dots, x_k) = P(\{\omega\colon \xi_1(\omega) < x_1, \dots, \xi_k(\omega) < x_k\})$$

képlet adja meg minden $(x_1, \dots, x_k) \in R^k$ *vektorra.*

Tétel egy valószínűségi változó eloszlása és eloszlásfüggvényének alapvető tulajdonságairól. *Legyen* $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))$ *egy* k*-dimenziós valószínűségi változó egy* $(\Omega, \mathcal{A}, P)$ *valószínűségi mezőn. Ekkor e valószínűségi változó előbb definiált* μ *eloszlása valószínűségi mérték az* $(R^k, \mathcal{B}^k)$ *téren, és* $\xi(\omega)$ *mértéktartó transzformáció az* $(\Omega, \mathcal{A}, P)$ *valószínűségi mezőről az* $(R^k, \mathcal{B}^k, \mu)$ *mértéktérre.*

Egy ξ *valószínűségi változó* μ *eloszlása meghatározza* ξ *eloszlásfüggvényét az*

$$F(x_1, \dots, x_k) = \mu(B(x_1, \dots, x_k))$$

képlet alapján, ahol $B(x_1, \dots, x_k) = \{(u_1, \dots, u_k)\colon u_1 < x_1, \dots, u_k < x_k\}$*. Megfordítva, a mértékelmélet bizonyos eredményei alapján a* ξ *valószínűségi változó* F *eloszlásfüggvénye egyértelműen meghatározza a* ξ *valószínűségi változó* μ *eloszlását.*

A fenti eredmények alapján, tekinthetjük a $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))$ vektor értékű valószínűségi változót, mint az $(\Omega, \mathcal{A}, P) = (X, \mathcal{X}, \alpha)$ mérhető tér mértéktartó transzformációját az $(Y, \mathcal{Y}, \nu) = (R^k, \mathcal{B}^k, \mu)$ térre, ahol μ a ξ véletlen vektor eloszlása. Ilyen szereposztással véve egy $f(x_1, \dots, x_k)$ mérhető függvényt az $(R^k, \mathcal{B}^k)$ téren és alkalmazva a *mértéktartó transzformációk szerinti integráltranszformációkról* szóló tételt a következő azonosságot kapjuk.

Tétel véletlen vektor várható értékének kiszámításáról. *Legyen egy* $(\xi_1, \dots, \xi_k)$ *véletlen vektor eloszlása* μ*. Ekkor érvényes az*

$$\int f(x_1, \dots, x_k)\mu(\,dx_1, \dots, \,dx_k) = \int f(\xi_1(\omega), \dots, \xi_k(\omega))P(\,d\omega) = Ef(\xi_1, \dots, \xi_k)$$

azonosság.

Ez azt jelenti, hogy egy véletlen vektor valamely függvényének a várható értékét ki tudjuk számolni a véletlen vektor eloszlásának a segítségével. Sőt, mivel az eloszlásfüggvény meghatározza az eloszlást, ezért elég az eloszlásfüggvényt ismerni. A továbbiakban a következő, az irodalomban elterjedt jelölést használom. Ha ismerem egy valószínűségi változó, vagy általánosabban egy véletlen vektor F eloszlásfüggvényét, akkor $dF(x_1,\dots,x_k)$ vagy $F(\,dx_1,\dots,\,dx_k)$-val jelölöm az ezen eloszlásfüggvény szerint egyértelműen meghatározott $\mu(\,dx_1,\dots,\mu\,dx_k)$ (úgynevezett Stieltjes mérték szerinti) integrált.

A mértékelmélet egy további fontos, más előadásjegyzetemben tárgyalt eredménye annak a pontos leírása, hogy milyen tulajdonságokkal kell rendelkeznie egy függvénynek ahhoz, hogy megjelenhessen, mint egy alkalmas valószínűségi mezőben megfelelő módon definiált valószínűségi változó eloszlásfüggvénye.

A további számunkra fontos kérdések azzal függnek össze, hogy hogyan lehet jól kiszámolni egy eloszlás szerinti integrált fontos esetekben. Sok fontos esetben olyan eloszlások szerint kell integrálokat számolni, amelyeknek létezik sűrűségfüggvénye. Az ilyen integrálok vizsgálata érdekében fel kell idézni a sűrűségfüggvények definicióját és legfontosabb tulajdonságait.

Többváltozós sűrűségfüggvények legfontosabb tulajdonságai.

Többváltozós sűrűségfüggvény definiciója. *Legyen $F(x_1,\dots,x_k)$ egy k-változós eloszlásfüggvény. Azt mondjuk, hogy ennek az eloszlásfüggvénynek az $f(x_1,\dots,x_k)$ függvény a sűrűségfüggvénye, ha teljesül az*

$$F(x_1,\dots,x_k)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_k} f(u_1,\dots,u_k)\,du_1\dots\,du_k$$

azonosság minden valós $x_1,\dots,x_k$ szám-k-asra.

Ha egy $(\xi_1,\dots,\xi_k)$ véletlen vektornak az $F(x_1,\dots,x_k)$ függvény az eloszlásfüggvénye, és az $f(x_1,\dots,x_k)$ függvény az $F(x_1,\dots,x_k)$ eloszlásfüggvény sűrűségfüggvénye, akkor időnként azt mondjuk, hogy $f(x_1,\dots,x_k)$ a $(\xi_1,\dots,\xi_k)$ véletlen vektor sűrűségfüggvénye.

Természetesen, az egyváltozós esethez hasonlóan, magasabb dimenzióban sem létezik minden eloszlásfüggvénynek sűrűségfüggvénye. Másrészt az eloszlásfüggvényekhez hasonlóan a sűrűségfüggvényeknek is van jellemzése.

Sűrűségfüggvény jellemzése. *Egy $f(x_1,\dots,x_k)$ k-változós függvény akkor és csak akkor sűrűségfüggvénye egy alkalmas eloszlásfüggvénynek, ha $f(x_1,\dots,x_k)\ge 0$ (majdnem) minden $(x_1,\dots,x_k)$ pontra, és*

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

Eloszlás és sűrűségfüggvény kapcsolata. *Az eloszlás és sűrűségfüggvény kölcsönösen meghatározzák egymást. Elég sima (elég sokszor differenciálható) $F(x_1,\dots,x_k)$ eloszlásfüggvény $f(x_1,\dots,x_k)$ sűrűségfüggvénye létezik, és az kiszámolható az*

$$f(x_1,\dots,x_k)=\frac{\partial^k F(x_1,\dots,x_k)}{\partial x_1\cdots\partial x_k}$$

képlet segítségével.

A következő egyszerű állítás nagyon fontos lesz számunkra a későbbiekben.

Független valószínűségi változókból álló véletlen vektor sűrűségfüggvénye. *Legyenek $\xi_1,\dots,\xi_k$ független valószínűségi változók $F_j(x)$ eloszlás és $f_j(x)$ sűrűségfüggvénnyel, $1\le j\le k$. Ekkor a $(\xi_1,\dots,\xi_k)$ véletlen vektor eloszlásfüggvénye az*

$$F(x_1,\dots,x_k)=F(x_1)\cdots F(x_k)$$

és (létező) sűrűségfüggvénye az

$$f(x_1,\dots,x_k)=f(x_1)\cdots f(x_k)$$

függvény.

Ugyancsak fontos számunkra a következő eredmény, amely azt adja meg, hogy hogyan lehet kiszámolni sok minket érdeklő mennyiséget sűrűségfüggvények segítségével.

Véletlen vektor függvényei és azok várható értékének kiszámolása sűrűségfüggvény segítségével. *Legyen egy $(\xi_1,\dots,\xi_k)$ véletlen vektornak $f(x_1,\dots,x_k)$ sűrűségfüggvénye. Ekkor tetszőleges (Borel mérhető) halmazra a k-dimenziós térben*

$$P((\xi_1,\dots,\xi_k)\in B)=\left(\int\cdots\int\right)_B f(u_1,\dots,u_k)\,du_1\dots du_k$$

Általánosabban, legyen $g(x_1,\dots,x_k)$ tetszőleges (mérhető) k-változós függvény. Ekkor

$$Eg(\xi_1,\dots,\xi_k)=\int\cdots\int g(u_1,\dots,u_k)f(u_1,\dots,u_k)\,du_1\dots du_k,$$

ha az $Eg(\xi_1,\dots,\xi_k)$ várható érték létezik, (azaz véges szám).

Ahhoz, hogy a fenti eredményeket használni tudjuk, tudnunk kell többváltozós integrálokkal számolni. Ehhez egyrészt fontos a Fubini tétel, amely szerint többváltozós integrálokat ki lehet számolni szukcesszív egyváltozós integrálás segítségével. Matematikai képletben megfogalmazva:

$$\int\cdots\int f(x_1,\dots,x_k)\,dx_1\dots dx_k=\int\dots\left(\int\left(\int f(x_1,\dots,x_k)\,dx_1\right)dx_2\right)\dots dx_k.$$

Az előbb megfogalmazott azonosság valójában a Fubini tétel speciális esete. A Fubini tétel azt mondja ki, hogy ha adva van egy szorzattér, azon egy szorzatmérték, akkor egy a szorzattéren definált függvény szorzatmérték szerinti integrálját ki lehet számolni szukcessziv integrálással is. Pontosabban megfogalmazva

Fubini tétel. *Legyenek adva $(X_j,\mathcal{X}_j,\mu_j)$, $1\le j\le k$, mértékterek és tekintsük ezek direkt szorzatát. Részletesebben kifejtve tekintsük azt az $(X,\mathcal{X},\mu)$ mértékteret, amelyre $X=X_1\times\dots\times X_k$, azaz X az $(x_1,\dots,x_k)$, $x_j\in X_j$, $1\le j\le k$, alakú pontokból áll, $\mathcal{X}$*

az $X_1 \times \cdots \times X_{j-1} \times B_j \times X_j \times \cdots \times X_k$*,* $B_j \in \mathcal{X}_j$*,* $1 \le j \le k$*, alakú hengerhalmazok által generált legszűkebb* σ*-algebra,* μ *az a (bizonyos mértékelméleti eredmények alapján létező és egyértelműen meghatározott) mérték a* $\mathcal{X}$ σ*-algebrán, amelyre* $\mu(B_1 \times \cdots \times B_k) = \mu_1(B_1) \cdots \mu_k(B_k)$ *minden* $B_j \in \mathcal{B}_j$*,* $1 \le j \le k$*, halmazsorozatra. Minden az* $(X, \mathcal{X}, \mu)$ *mértéktéren vett integrálható* $f(x) = f(x_1, \dots, x_k)$ *függvényre*

$$\int f(x_1, \dots, x_k)\, d\mu(x_1, \cdots, x_k) = \int \cdots \left(\int \left(\int f(x_1, \dots, x_k)\, d\mu_1(x_1) \right) d\mu_2(x_2) \right) \cdots\, d\mu_k(x_k).$$

Speciálisan, ha $f(x_1, \dots, x_k) = f_1(x_1) \cdots f_k(x_k)$*, akkor*

$$\int f(x_1) \cdots f(x_k)\, d\mu(x_1, \dots, x_k) = \int f(x_1)\, d\mu_1(x_1) \int f(x_2)\, d\mu_2(x_2) \cdots \int f(x_k)\, d\mu_k(x_k).$$

Ez az eredmény tartalmazza az előzőleg megfogalmazott eredményt, ha a k-dimenziós Euklideszi teret tekintjük a Borel σ-algebrával és a Lebesgue mértékkel, mint a számegyenesen vett Lebesge mérték k-szors direkt szorzatát önmagával.

Ezenkívül szükségünk lesz az egyváltozós integrálok kiszámolásában fontos szerepet játszó helyettesítéses integrál többváltozós megfelelőjére is. Ismertetem ezt az eredményt.

A következő problémával foglalkozunk. Ha adva van az n-dimenziós tér egy A tartományának szép, sima és invertálható $y_k = T_k(x_1, \dots, x_n)$, $1 \le k \le n$, leképezése az n-dimenziós tér egy másik B tartományába valamint egy

$$\int_B f(y_1, \dots, y_n)\, dy_1 \dots dy_n$$

alakú integrál a B tartományon, akkor hogyan tudjuk ezt az integrált átírni az $y_k = T_k(x_1, \dots, x_k)$, $k = 1, \dots, n$, leképezések segítségével, mint egy alkalmas az A tartományon értelmezett függvény integrálját? Célunk az, hogy ily módon a vizsgálandó integrált egy könnyebben kezelhető integrállá alakítsuk át.

Annak érdekében, hogy az erre a kérdésre adott választ meg tudjam fogalmazni először felidézem egy sima transzformáció Jacobian-jának a definicióját.

Jacobian definiciója. *Legyen* $y_k = T_k(x_1, \dots, x_n)$*,* $1 \le k \le n$*, az* n*-dimenziós tér egy tartományának sima transzformációja az* n*-dimenziós tér egy másik tartományába. Vezessük be a* $\mathbf{T} = (T_1, \dots, T_n)$ *jelölést. A* $\mathbf{T}$ *transzformáció* $\mathcal{J}(\mathbf{T}(x_1, \dots, x_n))$ *Jacobian-ja egy* $(x_1, \dots, x_n)$ *pontban a*

$$\left(\frac{\partial T_k(x_1, \dots, x_k)}{\partial x_l} \right), \quad 1 \le l, k \le n,$$

$n \times n$-es (az $(x_1, \dots, x_n)$ pontban vett derivált) mátrix determinánsának az abszolut értéke.

(A Jacobian szemléletes tartalma: Ez adja meg, hogy az $(x_1, \dots, x_n)$ pont kis környezetének a térfogatát a $\mathbf{T} = (T_1, \dots, T_n)$ transzformáció hányszorosára nagyítja.)

Integráltranszformációról szóló tétel. *Legyen adva az n-dimenziós tér egy A tartományának egy sima $y_k = T_k(x_1, \dots, x_n)$, $1 \le k \le n$, transzformáltja az n-dimenziós tér egy másik B tartományába, amelyik invertálható, azaz az $y_k = T_k(x_1, \dots, x_n)$, $1 \le k \le n$, egyenletrendszernek egyetlen $(x_1, \dots, x_n) \in A$ megoldása van minden $(y_1, \dots, y_n) \in B$ pontra. Legyen továbbá adva egy (integrálható) $f(y_1, \dots, y_n)$ függvény a B tartományon. Ekkor*

$$
\begin{aligned}
&\int_B f(y_1, \dots, y_n)\, dy_1 \dots dy_n \\
&\qquad = \int_A f(T_1(x_1, \dots, x_n), \dots, T_n(x_1, \dots, x_n)) \mathcal{J}(\mathbf{T}(x_1, \dots, x_n))\, dx_1 \dots dx_n,
\end{aligned}
$$

ahol $\mathcal{J}(\mathbf{T}(x_1, \dots, x_n))$ jelöli a $\mathbf{T}(x_1, \dots, x_n)$ leképezés Jacobianját.

Az előbbi integrál transzformációs képlet a következő kapcsolatban van az ismertetés elején tárgyalt *mértéktartó transzformációk szerinti integráltranszformációkról* szóló tétellel. Tekintsük az $(A, \mathcal{B}^n, \mu)$ és $(B, \mathcal{B}^n, \lambda)$ mértéktereket valamint egy $(y_1, \dots, y_n) = \mathbf{T}(x_1, \dots, x_n)$ transzformációt az A halmazból a B halmazba, ahol A és B az *integráltranszformációról szóló tétel*-ben szereplő halmazok az R^n Euklideszi térben, $\mathbf{T}$ az ott bevezetett (sima) transzformáció az A halmazból a B halmazba, $\mathcal{B}$ a Borel σ-algebra megszorítása az A illetve a B halmazra, λ a Lebesgue mérték az R^n térnek a B halmazra vett megszorításán, μ a Lebesgue mérték szerint $\mathcal{J}(\mathbf{T}(x_1, \dots, x_n))$ sűrűségfüggvénnyel rendelkező mérték az A halmaz (mérhető részhalmazain, ahol $\mathcal{J}(\mathbf{T}(x_1, \dots, x_n))$ a $\mathbf{T}$ transzformáció Jacobianja, azaz

$$
\mu(D) = \int_D \mathcal{J}(\mathbf{T}(x_1, \dots, x_n))\, dx_1 \dots dx_n \quad \text{minden } D \subset A \text{ halmazra.}
$$

Be lehet látni, és ez az *integráltranszformációról szóló tétel* bizonyításának a lényege, hogy $\mathbf{T}$ mértéktartó transzformáció az $(A, \mathcal{B}^n, \mu)$ térből a $(B, \mathcal{B}^n, \lambda)$ térbe. Ezt felhasználva a tétel állítása következik a *tétel mértéktartó transzformációk szerinti integráltranszformációkról* eredményéből.

Eredmények arról, hogy hogyan számítjuk ki valószínűségi változók összegének vagy szorzatának a várható értékét, szórásnégyzetét.

Tétel valószínűségi változók összegének a várható értékéről. *Legyenek $\xi_1, \dots,$ ξ_n olyan valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyekre $E|\xi_j| < \infty$ minden $1 \le j \le n$ indexre, és $a_1, \dots, a_n$ tetszőleges valós számok. Ekkor*

$$
E(a_1\xi_1 + \dots + a_n\xi_n) = \sum_{j=1}^{n} a_j E\xi_j.
$$

Fontos észrevétel. A fenti azonosság, azaz a várható érték additivitása, akkor is érvényes, ha a tekintett valószínűségi változók nem függetlenek.

Tétel független valószínűségi változók összegének a szórásnégyzetéről. *Legyenek $\xi_1, \dots, \xi_n$ olyan független valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyekre $E\xi_j^2 < \infty$ minden $1 \le j \le n$ indexre, és $a_1, \dots, a_n$ tetszőleges valós számok. Ekkor*

$$\mathrm{Var}\,(a_1\xi_1 + \cdots + a_n\xi_n) = \sum_{j=1}^{n} a_j^2 \mathrm{Var}\,\xi_j.$$

Továbbá, $\mathrm{Var}\,\xi_j = E(\xi_j - E\xi_j)^2 = E\xi_j^2 - (E\xi_j)^2$.

Tétel valószínűségi változók összegének a szórásnégyzetéről. *Legyenek $\xi_1, \dots,$ ξ_n olyan (nem feltétlenül független) valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyekre $E\xi_j^2 < \infty$ minden $1 \le j \le n$ indexre, és $a_1, \dots, a_n$ tetszőleges valós számok. Ekkor*

$$\begin{aligned}\mathrm{Var}\,(a_1\xi_1 + \cdots + a_n\xi_n) &= \sum_{j=1}^{n} a_j^2 \mathrm{Var}\,\xi_j + \sum_{1 \le i,j \le n,\ i \ne j} a_i a_j \mathrm{Cov}\,(\xi_j, \xi_j) \\ &= \sum_{j=1}^{n} a_j^2 \mathrm{Var}\,\xi_j + 2 \sum_{1 \le i < j \le n} a_i a_j \mathrm{Cov}\,(\xi_j, \xi_j)\end{aligned}$$

Továbbá, $\mathrm{Var}\,\xi_j = E(\xi_j - E\xi_j)^2 = E\xi_j^2 - (E\xi_j)^2$, *és* $\mathrm{Cov}\,(\xi_i, \xi_j) = E(\xi_i - E\xi_i)(\xi_j - E\xi_j) = E\xi_i\xi_j - E\xi_i E\xi_j$.

Tétel független valószínűségi változók szorzatának a várható értékéről. *Legyenek $\xi_1, \dots, \xi_n$ olyan független valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyekre $E|\xi_j| < \infty$ minden $1 \le j \le n$ indexre. Ekkor*

$$E(\xi_1 \cdots \xi_n) = \prod_{j=1}^{n} E\xi_j.$$

Hogyan számoljuk ki független valószínűségi változók összegének a sűrűségfüggvényét? Sűrűségfüggvények konvolúciója.

Tétel arról, hogyan számíthatjuk ki független, diszkrét eloszlású valószínűségi változók összegének az eloszlását. *Legyen ξ és η két független diszkrét eloszlású valószínűségi változó, amelyek valamilyen $x_1, x_2, \dots,$ illetve $y_1, y_2, \dots$ értékeket vesznek fel $p_j = P(\xi = x_j)$, és $q_j = P(\eta_j = y_j)$, $j = 1, 2, \dots$ valószínűségekkel. Ekkor $\xi + \eta$ diszkrét eloszlású valószínűségi változó, és*

$$P(\xi + \eta = z) = \sum_{(j,k):\ x_j + y_k = z} P(\xi = x_j) P(\eta = y_k).$$

(Olyan z számokra, amelyekre nincs olyan (j,k) pár, amelyre $x_j+y_k=z$, a fenti képlet jobb-oldalán szereplő üres összeg definició szerint nullával egyenlő, azaz $P(\xi+\eta=z)=0$ az ilyen z számokra.)

Abban a speciális esetben, ha mind ξ mind η egész értékeket vesz fel, azaz mindegyik x_j és y_j egész szám, akkor $\xi+\eta$ is egész értékeket vesz fel, és a fenti képlet azt adja, hogy

$$P(\xi+\eta=n)=\sum_{k=-\infty}^{\infty}P(\xi=k)P(\eta=n-k)\quad \textit{minden } n=0,\pm1,\pm2,\dots \textit{ számra.}$$

Ha mind ξ mind η nem negatív egész értékeket vesz fel, akkor a fenti képlet tovább egyszerűsödik. Ekkor $\xi+\eta$ is nem negatív egész értékeket vesz fel, és

$$P(\xi+\eta=n)=\sum_{k=0}^{n}P(\xi=k)P(\eta=n-k)\quad \textit{minden } n=0,1,2,\dots \textit{ számra.}$$

Hasonló képlet érvényes két sűrűségfüggvénnyel rendelkező független valószínűségi változó összegének a sűrűségfüggvényére.

Tétel arról, hogyan számolhatjuk ki két sűrűségfüggvénnyel rendelkező független valószínűségi változó összegének a sűrűségfüggvényét. *Legyen ξ és η két független valószínűségi változó $f(x)$ és $g(x)$ sűrűségfüggvénnyel. Ekkor a $\xi+\eta$ összegnek is létezik sűrűségfüggvénye. Jelöljük ezt $h(x)$-szel. A $\xi+\eta$ $h(x)$ sűrűségfüggvényét a következő képlet adja meg.*

$$h(x)=\int_{-\infty}^{\infty}f(y)g(x-y)\,dy=\int_{-\infty}^{\infty}f(x-y)g(y)\,dy\quad -\infty<x<\infty.$$

Ha ξ és η csak nem negatív értékeket vesz fel, azaz $f(x)=0$ és $g(x)=0$, ha $x<0$, akkor a fenti képlet egyszerűsödik. Ekkor

$$h(x)=\int_{0}^{x}f(y)g(x-y)\,dy=\int_{0}^{x}f(x-y)g(y)\,dy\quad \textit{ha } 0\le x<\infty.$$

és $h(x)=0$, ha $x<0$.

A következő definiciót szokás bevezetni.

Két integrálható függvény konvolúciójának a definiciója. *Legyen $f(x)$ és $g(x)$ két integrálható függvény a számegyenesen. Ekkor e függvények konvolúciója az*

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

függvény.

Be lehet látni, hogy két integrálható függvény konvolúciója mindig létezik, és az is integrálható függvény. A konvolúció fogalmának bevezetése után a korábban kimondott tételt úgy is megfogalmazhatjuk, hogy két független $f(x)$ és $g(x)$ sűrűségfüggvénnyel rendelkező valószínűségi változó összegének is létezik sűrűségfüggvénye, és az a $h(x) = f * g(x)$ függvény, ahol $*$ konvolúciót jelöl.

Felidézek még egy valószínűségi vizsgálatokban fontos mértékelméleti problémát és az ezzel kapcsolatos legfontosabb eredményeket. Azt a kérdést tárgyalom, hogy egy konvergens függvénysorban mikor cserélhető fel a limeszképzés és az integrálás, azaz mikor mondhatjuk, hogy a függvénysorozat tagjainak az integráljai konvergálnak, és a határérték megegyezik a függvénysorozat limeszfüggvényének (létező) integráljával.

A mértékelmélet néhány fontos eredménye konvergens függvénysorozatok integráljairól.

A valószínűségszámítás eredményeinek bizonyításában gyakran van szükségünk olyan eredményekre, melyek azt mondják ki, hogy konvergens függvények sorozatának az integrálja valamely mérték szerint konvergál a határfüggvény integráljához eme mérték szerint. Az ilyen eredmények vizsgálata a mértékelmélet témakörbe tartozik. Mégis, hasznosnak látszik a legfontosabb eredmények ismertetése. Ezek az úgynevezett Lebesgue (dominated convergence) tétel, a Beppo–Levy tétel és a Fatou lemma. Leírok néhány egyszerű példát is, melyek segíthetnek ezen eredmények tartalmát jobban megérteni.

Lebesgue tétel. *Legyen f_n, $n = 1, 2, \dots,$ mérhető függvények sorozata egy $(X, \mathcal{X}, \mu)$ mértéktéren. Tegyük fel, hogy $\lim\limits_{n\to\infty} f_n(x) = f_0(x)$ a μ mérték szerint majdnem minden $x \in X$ pontban, és létezik olyan "domináns" g függvény az $(X, \mathcal{X}, \mu)$ téren, amelyre $|f_n(x)| \le g(x)$ a μ mérték szerint majdnem minden $x \in X$ pontban minden $n = 1, 2, \dots,$ indexre, és $\int g(x)\,d\mu(x) < \infty$. Ekkor $\lim\limits_{n\to\infty} \int f_n(x)\,d\mu(x) = \int f_0(x)\,d\mu(x)$.*

Beppo-Levy tétel. *Legyen f_n, $n = 1, 2, \dots,$ mérhető függvények sorozata egy $(X, \mathcal{X}, \mu)$ mértéktéren. Tegyük fel, hogy az $f_n(x)$ sorozat monoton növekszik, azaz $f_1(x) \le f_2(x) \le f_3(x) \le \cdots$ majdnem minden $x \in X$ pontban. Legyen $\lim\limits_{n\to\infty} f_n(x) = f_0(x)$ az f_n függvények limesze. Tegyük fel, hogy teljesül az $\int f_1(x)\,d\mu(x) > -\infty$ feltétel. Ekkor $\lim\limits_{n\to\infty} \int f_n(x)\,d\mu(x) = \int f_0(x)\,d\mu(x)$. Abban a speciális esetben, ha $\lim\limits_{n\to\infty} \int f_n(x)\,d\mu(x) = \infty$ a $\int f_0(x)\,d\mu(x) = \infty$ reláció teljesül.*

Fatou lemma. *Legyen f_n, $n = 1, 2, \dots,$ mérhető függvények sorozata egy $(X, \mathcal{X}, \mu)$ mértéktéren. Tegyük fel, hogy $\lim\limits_{n\to\infty} f_n(x) = f_0(x)$ a μ mérték szerint majdnem minden $x \in X$ pontban, és létezik olyan "alsó korlát" g függvény az $(X, \mathcal{X}, \mu)$ téren, amelyre $f_n(x) \ge g(x)$ a μ mérték szerint majdnem minden $x \in X$ pontban minden $n = 1, 2, \dots$ indexre, és $\int g(x)\,d\mu(x) > -\infty$. Ekkor $\liminf\limits_{n\to\infty} \int f_n(x)\,d\mu(x) \ge \int f_0(x)\,d\mu(x)$.*

A fenti tételekben szereplő feltételek jobb megértése érdekében tekintsük a következő példát. Legyen az $(X, \mathcal{X}, \mu)$ mértéktérben az X halmaz a $(0, 1]$ intervallum, az $\mathcal{X}$ σ-algebra a Borel σ-algebra a $[0, 1]$ inervallumon, és a μ mérték a Lebesgue mérték.

Definiáljuk a következő $f_n(x)$, $n=1,2,\dots$, függvényeket ezen a téren a következő módon: $f_n(x)=n$, ha $0<x\le\frac1n$, $f_n(x)=0$, ha $\frac1n\le x\le 1$. Ekkor $\lim\limits_{n\to\infty} f_n(x)=f_0(x)$ minden $x\in(0,1]$ pontban, ahol $f_0(x)\equiv 0$. Az $\int f_n(x)\,d\mu(x)=1$ reláció érvényes minden $n=1,2,\dots$ indexre, és $\int f_0(x)\,d\mu(x)=0$. Ez azt jelenti, hogy ebben az esetben az integrálás és limeszképzés nem felcserélhető.

Ennek megfelelően ez a példa nem teljesíti sem a Lebesgue tétel sem a Beppo–Levy tétel feltételeit, amelyek biztosítanák az integrálás és limeszképzés felcserélhetőségét. A Lebesgue tétel feltételei azért nem teljesülnek, mert az $f_n(x)$, $1\le n<\infty$, függvények legkisebb domináns függvénye az $F(x)=\sup\limits_{n\ge1} f_n(x)$, $0<x\le 1$ függvény nem integrálható. A Beppo–Levy tétel feltételei sem teljesülnek, mert az $f_n(x)$ függvénysorozat nem monoton. A Fatou lemma feltételei teljesülnek, mert mindegyik f_n függvény nem negatív. De a Fatou lemma csak egy viszonylag gyenge állítást mond ki. Azt állítja, hogy amennyiben a konvergens függvénysorozat elemei nem negatívak, vagy teljesítik azt az enyhébb megkötést, hogy van egy integrálható alsó korlátjuk, akkor a határfüggvény integrálja alsó becslése a tekintett integrálok limesz inferiorjának.

Tekintsünk egy másik példát. Legyen $f_n(x)=-\frac1x$, ha $0<x\le\frac1n$, és $f_n(x)=0$, ha $\frac1n<x\le 1$, $n=1,2,\dots$. Ez egy olyan monoton függvénysorozat, amelyre $\lim\limits_{n\to\infty} f_n(x)=f_0(x)$, $f_0(x)\equiv 0$. Ebben a példában $\int f_n(x)\,dx=-\infty$ minden $n=1,2,\dots$ számra, és $\int f_0(x)\,dx=0$, tehát az integrálás és limeszképzés ebben a példában sem cserélhető fel. Ekkor a Beppo–Levy tétel azért nem alkalmazható, mert a benne szereplő feltételek közül az $\int f_1(x)\,d\mu(x)>-\infty$ reláció nem teljesül.

A fenti eredményeket át lehet fogalmazni függvénysorokra is. Így a Lebesgue tétel vagy a Beppo–Levy tétel következményeként az alábbi eredményt fogalmazhatjuk meg.

Tétel konvergens függvénysorok tagonkénti integrálhatóságáról. *Teljesüljön a $G(x)=\sum\limits_{k=1}^{\infty} g_k(x)$ reláció a μ mérték szerint majdnem minden pontban egy $(X,\mathcal{X},\mu)$ mértéktérben. Ha létezik olyan a μ mérték szerint integrálható (azaz az $\int|H(x)|\,d\mu(x)<\infty$ feltételt teljesítő) $H(x)$ függvény, amelyre $\left|\sum\limits_{k=1}^{n} g_k(x)\right|\le H(x)$ minden $n=1,2,\dots$ indexre majdnem minden x pontban, vagy $g_k(x)\ge 0$ minden $k=1,2,\dots$ indexre majdnem minden x pontban, és $\int g_1(x)\,d\mu(x)>-\infty$, akkor $\int G(x)\,d\mu(x)=\sum\limits_{k=1}^{\infty}\int g_k(x)\,d\mu(x)$.*

Ez a tétel egyszerűen következik a Lebesgue tételből vagy a Beppo–Levy tételből, ha azt az $f_n(x)=\sum\limits_{k=1}^{n} g_k(x)$, $n=1,2,\dots$ függvénysorozatra alkalmazzuk.

Megjegyzés: A Lebesgue tétel és a Fatou lemma feltételei gyengíthetőek abban az esetben, ha a μ mérték, amely szerint integrálunk véges. Ekkor az a feltétel, hogy az $f_n(x)$ függvénysorozat majdnem minden x pontban konvergál egy $f_0(x)$ függvényhez helyettesíthető azzal a gyengébb feltétellel, hogy az $f_n(x)$ függvénysorozat mértékben konvergál az $f_0(x)$ függvényhez. A Beppo–Levy tételnek nincs ilyen módosítása, mert az e tétel feltételeiben megfogalmazott monotonitás biztosítja a konvergenciát majdnem mindenütt. (Ha a μ mérték 1-re normált, akkor a mértékelméletben bevezetett mértékben való konvergencia fogalma megegyezik a valószínűségszámításban definiált sztochasztikus konvergencia fogalmával.) Az eredmények bizonyítása nem nehezebb ebben az új, általánosabb esetben. Gyakorlati alkalmazásokban a tétel eredeti formájában megfogalmazott alakja, ahol majdnem mindenütt való konvergenciát használunk ugyanolyan jól használható.