A Valószínűségszámítás I. előadássorozat nyolcadik előadása.
2007. március 27.

A sűrűségfüggvény fogalma

Az előző előadásban kimondott Fontos Tételnek nevezett eredmény, illetve annak általánosításából következik, hogy egy valószínűségi változónak, illetve egy valószínűségi változó függvényének a várható értékét ki lehet számítani csak a valószínűségi változó eloszlásfüggvényének az ismeretében.

Gyakorlati szempontból, annak érdekében, hogy a leggyakrabban előforduló esetekben jobban tudjunk számolni, érdemes bevezetni egy eloszlásfüggvény sűrűségfüggvényének a definicióját. Ez lehetővé teszi, hogy a problémákban felmerülő és sokszor kényelmetlenül kezelhető Lebesgue integrálokat átírjuk, mint (közönséges) Riemann integrálokat.

Sűrűségfüggvény definiciója. *Egy $F(x)$ eloszlásfüggvénynek az $f(u)$, $-\infty < u < \infty$, függvény sűrűségfüggvénye, ha minden $-\infty < x < \infty$ számra teljesül az*

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

azonosság.

Megjegyzés: Többször fogunk beszélni egy valószínűségi változó sűrűségfüggvényéről is. Ez azt jelenti, hogy e valószínűségi változó eloszlásfüggvényének a sűrűségfüggvényét tekintjük. Sok érdekes és fontos eloszlásfüggvénynek létezik sűrűségfüggvénye, de nem mindegyiknek. Például egy diszkrét eloszlású valószínűségi változónak nincs sűrűségfüggvénye.

Tétel eloszlásfüggvények által definiált integrálok kifejezésére sűrűségfüggvények segítségével . *Ha egy $F(x)$ eloszlásfüggvénynek létezik $f(u)$ sűrűségfüggvénye, akkor minden (mérhető) $g(\cdot)$ függvényre teljesül a következő azonosság:*

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

Ez az azonosság úgy értendő, hogy az azonosság két oldalán szereplő kifejezés egyszerre létezik vagy nem létezik, ha mind a két kifejezést mint Lebesgue integrált értelmezzük. Ha a jobboldalon szereplő integrál létezik mint (közönséges) Riemann integrál, akkor ez az integrál tekinthető úgy is mint (egy a Riemann integrál értékével megegyező) Lebesgue integrál. Tehát az azonosság ebben a fontos speciális esetben is érvényes.

Annak érdekében, hogy megértsük, hogyan tudjuk egy eloszlásfüggvény sűrűségfüggvényét kiszámítani, idézzük fel a klasszikus analízis egyik alapvető eredményét, a Newton–Leibniz formulát.

Kissé(?) pontatlanul a Newton–Leibniz formulát úgy fogalmazhatjuk meg, hogy a differenciálás és integrálás egymás inverzei, azaz egy függvény differenciálhányadosának az integrálja, illetve a függvény integráljának a differenciálhányadosa egyenlő a kiinduló függvénnyel. Pontosabb megfogalmazás esetén a következő tényeket figyelembe kell venni.

a) Egy függvény differenciálhányadosa nem változik, ha hozzáadunk egy konstanst. Ennek a ténynek az felel meg az integrálásnál, hogy egy függvény $\int_a^x f(u)\,du$ (az x változó szerinti) határozatlan integrálja függ az integrálás a alsó határától, ha az a konstanst megváltoztatjuk, akkor a határozatlan integrál értékéhez egy konstans adódik hozzá.
b) Nem minden függvény differenciálható. E tény integrálásra vonatkozó megfelelője az, hogy egy határozatlan integrál folytonos függvény, sőt ennél erősebb simasági tulajdonságot is teljesít. (Ennek az erősebb simasági tulajdonságnak a pontos megfogalmazása a mértékelmélet tárgya, ezzel a kérdéssel itt nem foglalkozunk.) Ezenkívül vegyük észre azt is, hogy egy függvény határozatlan integrálja nem változik, ha a függvény értékét véges sok pontban megváltoztatjuk.

Megfogalmazom a Newton–Leibniz formulát részletesebben. A modern matematikában (a mértékelméletben) ennek az eredménynek pontosabb, általánosabb változatát bizonyítják be.

Newton–Leibniz formula. *Legyen $F(x)$ folytonos függvény egy $[a,b]$ véges intervallumon, amely véges sok pont kivételével folytonosan deriválható. Legyen $f(u)=\left.\frac{dF(x)}{dx}\right|_{x=u}$ minden pontban, ahol az $F(\cdot)$ differenciálható. Ekkor $F(x)-F(a)=\int_a^x f(u)\,du$ minden $a\le x\le b$ számra. Továbbá, ha a fenti feltétel teljesül minden véges $[a,b]$ intervallumban, és létezik a $F(-\infty)=\lim\limits_{x\to-\infty}F(x)$ határérték, akkor $F(x)-F(-\infty)=\int_{-\infty}^x f(u)\,du$ minden $-\infty<x<\infty$ számra.*

Megfordítva, ha $f(u)$, (pontosabban annak abszolut értéke) integrálható függvény egy $[a,b]$ intervallumban, és $F(x)=\int_a^x f(u)\,du$, $a\le x\le b$, akkor $f(u)=\left.\frac{dF(x)}{dx}\right|_{x=u}$ (majdnem) minden $a\le u\le b$ pontban. Ha az $f(u)$ függvény integrálható az egész számegyenesen, akkor a fenti állítás igaz $a=-\infty$ választással is.

1. megjegyzés. A Newton–Leibniz formula jelentősége számunkra az, hogy eszerint az eredmény szerint a sűrűségfüggvény (feltéve, hogy az létezik) egyenlő az eloszlásfüggvény deriváltjával. Megfordítva, a sűrűségfüggvény ismeretében meg tudjuk határozni az eloszlásfüggvényt. Annak értéke az x pontban megegyezik a sűrűségfüggvény integráljával a $[-\infty,x]$ intervallumban.

2. megjegyzés. A mértékelméletben bebizonyították a Newton–Leibniz formula élesebb formáját is. Pontosan leírták azon függvények osztályát, az úgynevezett abszolut folytonos függvényeket, amelyekre igaz, hogy a függvény egyenlő a deriváltjának a (Lebesgue) integráljával. Az általunk kimondott eredmény csak egy elégséges feltételt ad arra, hogy egy $F(\cdot)$ függvény teljesítse ezt a tulajdonságot. Viszont ez az eredmény alkalmazható a gyakorlatban előforduló feladatok majdnem mindegyikére.

A kimondott tételben megengedtük, hogy az $F(\cdot)$ függvény néhány pontban ne legyen differenciálható. Ezt azért tettük, hogy ne zárjunk ki néhány érdekes esetet.

Ilyen eset például a következő $F(\cdot)$ eloszlásfüggvény: $F(x) = 0$, ha $x \le 0$, $F(x) = x$, ha $0 \le x \le 1$, és $F(x) = 1$, ha $x \ge 1$. Ennek az eloszlásfüggvénynek a sűrűségfüggvénye az az $f(\cdot)$ függvény, amelyre $f(u) = 1$, ha $0 \le u \le 1$, és $f(u) = 0$ különben. Ebben az esetben az $F(\cdot)$ függvény folytonosan deriválható mindenütt, kivéve az $x = 0$ és $x = 1$ pontot.

3. megjegyzés: A Newton–Leibniz formula második részében megfogalmazott állítás szerint egy $f(\cdot)$ függvény integráljának a deriváltja csak *majdnem* mindenütt egyezik meg az eredeti $f(\cdot)$ függvénnyel. Valóban, kissé óvatosabban kell megfogalmazni az állításokat, mert ha például egy függvényt véges sok pontban megváltoztatunk, akkor annak integrálja megegyezik az eredeti függvény integráljával. Ez azt jelenti, hogy nem várhatjuk azt, hogy egy függvény integráljának a deriváltja minden pontban megegyezzen az eredeti függvénnyel. Viszont ez az állítás igaz *majdnem minden* pontban. Azt, hogy mit jelent a „majdnem minden” kitétel pontosan elmagyarázzák a mértékelméletben, de ez nem témája a mostani előadásnak.

Az eloszlásfüggvényekhez hasonlóan a sűrűségfüggvényeket is lehet pontosan jellemezni. Erről szól az alábbi tétel.

Tétel sűrűségfüggvények jellemzéséről. *Egy $f(\cdot)$ (mérhető) függvény akkor és csak akkor sűrűségfüggvénye egy alkalmas eloszlásfüggvénynek, ha $f(u) \ge 0$ majdnem minden $-\infty < u < \infty$ számra, és $\int_{-\infty}^{\infty} f(u)\,du = 1$.*

Ezt a tételt nem nehéz bebizonyítani néhány alapvető mértékelméleti eredmény segítségével, de most elhagyom a bizonyítást. Külön tételben megfogalmazom azt az eredményt, amely megadja, hogyan lehet kiszámolni egy valószínűségi változó várható értékét a valószínűségi változó eloszlás vagy sűréségfüggvényének az ismeretében. Ez közvetlen következménye néhány korábban megfogalmazott eredménynek.

Tétel várható érték kiszámolásáról a sűrűségfüggvény segítségével. *Jelölje $F(\cdot)$ illetve $f(\cdot)$ egy ξ valószínűségi változó eloszlás illetve sűrűségfüggvényét. Legyen $g(\cdot)$ mérhető függvény a számegyenesen. Ekkor*

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

és

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

Továbbá, ha ismerjük egy valószínűségi változó $F(x)$ eloszlásfüggvényét, akkor könnyen kiszámolhatjuk annak sűrűségfüggvényét is. Ezt az eredményt, amely a Newton–Leibniz formula egyszerű következménye külön lemmában mondom ki.

Lemma sűrűségfüggvény kiszámításáról. *Ha egy valószínűségi változónak van sűrűségfüggvénye, és a valószínűségi változó eloszlásfüggvénye $F(x)$, akkor a valószínűségi változó sűrűségfüggvénye az $f(x) = \frac{dF(x)}{dx}$, $-\infty < x < \infty$, függvény.*

Az eloszlásfüggvény jellemzéséhez hasonlóan jellemezni lehet azokat a függvényeket, amelyek tekinthetőek, mint egy alkalmas eloszlásfüggvény sűrűségfüggvénye. Ez az eredmény egyszerű következménye az eloszlásfügvények jellemzéséről és az eloszlás és sűrűségfüggvények kapcsolatáról szóló eredménynek. Ezért az állítás bizonyítását elhagyom.

Tétel sűrűségfüggvények jellemzéséről. *Egy a számegyenesen definiált $f(x)$ függvényhez akkor és csak akkor létezik valamely $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált valószínűségi változó, amelynek ez a sűrűségfüggvénye, ha ez a függvény teljesíti a következő két feltételt.*

a) $f(x) \geq 0$ *a számegyenes (majdnem) minden x pontjában.*

b)

$$\int_{-\infty}^{\infty} f(x)\, dx = 1.$$

Feladat:

1.) Legyen ξ valószínűségi változó $f(u)$ sűrűségfüggvénnyel, a és b valós számok. Határozzuk meg az $a\xi + b$ és ξ^2 valószínűségi változók sűrűségfüggvényét.

Megoldás: Legyen $F(x) = P(\xi < x) = \int_{-\infty}^{x} f(u)\, du$ a ξ valószínűségi változó eloszlásfüggvénye. Ekkor az $\eta = a\xi + b$ valószínűségi változó eloszlásfüggvénye a $G(x) = P(\eta < x) = P\left(\xi < \frac{x-b}{a}\right) = F\left(\frac{x-b}{a}\right)$ függvény, ha $a > 0$, és $G(x) = P(\eta < x) = P\left(\xi > \frac{x-b}{a}\right) = 1 - F\left(\frac{x-b}{a}\right)$, ha $a < 0$. Innen η sűrűségfüggvénye $g(x) = \frac{dG(x)}{dx} = \frac{1}{|a|} f\left(\frac{x-b}{a}\right)$.

A $\zeta = \xi^2$ valószínűségi változó eloszlásfüggvénye

$$G(x) = P(\xi^2 < x) = P\left(-\sqrt{x} < \xi < \sqrt{x}\right) = F\left(\sqrt{x}\right) - F\left(-\sqrt{x}\right),$$

ha $x \geq 0$, és $G(x) = 0$, ha $x < 0$. Innen deriválással ζ sűrűségfüggvénye $g(x) = \frac{1}{2\sqrt{x}} \left(f\left(\sqrt{x}\right) + f\left(-\sqrt{x}\right)\right)$, ha $x > 0$, $g(x) = 0$, ha $x \leq 0$.

Néhány fontos folytonos eloszlás.

a.) *Normális eloszlásfüggvény.*

Bevezetem a standard normális eloszlásfüggvény definicióját. Későbbi eredményekből fog kiderülni, hogy ez az eloszlás miért játszik fontos szerepet a valószínűségszámításban.

A standard normális eloszlás definiciója. *A $\Phi(x)$ standard normális eloszlás az az eloszlás, amelynek sűrűségfüggvénye a $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$, $-\infty < u < \infty$ függvény.*

E definició helyességének igazolásához meg kell mutatni, hogy a fent definiált $\varphi(\cdot)$ függvény valóban sűrűségfüggvény. Ennek érdekében be kell bizonyítani a következő eredményt.

Tétel.

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = 1.$$

Bizonyítás. Vezessük be az $I = \displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$ jelölést. Ekkor

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-v^2/2}\, dv = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-(u^2+v^2)/2}\, du\, dv \\ &= \int_0^{2\pi} \int_0^{\infty} \frac{1}{2\pi} r e^{-r^2/2}\, dr\, d\varphi = \int_0^{\infty} r e^{-r^2/2}\, dr = \left[-e^{-r^2/2}\right]_0^{\infty} = 1. \end{aligned}$$

Ebben a számolásban az I^2 mennyiséget kifejező az (u,v) térben megadott kettős integrált kifejeztük mint az (r,φ) polárkoordinátarendszerben, $u = r\cos\varphi$, $v = r\sin\varphi$, $0 \le r < \infty$, $0 \le \varphi < 2\pi$, kifejezett integrált. Ezen számolás során felhasználjuk, hogy a polárkkoordináta rendszerbe való áttérést leíró transzformáció Jacobianja r, azaz formálisan $du\, dv = r\, dr\, d\varphi$. Ezért jelenik meg egy r szorzó az integrálban a polárkoordinátarendszerbe való áttéréskor. Ezt az integráltranszformációkról szóló eredményt, illetve annak szemléletes tartalmát a kiegészítésben magyarázom el.

Megjegyzés: Természetesnek látszódna a fent kimondott tételt úgy bebizonyítani, hogy kiszámítjuk a $\varphi(x)$ függvény $\Phi(x) = \int_{-\infty}^{x} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\, du$ primitív függvényét, és ebbe a képletbe $x = \infty$ értéket helyettesítünk. Erre azonban nem vagyunk képesek, mert a fenti primitív függvényt nem tudjuk kiszámítani. Be lehet látni, hogy ez a primitív függvény nem írható fel szokásos függvények és műveletek segítségével zárt alakban. Ez azonban mély matematikai eredmények felhasználását igényli. (A Galois elméletet kell használni.)

Megmutatom, hogy egy standard normális eloszlású ξ valószínűségi változó várható értéke nulla, szórásnégyzete 1. Ez az oka a standard jelzőnek a standard normális eloszlás definiciójában.

Valóban, $E\xi = 0$, mert $E\xi = \int_{-\infty}^{\infty} u \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$, és ez az integrál nulla, mert az integrandus páratlan függvény. Ezután parciális integrálással kapjuk $f(x) = x$ és $g(x) = \frac{1}{\sqrt{2\pi}} x e^{-x^2/2} = \frac{d}{dx}\left(-\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right)$ választással, hogy

$$E\xi^2 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-x^2/2}\, dx = \left[-\frac{x}{\sqrt{2\pi}} e^{-x^2/2}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = 1.$$

Innen $\mathrm{Var}\,\xi = E\xi^2 - (E\xi)^2 = 1$.

Feladatok:

2.) Ha ξ standard normális eloszlású valószínűségi változó, m, σ valós számok, akkor az $\sigma\xi + m$ valószínűségi változó várható értéke m szórásnégyzete σ^2, sűrűségfüggvénye pedig $g(u) = \frac{1}{\sqrt{2\pi}|\sigma|} e^{-(u-m)^2/2\sigma^2}$.

Megoldás: $E(\sigma\xi + m) = \sigma E\xi + m = m$, és $\mathrm{Var}\,(\sigma\xi + m) = \sigma^2 \mathrm{Var}\,\xi = \sigma^2$. Továbbá, az első feladat eredménye alapján a $\sigma\xi + m$ valószínűségi változó sűrűségfüggvénye az $f(x) = \frac{1}{\sigma}\varphi\left(\frac{x-m}{\sigma}\right)$, ahol $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, a standard normális sűrűségfüggvény. Innen következik a feladat állítása.

A fenti feladat eredménye alapján egy valószínűségi változót akkor nevezünk normális eloszlásúnak, ha van sűrűségfüggvénye, és az $g(u) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(u-m)^2/2\sigma^2}$ alakban adható meg alkalmas m és $\sigma > 0$ számokkal. A képletben szereplő m és σ^2 számok megegyeznek a normális eloszlású valószínűségi változó várható érékével és szórásnégyzetével. Ez speciálisan azt is jelenti, hogy egy normális eloszlást két paraméter meghatározza annak várható éréke és szórásnégyzete.

3.) Ha ξ standard normális eloszlású valószínűségi változó, akkor$E\xi^{2k-1} = 0$ $E\xi^{2k} = 1 \cdot 3 \cdots (2k-1)$ minden $k = 1, 2, \dots$ számra.

Megoldás:

$$E\xi^{2k-1} = \int_{-\infty}^{\infty} x^{2k-1} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = 0,$$

mert az integrandus páratlan függvény. Másrészt parciális integrálással $f(x) = x^{2k-1}$ és $g(x) = x\frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{d}{dx}\left(-\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right)$ kapjuk, hogy

$$\begin{aligned} E\xi^{2k} &= \int_{-\infty}^{\infty} x^{2k} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \int_{-\infty}^{\infty} (2k-1) x^{2k-2} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \\ &= (2k-1) E\xi^{2k-2}. \end{aligned}$$

Innen k szerinti indukcióval kapjuk a feladat második állítását.

4.) Legyen ξ standard normális eloszlású valószínűségi változó, azaz legyen ξ sűrűségfüggvénye $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, és legyen t valós szám. Számoljuk ki az $e^{t\xi}$ valószínűségi változó $Ee^{t\xi}$ várható értékét.

Megoldás:

$$\begin{aligned} Ee^{t\xi} &= \int_{-\infty}^{\infty} e^{tu} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tu-u^2/2-t^2/2+t^2/2}\, du \\ &= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(t-u)^2/2}\, du = e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = e^{t^2/2}. \end{aligned}$$

5.) Legyen ξ normális eloszlású valószínűségi változó $m = 2$ várható értékkel és $\sigma^2 = 3$ szórásnégyzettel. Számoljuk ki a ξ valószínűségi változó $E\xi^4$ negyedik momentumát.

Megoldás: Tudjuk, hogy a ξ valószínűségi változó sűrűségfüggvénye

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} = \frac{1}{\sqrt{6\pi}} e^{-(x-2)^2/6},$$

ahonnan $E\xi^4 = \int_{-\infty}^{\infty} x^4 \frac{1}{\sqrt{6\pi}} e^{-(x-2)^2/6}\, dx$. Ezt az integrált ki tudjuk számolni hasonlóan a standard normális eloszlás momentumaihoz, sőt a számolást egyszerűen visszavezethetjük az ilyen momentumok kiszámolására, az $u = \frac{x-2}{\sqrt{3}}$ helyettesítés segítségével. Valójában egyszerűbben is megoldhatjuk a feladatot a következő módon. Vezessünk be egy η standard normális eloszlású valószínűségi változót, és legyen $\bar{\xi} = \sqrt{3}\eta + 2$. Ekkor $\bar{\xi}$ és ξ eloszlása megegyezik, mivel mindkettő normális eloszlású $m = 2$ várható értékkel és $\sigma^2 = 3$ szórásnégyzettel. Innen $E\xi^4 = E\bar{\xi}^4 = E(\sqrt{3}\eta + 2)^4 = 9E\eta^4 + 24\sqrt{3}E\eta^3 + 72E\eta^2 + 32\sqrt{3}E\eta + 16 = 9E\eta^4 + 72E\eta^2 + 16 = 27 + 72 + 16 = 105$ a második feladat eredménye alapján.

b.) *Egyenletes eloszlásfüggvény.*

Egyenletes eloszlásfüggvény definiciója. *Egy ξ valószínűségi változó egyenletes eloszlású egy $[a, b]$ intervallumban, $-\infty < a < b < \infty$, ha sűrűségfüggvénye $f(u) = \frac{1}{b-a}$, ha $a \le u \le b$, és $f(u) = 0$ egyébként.*

Kiszámítjuk egy az $[a, b]$ intervallumban egyenletes eloszlású ξ valószínűségi változó várható értékét és szórásnégyzetét.

$$E\xi = \int_a^b u \frac{1}{b-a}\, du = \frac{1}{b-a} \left[\frac{u^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2},$$

$$\begin{aligned} \operatorname{Var} \xi &= E\left(\xi - \frac{b+a}{2}\right)^2 = \int_a^b \left(u - \frac{b+a}{2}\right)^2 \frac{1}{b-a}\, du = \frac{1}{b-a} \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} u^2\, du \\ &= \frac{1}{b-a} \left[\frac{u^3}{3}\right]_{-\frac{b-a}{2}}^{\frac{b-a}{2}} = \frac{2}{3(b-a)} \left(\frac{b-a}{2}\right)^3 = \frac{(b-a)^2}{12}. \end{aligned}$$

Feladatok:

6.) Legyen ξ egyenletes eloszlású valószínűségi változó a $[0,1]$ intervallumban, azaz legyen $f(\cdot)$ sűrűségfüggvénye $f(x) = 1$, ha $0 \le x \le 1$, $f(x) = 0$ egyébként. Számítsuk ki a ξ^2 valószínűségi változó várható értékét és szórásnégyzetét.
Első megoldás:

$$\operatorname{Var} \xi^2 = E\left(\left(\xi^2\right)^2\right) - \left(E\xi^2\right)^2 = E\xi^4 - \left(E\xi^2\right)^2 = \int x^4 f(x)\,dx - \left(\int x^2 f(x)\,dx\right)^2,$$

ahol $f(x) = 1$, ha $0 \le x \le 1$, $f(x) = 0$ egyébként, azaz $f(x)$ a ξ valószínűségi változó sűrűségfüggvénye. Innen $E\xi^2 = \int_0^1 x^2\,dx = \frac{1}{3}$, és $\operatorname{Var}\xi^2 = \int_0^1 x^4\,dx - \left(\int_0^1 x^2\,dx\right)^2 = \frac{1}{5} - \left(\frac{1}{3}\right)^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$.
Második megoldás: Számítsuk ki először a ξ^2 valószínűségi változó eloszlás és sűrűségfüggvényét. A ξ^2 valószínűségi változó $F(\cdot)$ eloszlásfüggvénye

$$F(x) = P(\xi^2 < x) = P(-\sqrt{x} < \xi < \sqrt{x}) = P(0 \le \xi \le \sqrt{x}) = \sqrt{x}, \quad \text{ha } 0 \le x \le 1,$$

$F(x) = 0$, ha $\xi < 0$, $F(x) = 1$, ha $x > 1$. Innen a ξ^2 valószínűségi változó $f(\cdot)$ sűrűségfüggvénye $f(x) = \frac{1}{2\sqrt{x}}$, ha $0 \le x \le 1$, és $f(x) = 0$ egyébként. Ezért

$$E\xi^2 = \int_0^1 x \frac{1}{2\sqrt{x}}\,dx = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3},$$

és

$$\operatorname{Var}\xi^2 = \int_0^1 x^2 \frac{1}{2\sqrt{x}}\,dx - \left(\int_0^1 x \frac{1}{2\sqrt{x}}\,dx\right)^2 = \frac{1}{2} \cdot \frac{2}{5} - \left(\frac{1}{2} \cdot \frac{2}{3}\right)^2 = \frac{4}{45}.$$

c.) *Exponenciális eloszlásfüggvény.*

Exponenciális eloszlásfüggvény definiciója. *Egy ξ valószínűségi változó exponenciális eloszlású λ paraméterrel, $\lambda > 0$, ha eloszlásfüggvénye, $F(x) = P(\xi < x) = 1 - e^{-\lambda x}$, ha $x \ge 0$, és $F(x) = P(\xi < x) = 0$, ha $x < 0$. Ezzel ekvivalens jellemzés: Egy valószínűségi változó exponenciális eloszlású $\lambda > 0$ paraméterrel, ha létezik $f(u)$ sűrűségfüggvénye, és az $f(u) = \lambda e^{-\lambda u}$, ha $u \ge 0$, $f(u) = 0$, ha $u \le 0$ alakú.*

Számoljuk ki egy λ paraméterű ξ exponenciális eloszlású valószínűségi változó várható értékét és szórásnégyzetét.

Parciális integrálással kapjuk, hogy

$$E\xi = \int_0^\infty u\lambda e^{-\lambda u}\,du = \frac{1}{\lambda}\int_0^\infty u e^{-u}\,du = \frac{1}{\lambda}\left(\left[-ue^{-u}\right]_0^\infty + \int_0^\infty e^{-u}\,du\right) = \frac{1}{\lambda},$$

$$E\xi^2 = \int_0^\infty u^2 \lambda e^{-\lambda u}\, du = \frac{1}{\lambda^2}\left(\left[-u^2 e^{-u}\right]_0^\infty + \int_0^\infty 2u e^{-u}\, du \right) = \frac{2}{\lambda^2}.$$

Ezért $\operatorname{Var}\xi = E\xi^2 - (E\xi)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$.

Feladatok:

7.) Számítsuk ki egy λ paraméterű exponenciális eloszlású valószínűségi változó momentumait.
Megoldás:

$$E\xi^k = \int_0^\infty u^k \lambda e^{-\lambda u}\, du = \frac{1}{\lambda^k}\left(\left[-u^k e^{-u}\right]_0^\infty + \int_0^\infty k u^{k-1} e^{-u}\, du \right) = \frac{k}{\lambda^k} E\xi_1^{k-1},$$

ahol ξ_1 egy exponenciális eloszlású valószínűségi változó $\lambda = 1$ paraméterrel. Innen k szerinti indukcióval $E\xi_1^k = k!$, $E\xi^k = \frac{k!}{\lambda^k}$.

8.) Mutassuk meg, hogy egy exponenciális eloszlású ξ valószínűségi változó teljesíti a következő úgynevezett örökifjú tulajdonságot:

$$P(\xi > x + y | \xi > y) = P(\xi > x)$$

minden $x \geq 0$ és $y \geq 0$ számra.
Megoldás: Mivel $P(\xi > u) = e^{-\lambda u}$ minden $u \geq 0$ számra, ezért $P(\xi > x + y | \xi > y) = \frac{e^{-\lambda(x+y)}}{e^{-\lambda y}} = e^{-\lambda x} = P(\xi > x)$.

Nem kötelező házi feladat:

Ha egy ξ valószínűségi változó teljesűti az örökifjú tulajdonságot, akkor az exponenciális eloszlású.

Ezenkívül megbeszéltük az örökifjú tulajdonság szemléletes tartalmát. Tekintsünk egy személyt vagy tárgyat amelynek élettartama véletlen ξ valószínűségi változó. Tekintsük a $\xi{+}y$ valószínűségi változó feltételes eloszlását azon feltétel mellett, hogy $\xi > y$, azaz azt milyen valószínűséggel fog az a személy vagy tárgy legalább még x ideig élni, feltéve, hogy megélte az y időpontot. Ha ez a feltételes eloszlás nem függ az y értéktől, akkor azt mondjuk, hogy a ξ valószínűségi változó teljesíti az örökifjú tulajdonságot. A fenti feladatban ezt a tulajdonságot fogalmaztuk meg formálisan.

Megjegyzés: A 6. előadás nyolcadik feladatában láttuk, hogy a geometriai eloszlás teljesíti az örökifjú tulajdonság diszkrét változatát, azaz a $P(\xi > x + y | \xi > y) = P(\xi > y)$ azonosságot akkor, ha x és y csak nem-negatív egész értékeket vehetnek fel.

9.) Mutassuk meg, hogy létezik olyan ξ valószínűségi változó teljesíti a következő szuper örökifjú tulajdonságot:

$$P(\xi > x + y | \xi > y) > P(\xi > x)$$

minden $x > 0$ és $y > 0$ számra.

Megoldás: Legyen a ξ valószínűségi változó olyan, hogy $P(\xi > x) = e^{-\sqrt{x}}$, ha $x \ge 0$ és $P(\xi > x) = 1$, ha $x < 0$. Ekkor $P(\xi > x + y | x > y) = e^{-\sqrt{x+y}+\sqrt{y}}$, ha $x > 0$, $y > 0$. Ezért elég megmutatni, hogy $e^{-\sqrt{x+y}+\sqrt{y}} > e^{-\sqrt{y}}$, illetve hogy $\sqrt{x+y} < \sqrt{x} + \sqrt{y}$, ha $x > 0$ és $y > 0$. Ez az egyenlőtlenség viszont (négyzetre emelés után) könnyen látható.

10.) Ledobunk a $[0,2]$ intervallumra 100 pontot egymástól függetlenül a $[0,2]$ intervallumba egyenletes eloszlással. A $[0,1]$ intervallumba eső pontok értékeit beírjuk egy jegyzőkönyvbe. Számítsuk ki a jegyzőkönyvbe írt számok összegének a várható értékét és szórásnégyzetét.

Megoldás: Először megmutatom, hogyan lehet ezt a feladatot úgy megoldani, hogy a benne szereplő várható értékeket Riemann–Stieltjes integrálok segítségével számoljuk ki, majd megmutatom, hogyan lehet ezeket a várható értékeket hagyományos módon, Riemann integrálok segítségével kiszámolni.

Vezessük be a következő ξ_j, $1 \le j \le 100$, valószínűségi változókat. $\xi_j = x$, ha a j-ik ledobott pont értéke $x \in [0,1]$, és nulla, ha a j-ik pont az $[1,2]$ intervallumba esik. Ekkor minket az $S = \sum\limits_{j=1}^{100} \xi_j$ valószínűségi változó várható értéke és szórásnégyzete érdekel. A ξ_j valószínűségi változó $F(x)$ eloszlásfüggvénye sem nem diszkrét, sem nem folytonos, hanem $F(x) = F_1(x) + F_2(x)$ alakban írható fel, ahol $F_1(x)$-nek $f(x) = \frac{1}{2}$, ha $0 \le x \le 1$ sűrűségfüggvénye van, azaz $F_1(x) = \int_{-\infty}^{x} f(u)\,du$ ezzel az $f(\cdot)$ sűrűségfüggvénnyel, $F_2(x)$ pedig egy olyan (nem valószínűségi) μ mértéknek az eloszlásfüggvénye, amelyre $\mu(\{0\}) = \frac{1}{2}$, és $\mu(R \setminus \{0\}) = 0$. Ezért $E\xi_j = \int xF(\,dx) = \int xf(x)\,dx + 0 \cdot \mu(\{0\}) = \int_0^1 \frac{1}{2}x\,dx = \frac{1}{4}$, $E\xi_j^2 = \int x^2 F(\,dx) = \int x^2 f(x)\,dx + 0^2 \cdot \mu(\{0\}) = \int_0^1 \frac{1}{2}x^2\,dx = \frac{1}{6}$, $\operatorname{Var}\xi_j = \frac{1}{6} - \frac{1}{16} = \frac{5}{48}$ minden $1 \le j \le 100$ indexre. Innen $ES = 25$, $\operatorname{Var} S = \frac{125}{12}$.

Az $E\xi_j$ és $E\xi_j^2$ mennyiségeket egyszerűbb meggondolások segítségével is kiszámolhatjuk. Legyen η_j a j-ik ledobott pont értéke, és vezessük be a következő $h(u)$ függvényt a $[0,1]$ intervallumon: $h(u) = u$, ha $0 \le u \le 1$, $h(u) = 0$, ha $1 \le u \le 2$. Ekkor $\xi_j = h(\eta_j)$, és η_j egyenletes eloszlású a $[0,2]$ intervallumon. Innen $E\xi_j = Eh(\eta_j) = \int_0^2 \frac{1}{2}h(u)\,du = \int_0^1 \frac{u}{2}\,du = \frac{1}{4}$, $E\xi_j^2 = Eh^2(\eta_j) = \int_0^2 \frac{1}{2}h^2(u)\,du = \int_0^1 \frac{u^2}{2}\,du = \frac{1}{6}$, $\operatorname{Var}\xi_j = E\xi_j^2 - (E\xi_j)^2 = \frac{1}{6} - \frac{1}{16} = \frac{5}{48}$.

Többváltozós eloszlásfüggvények, és valószínűségi változók függetlensége.

Megtárgyaljuk az eloszlásfüggvény többdimenziós változatát. E fogalom bevezetése után lehet természetes módon definiálni valószínűségi változók függetlenségét. Az itt szereplő fogalmak és bizonyítások hasonlóak a korábban tárgyalt egydimenziós esethez. Ezen eredmények bizonyítását csak nagyon vázlatosan fogom ismertetni. Vázolni fogom azokat a mértékelméleti eredményeket, amelyeket a bizonyítások felhasználnak. Elsősorban a fogalmak és eredmények tisztességes leírására fogok törekedni.

Többdimenziós eloszlásfüggvény definiciója. *Legyen adva k valós értékű $\xi_1, \dots, \xi_k$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ezek (együttes) eloszlásfüggvénye az*

$$F(x_1, \dots, x_k) = P(\xi_1 < x_1, \dots, \xi_k < x_k),$$

k változós függvény, ahol $-\infty < x_j < \infty$ minden $1 \le j \le k$ indexre.

Az előző előadásban megfogalmaztam az egydimenziós eloszlásfüggvények jellemzését egy tételben és az azt megelőző lemmában. Érvényes ezeknek az eredményeknek a természetes többdimenziós változata is. A többváltozós eloszlásfüggvényeket jellemző eredmény megfogalmazásához a következő kérdést kell tisztázni. Az előbb említett lemmában felsoroltuk az egyváltozós függvényeknek a tulajdonságait. Mik e tulajdonságok a többváltozós megfelelői? Többek között meg kell találnunk a lemmában szereplő az eloszlásfüggvény monotonítását kimondó a) feltételnek a többdimenziós megfelelőjét.

Ezt megértendő tekintsünk egy tetszőleges $\mathbf{K} = [a_1, b_1) \times \cdots \times [a_k, b_k)$ alakú téglatestet, $-\infty < a_j < b_j < \infty$, $j = 1, \dots, k$, a k dimenziós térben. Világos, hogy tetszőleges $(\xi_1, \dots, \xi_k)$ k-dimenziós valószínűségi változóra teljesül a

$$P\left((\xi_1(\omega), \dots, \xi_k(\omega)) \in \mathbf{K}\right) = P\left(\bigcap_{j=1}^{k} \{\omega\colon\ a_j \le \xi_j(\omega) < b_j\}\right) \ge 0$$

tulajdonság. Azt állítom, hogy az ilyen alakú események valószínűsége viszonylag egyszerűen kifejezhető a $(\xi_1, \dots, \xi_k)$ véletlen vektor $F(x_1, \dots, x_k) = P(\xi_1 < x_1, \dots, \xi_k < x_k)$ eloszlásfüggvényének a segítségével, és az így kapott kifejezés nem-negatív volta az előbb említett a) feltétel megfelelője a többdimenziós esetben.

Annak érdekében, hogy ezt a feltételt megfogalmazzuk, a következő állítást kell igazolni. Tekintsük egy $\mathbf{K} = [a_1, b_1) \times \cdots \times [a_k, b_k)$ alakú téglatestet, és ennek csúcspontjait, azaz az olyan $(u_1, \dots, u_k)$ pontokat, amelyek koordinátái az a_j vagy b_j számok. Annak valószínűsége, hogy egy $F(x_1, \dots, x_k)$ eloszlású $(\xi_1, \dots, \xi_k)$ véletlen vektor a $\mathbf{K} = [a_1, b_1) \times \cdots \times [a_k, b_k)$ téglatestbe esik kifejezhető mint az $F(\cdot)$ eloszlásfüggvénynek a téglatest csúcspontjaiban felvett értékeinek alkalmas lineáris kombinációja. Pontosabban,

$$P((\xi_1, \dots, \xi_k) \in \mathbf{K}) = \sum_{\substack{u_j = a_j \text{ vagy } b_j \\ j=1,\dots,k}} (-1)^{\chi(u_1, \dots, u_k)} F(u_1, \dots, u_k),$$

ahol $\chi(u_1,\dots,u_k)$ jelöli az a_j-k számát az $u_1,\dots,u_k$ sorozatban. Az $F(\cdot)$ függvény monotonitásának az többdimenziós esetben az felel meg, hogy a fent felírt azonosság jobboldalán szereplő kifejezés nem-negatív. Az egyszerűség kedvéért ezt az azonosságot csak $k=2$ esetben ellenőrizzük.

Ekkor azt kell megmutatni, hogy $P(a_1\le\xi_1<b_1,a_2\le\xi_2<b_2)=F(b_1,b_2)-F(b_1,a_2)-F(a_1,b_2)+F(a_1,a_2)$. Viszont $F(b_1,b_2)-F(b_1,a_2)=P(\xi_1<b_1,a_2\le\xi_2<b_2)$, és $F(a_1,b_2)-F(a_1,a_2)=P(\xi_1<a_1,a_2\le\xi_2<b_2)$. Ezt a két azonosságot kivonva egymásból kapjuk a kívánt állítást.

Feladat:

11.) Legyen egy (ξ_1,ξ_2) véletlen vektor eloszlása az $F(x_1,x_2)=P(\xi_1<x_1,\xi_2<x_2)$ függvény, $-\infty<a_1<b_1<\infty$, $-\infty<a_2<b_2<\infty$ valós számok. Fejezzük ki az $F(x_1,x_2)$ eloszlásfüggvény segítségével a $P(a_1\le\xi_1<b_1,a_2\le\xi_2<b_2)$, $P(a_1\le\xi_1\le b_1,a_2\le\xi_2\le b_2)$ és $P(a_1<\xi_1<b_1,a_2<\xi_2<b_2)$ valószínűségeket.

Megoldás: Először azt megmutatjuk meg, hogy $P(a_1\le\xi_1<b_1,a_2\le\xi_2<b_2)=F(b_1,b_2)-F(b_1,a_2)-F(a_1,b_2)+F(a_1,a_2)$. Valóban, $F(b_1,b_2)-F(b_1,a_2)=P(\xi_1<b_1,a_2\le\xi_2<b_2)$, és $F(a_1,b_2)-F(a_1,a_2)=P(\xi_1<a_1,a_2\le\xi_2<b_2)$. Ezt a két azonosságot kivonva egymásból megkapjuk a kívánt állítást.

$$\begin{aligned}
P(a_1\le\xi_1\le b_1,a_2\le\xi_2\le b_2)&=\lim_{n\to\infty}P\left(a_1\le\xi_1<b_1+\frac1n,a_2\le\xi_2<b_2+\frac1n\right)\\
&=\lim_{n\to\infty}\left(F\left(b_1+\frac1n,b_2+\frac1n\right)-F\left(b_1+\frac1n,a_2\right)-F\left(a_1,b_2+\frac1n\right)+F(a_1,a_2)\right)\\
&=\lim_{n\to\infty}F\left(b_1+\frac1n,b_2+\frac1n\right)-\lim_{n\to\infty}F\left(b_1+\frac1n,a_2\right)\\
&\qquad-\lim_{n\to\infty}F\left(a_1,b_2+\frac1n\right)+F(a_1,a_2).
\end{aligned}$$

Megjegyzem, hogy az eloszlásfüggvény tulajdonságaiból következik, hogy a felírt határértékek, például a $\lim\limits_{n\to\infty}F\left(b_1+\frac1n,b_2+\frac1n\right)$ kifejezés valóban létezik.

Hasonlán,

$$\begin{aligned}
P(a_1<\xi_1<b_1,a_2<\xi_2<b_2)&=\lim_{n\to\infty}P\left(a_1+\frac1n\le\xi_1<b_1,a_2+\frac1n\le\xi_2<b_2\right)\\
&=\lim_{n\to\infty}F(b_1,b_2)-F\left(b_1,a_2+\frac1n\right)\\
&\qquad-\lim_{n\to\infty}F\left(a_1+\frac1n,b_2\right)+\lim_{n\to\infty}F\left(a_1+\frac1n,a_2+\frac1n\right).
\end{aligned}$$

Feladat:

12.) Legyen az $F(x_1,\dots,x_k)=P(\xi_1<x_1,\dots,\xi_k<x_k)$ függvény a $(\xi_1,\dots,\xi_k)$ véletlen vektor eloszlásfüggvénye. Mutassuk meg, hogy

$$I((a_j\le\xi_j<b_j,\ 1\le j\le k)=\sum_{\substack{u_j=\,a_j\text{ vagy }b_j\\ j=1,\dots,k}}(-1)^{\chi(u_1,\dots,u_k)}I(u_1,\dots,u_k),$$

minden (a_j,b_k), $j=1,\dots,k$, számpárokra, ezért a várható érték additivitásából következik, hogy

$$P((a_j\le\xi_j<b_j,\ 1\le j\le k)=\sum_{\substack{u_j=\,a_j \text{ vagy } b_j\\ j=1,\dots,k}}(-1)^{\chi(u_1,\dots,u_k)}F(u_1,\dots,u_k).$$

A fenti formulákban $\chi(u_1,\dots,u_k)$ jelöli az a_j-k számát az $u_1,\dots,u_k$ sorozatban, $I(A)$, $A\subset R^k$ pedig egy A halmaz indikátorfüggvénye, azaz $I(A)(x)=1$, ha $x\in A$, és $I(A)(x)=0$, ha $x\notin A$.
Megoldás: Adva valamely $u_1,\dots,u_k$ számok, definiáljuk a

$$B(u_1,\dots,u_k)=\{(t_1,\dots,t_k)\colon\ t_j<u_j,\ 1\le j\le k\}$$

halmazt, és legyen $I_{B(u_1,\dots,u_k)}(\omega)$ az $\{\omega\colon\ (\xi_1(\omega),\dots,\xi_k(\omega))\in B(u_1,\dots,u_k)\}$ halmaz indikátorfüggvénye. Legyen továbbá $I_{\mathbf{K}}(\omega)$ az $\{\omega\colon\ a_j\le\xi_j<b_j,\ 1\le j\le k\}$ halmaz indikátorfüggvénye. Elég megmutatni, hogy

$$I_{\mathbf{K}}(\omega)=\sum_{\substack{u_j=\,a_j \text{ vagy } b_j\\ j=1,\dots,k}}(-1)^{\chi(u_1,\dots,u_k)}I_{B(u_1,\dots,u_k)}(\omega),$$

mert véve ezután mind a két oldal várható értékét megkapjuk a kívánt azonosságot. Az is világos, hogy $\{\omega\colon\ a_j\le\xi_j<b_j,\ 1\le j\le k\}$ esetében a bizonyítandó azonosság mind a két oldala 1-gyel egyenlő. Azt kell észrevenni, hogy ekkor a jobboldalon az $I_{B(b_1,\dots,b_k)}(\omega)$ tag 1, és az összes többi tag a jobboldalon nulla.
Ha $\omega\in\Omega$ olyan, hogy az $a_j\le\xi_j<b_j$ egyenlőtlenségek nem teljesülnek minden $1\le j\le k$ indexre, akkor a baloldal nulla. Azt kell belátni, hogy ugyanez érvényes a jobboldalon. Világos, hogy a jobboldal nulla, ha $\xi_j(\omega)\ge b_j$ valamilyen j indexre. Azt az esetet kell még nézni, amikor $\xi_j(\omega)<b_j$ minden $1\le j\le k$ indexre, bizonyos j indexekre $a_j\le\xi_j<b_j$, de létezik legalább egy olyan l index, melyre $\xi_j<a_j$. Legyen L a legkisebb ilyen index. Ha párosítjuk az olyan $(-1)^{\chi(u_1,\dots,u_k)}I_{B(u_1,\dots,u_k)}(\omega)$ kifejezéseket, melyeknek az u_j koordinátátái megegyeznek $j\ne L$ esetben, de $u_L=a_L$ a pár egyik és $u_L=b_L$ a pár másik tagjára, akkor az összes ilyen pár hozadéka zéró. Ezért az azonosság ekkor is érvényes, mert mind a bal mind a jobboldal zéró ebben az esetben.

Be lehet látni, hogy az előző előadáson megfogalmazott a valószínűségi változók eloszlásának jellemzéséről szóló tétenek és az azt megelőző Lemmának igaz a következő többdimenziós általánosítása.

Tétel többváltozós valószínűségi változók eloszlásának jellemzéséről. *Egy $F(u_1,\dots,u_k)$ függvény akkor és csak akkor eloszlásfüggvénye alkalmas $\xi_1,\dots,\xi_k$ valószínűségi változóknak valamely $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, ha ez az F függvény teljesíti a következő (i)—(iv) tulajdonságokat.*

(i) $F(u_1,\dots,u_k)$ minden változójának balról folytonos függvénye.

(ii) $\lim\limits_{\substack{u_j\to\infty \\ \text{minden } j=1,\dots,k \text{ számra}}} F(u_1,\dots,u_k)=1.$

(iii) $\lim\limits_{\substack{u_j\to-\infty \\ \text{valamely } 1\le j\le k \text{ számrra}}} F(u_1,\dots,u_k)=0.$ *(Ez úgy értendő, hogy az összes* u_s, $1\le s\le k$, $s\neq j$ *koordinátát rögzítjük, és* $u_j\to-\infty$.*)*

Végül definiáljuk egy az R^k *téren definiált* F *függvényre és egy* $\mathbf{K}=[a_1,b_1)\times\cdots\times[a_k,b_k)$ *téglatestre a*

$$\mu(\mathbf{K})=\mu_F(\mathbf{K})=\sum_{\substack{u_j=\,a_j \text{ vagy } b_j \\ j=1,\dots,k}}(-1)^{\chi(u_1,\dots,u_k)}F(u_1,\dots,u_k)$$

mennyiséget, ahol $\chi(u_1,\dots,u_k)$ *jelöli az* a_j*-k számát az* $u_1,\dots,u_k$ *sorozatban. Ekkor*

(iv) $\mu_F(\mathbf{K})\ge 0$ *minden* $\mathbf{K}$ *téglatestre.*

A fenti tétel bizonyításának legnehezebb része annak megmutatása, hogy ha az $F(x_1,\dots,x_k)$ függvény teljesíti az (i)–(iv) tulajdonságokat, akkor léteznek olyan $\xi_1,\dots,\xi_k$ valószínűségi változók alkalmas $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, amelyek (együttes) eloszlásfüggvénye az $F(x_1,\dots,x_k)$ függvény. Ennek bizonyítása az előző előadás Tétel B eredményének következő általánosításán alapul:

Az előző előadás Tétel B eredményének általánosítása. *Legyen* $F(x_1,\dots,x_k)$ *olyan* k *változós függvény, amelyik teljesíti az (i)–(iv) tulajdonságokat. Ekkor létezik és egyértelműen meghatározott egy olyan* μ_F *Stieltjes mérték az* R^k k*-dimenziós tér Borel mérhető halmazainak* $\mathcal{B}^k$ σ*-algebráján, amelyik nem negatív* σ*-additiv halmazfüggvény ezen a* σ*-algebrán,* $\mu_F(R^k)=1$, *és*

$$\mu_F(\{u_1,\dots,u_k\}\colon\ u_j<x_j,\ 1\le j\le k\})=F(x_1,\dots,x_k)$$

minden $x_1,\dots,x_k$ *valós számra.*

E tétel segítségével a következő módon konstruálhatunk $F(x_1,\dots,x_k)$ eloszlásfüggvénnyel rendelkező valószínűségi változókat. Legyen $(\Omega,\mathcal{A},P)=(R^k,\mathcal{B}^k,\mu_F)$, ahol $\mathcal{B}^k$ jelöli a Borel σ-algebrát az R^k k-dimenziós térben, és μ_F az előző tételben jellemzett az F eloszlásfüggvény által meghatározott Stieltjes mérték. Ebben a valószínűségi mezőben az $(x_1,\dots,x_k)\in R^k$ pontok alkotják az ω elemi eseményeket. A $(\xi_1,\dots,\xi_k)$ valószínűségi változókat a következő módon definiáljuk. Legyen $\xi_j(x_1,\dots,x_k)=x_j$, $1\le j\le k$. Ezen valószínűségi változók együttes eloszlása az $F(x_1,\dots,x_k)$ függvény.

Érvényes továbbá az előző előadáson megfogalmazott eredmények következő többdimenziós változata, amely lehetővé teszi azt, hogy kiszámítsuk véletlen vektorok függvényeinek a várható értékét.

Tétel. *Legyenek* $\xi_1(\omega),\dots,\xi_k(\omega)$ *valószínűségi változók egy* $(\Omega,\mathcal{A},P)$ *valószínűségi mezőn, amelyeknek* $F(x_1,\dots,x_k)=P(\xi_1(\omega)<x_1,\dots,\xi_k(\omega)<x_k)$, *az eloszlásfüggvénye. Legyen* $g(x_1,\dots,x_k)$ *tetszőleges (mérhető)* k*-változós függvény, és definiáljuk az*

$\eta(\omega)=g(\xi_1(\omega),\dots,\xi_k(\omega))$ valószínűségi változót. Ekkor

$$E\eta(\omega)=Eg(\xi_1(\omega),\dots,\xi_k(\omega))=\int g(x_1,\dots,x_k)\mu_F(\,dx_1,\dots,\,dx_k),$$

ahol a fenti integrál Lebesgue–Stieltjes integrált jelöl az F függvény által meghatározott μ_F Stieltjes mérték szerint. Ez a formula akkor érvényes, ha

$$\int|g(x_1,\dots,x_k)|\mu_F(\,dx_1,\dots,\,dx_k)<\infty.$$

Ellenkező esetben az $E\eta$ várható értéket nem definiáltuk.

Az egydimenziós esethez hasonlón a többdimenziós esetben is érdemes bevezetni eloszlásfüggvények sűrűségfüggvényének a fogalmát. Amennyiben valamely valószínűségi változók együttes eloszlásának van sűrűségfüggvénye, akkor ezen valószínűségi változók valamely függvényének a várható értékét ki lehet számolni ezen sűrűségfüggvény szerinti alkalmas integrál segítségével. Az így kapott formula általában jobban használható konkrét feladatokban mint az eloszlások szerinti integrál. Megadom a többdimenziós sűrűségfüggvény definicióját és az előbb jelzett eredmény pontos megfogalmazását.

Többdimenziós eloszlásfüggvény sűrűségfüggvényének definiciója. *Azt mondjuk, hogy egy $F(x_1,\dots,x_k)$ többváltozós eloszlásfüggvénynek létezik $f(u_1,\dots,u_k)$ 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 $-\infty<x_j<\infty$, $1\le j\le k$ számra.

Tétel. *Ha $(\xi_1,\dots,\xi_k)$ valószínűségi változók $F(x_1,\dots,x_k)$ eloszlásfüggvényének létezik $f(u_1,\dots,u_k)$ sűrűségfüggvénye, akkor ez teljesíti minden $g(\cdot)$ k-változós (mérhető) függvényre a következő azonosságot:*

$$\begin{aligned}Eg(\xi_1,\dots,\xi_k)&=\int g(u_1,\dots,u_k)\mu_F(\,du_1,\dots,\,du_k)\\&=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(u_1,\dots,u_k)f(u_1,\dots,u_k)\,du_1\dots\,du_k,\end{aligned}$$

ahol μ_F jelöli az F eloszlásfüggvény által meghatározott Stieltjes mértéket. Ez az azonosság úgy értendő, hogy az annak két oldalán szereplő kifejezés egyszerre létezik vagy nem létezik, ha mind a két kifejezést mint Lebesgue integrált értelmezzük. Ha a jobboldalon szereplő integrál létezik mint (közönséges) Riemann integrál, akkor ez az integrál tekinthető úgy is mint (egy a Riemann integrál értékével megegyező) Lebesgue integrál. Tehát az azonosság ebben a fontos speciális esetben is érvényes.

Az egydimenziós esethez hasonlóan a többdimenziós sűrűségfüggvényeket is lehet jellemezni. Igaz a következő tétel.

Tétel. *Egy k-változós $f(u_1,\dots,u_k)$ függvény akkor és csak akkor sűrűségfüggvénye egy alkalmas k-dimenziós eloszlásfüggvénynek, ha $f(u_1,\dots,u_k)\ge 0$ majdnem minden $(u_1,\dots,u_k)$ pontban, és*

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

Megjegyzés: Egydimenziós eloszlások esetében, elég sima eloszlásfüggvényeknek létezik sűrűségfüggvényük, és az egyenlő az eloszlásfüggvény deriváltjával. Ez a Newton–Leibniz formulából következik. Létezik ehhez az eredményhez hasonló eredmény a többdimenziós esetben is. Eszerint az eredmény szerint, ha $F(x_1,\dots,x_k)$ elég sima k-dimenziós eloszlásfüggvény, akkor létezik $f(x_1,\dots,x_k)$ sűrűségfüggvénye, és az az

$$f(x_1,\dots,x_k)=\frac{\partial^k F(x_1,\dots,x_k)}{\partial x_1\partial x_2\cdots\partial x_k}$$

képlet segítségével számítható ki.

Kiegészítés

A normális sűrűségfüggvény kiintegrálásában ($-\infty$-től ∞-ig) a klasszikus analízis két fontos eredményét használtuk fel. Ezek egyike arról szól, hogyan lehet két egyváltozós integrál szorzatát egy kétváltozós integrálként felírni, a másik pedig arról, hogyan lehet koordinátatranszformációkat alkalmazni többváltozós integrálokban. Ez utóbbi eredmény az $\int_{g(a)}^{g(b)} f(u)\,du=\int_a^b f(g(u))g'(u)\,du$ azonosság többdimenziós megfelelője.

Az első említett eredmény arról szól, hogy két egyváltozós integrál szorzata felírható, mint alkalmas kétváltozós integrál. Pontosabban, ha $f(x)$, $g(x)$ két egyváltozós függvény valamely $[a,b]$ illetve $[c,d]$ intervallumon, $-\infty\le a<b\le\infty$, $-\infty\le c<d\le\infty$ akkor $\int_a^b f(x)\,dx\int_c^d g(x)\,dx=\int\int_{[a,b]\times[c,d]} f(x)g(y)\,dx\,dy$. Általánosabban, ha $f(x,y)$ kétváltozós függvény az $[a,b]\times[c,d]$ téglalapon, akkor $\int_a^b\left(\int_c^d f(x,y)\,dy\right)dx=\int\int_{[a,b]\times[c,d]} f(x,y)\,dx\,dy$. Ebben az azonosságban a baloldalon egyváltozós integrálok szukcesszív alkalmazása szerepel, a jobb-oldalon pedig kétváltozós integrál. Emlékeztetőül felidézem, hogy a többváltozós Riemann integrálokat az egyváltozós integrálokhoz hasonlóan definiálhatjuk. Nevezetesen, adva egy szép $A\subset R^k$ halmaz, és azon egy $f(x_1,\dots,x_k)$ függvény, akkor az $\int_A f(x_1,\dots,x_k)\,dx_1\dots\,dx_k$ integrál definicója érdekében állítsuk elő az A halmazt kis átmérőjű K_j, $\bigcup K_j=A$, halmazok particiójaként, és

írjuk fel a definiálandó integrál $\sum f(u_1^{(j)},\dots,u_k^{(j)})\mathrm{Vol}\,(K_j)$ integrál–közelítő összegeit, ahol $\mathrm{Vol}\,(K_j)$ a K_j halmaz térfogatát jelöli. Az integrált ilyen integrál-közelítő összegek limeszeként definiáljuk, ha a K_j halmazok átmérőienek a szuprémuma nullához tart.

A többváltozós integrál definiciója hasonló az egyváltozós integráléhoz. Egy apró különbség van, amelyikre talán érdemes felhívni a figyelmet. A $\mathrm{Vol}\,(K_j)$ térfogatot és nem előjeles térfogatot jelöl. Az egyváltozós integrálokkal akkor lesz teljes az analógia, ha az $\int_a^b f(x)\,dx$ integrálokat csak $a<b$ esetben definiáljuk. Megjegyzem, hogy az idézett eredmény nemcsak Riemann, hanem általánosabb Lebesgue integrálokra is érvényes. Ezt az analízisban nagyon fontos eredményt Fubini tételnek hívják.

Szükségünk van a következő probléma megoldására is. 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, hogyan tudjuk ezt átírni és kiszámítani, mint egy alkalmas az A tartományon értelmezett függvény integrálját? Elsősorban az $n=2$ változós integrálok kiszámítása érdekel minket, azon belül is az az eset, amikor egy a síkon (pontosabban az $R^2\setminus\{0\}$ halmazon definiált integrált a változók polárkoordinátás átírásának a segítségével kívánunk kiszámolni. Azaz egy a $B=R^2\setminus\{0\}$ halmazon definiált $\int_{R^2} f(y_1,y_2)\,dy_1\,dy_2$ integrált akarunk átírni az $A=(0,\infty)\times(-\pi,\pi]$ halmazon vett integrál alakjában úgy, hogy az $(y_1,y_2)\in R^2\setminus\{0\}$ pontokat átírjuk $y_1=T_1(x_1,x_2)$, $y_2=T_1(x_1,x_2)$ alakban, ahol $(x_1,x_2)\in A$. Annak érdekében, hogy a szokásos jelölésrendszerhez hasonló jelenjen meg érvelésünkben, használjuk az $x_1=r$ és $x_2=\varphi$ betűket a megfelelő változók jelölésére. Ekkor a bevezetett T_1 és T_2 transzformációk az $(x_1,x_2)=(r,\varphi)\in A$ pontokat képezik le az $(y_1,y_2)\in R^2\setminus\{0\}$ halmazba az $y_1=T_1(x_1,x_2)=x_1\cos x_2=r\cos\varphi$ és $y_2=T_2(x_1,x_2)=x_1\sin x_2=r\sin\varphi$ képletek segítségével.

Annak érdekében, hogy a számunkra érdekes eredményt meg tudjuk 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 ki. Ezt később részletesebben elmagyarázom.)

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 $y_k = T_j(x_1, \dots, x_n)$, $1 \le k \le n$, 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*

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

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

Tekintsük az $x = r\cos\varphi$ és $y = r\sin\varphi$ képletekkel megadott polár koordinákra való áttérést, azaz azt, hogy egy síkon adott függvény integrálját hogyan írhatjuk át polár-koordinátarendszerben egy a az $A = \{(r, \varphi)\colon\ r > 0, -\pi < \varphi \le \pi\}$ tartományon definiált függvény integráljává az előző tétel segítségével.

Számoljuk ki a tekintett leképezés Jacobianját. Egyszerű számolás adja, hogy $\frac{\partial x}{\partial r} = \cos\varphi$, $\frac{\partial x}{\partial \varphi} = -r\sin\varphi$, $\frac{\partial y}{\partial r} = \sin\varphi$, $\frac{\partial y}{\partial \varphi} = r\cos\varphi$, ezért a Jacobian értéke

$$\mathcal{J}(r\cos\varphi, r\sin\varphi) = \left|\det \begin{vmatrix} \cos\varphi, & -r\sin\varphi \\ \sin\varphi, & r\cos\varphi \end{vmatrix}\right| = r(\cos^2\varphi + \sin^2\varphi) = r,$$

ahonnan
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = \int_{-\pi}^{\pi}\int_0^{\infty} rf(r\cos\varphi, r\sin\varphi)\, dr\, d\varphi.$$

Értsük meg a fent kimondott tételt. Ennek érdekében tekintsük egy A $n \times n$ méretű $A = (a_{j,l})$, $1 \le j, l \le n$, alakú mátrix $\det A$ determinánsát. Ennek a determinánsnak a következő a szemléletes jelentése: Vegyük az A mátrix $a^{(j)} = (a_{j,1}, \dots, a_{j,n})$ sorait, $1 \le j \le n$, mint vektorokat az n-dimenziós térben. Ekkor $\det A$ egyenlő az $a^{(j)}$, $1 \le j \le n$, vektorok által kifeszített parallelepipedon előjeles térfogatával. Az A mátrix sorvektorai által kifeszített parallelepipedon az n-dimenziós egységkocka képe, ha az A mátrix által meghatározott lineáris transzformációt alkalmazzuk rá. Ezért a $\det A$ kifejezés geometriai tartalma az, hogy az A mátrix által meghatározott lineáris transzformáció hányszorosára nagyítja egy n-dimenziós vektorok által kifeszített parallelepipedon előjeles térfogatát. Ennek a geometriai ténynek fontos következményei vannak. Az előbb tárgyalt eredménynek is ez áll a hátterében.

Tekintsük ugyanis a bizonyítandó azonosság két oldalán szereplő integrál egy-egy integrálközelítését. Ennek érdekében tekintsük az A halmaznak egy kis átmérőjű K_j, $j = 1, 2, \dots,$ halmazokból álló $\bigcup K_j = A$ partícióját valamint minden K_j halmazban egy $\xi^{(j)} = \left(\xi_1^{(j)}, \dots, \xi_n^{(j)}\right) \in K_j$ pontot. Legyen továbbá $\bar{K}_j = \mathbf{T}(K_j)$ a K_j halmaz, $\eta^{(j)} = \left(\eta_1^{(j)}, \dots, \eta_n^{(j)}\right) \in \bar{K}_j$ pedig az $\xi^{(j)}$ pont képe a $\mathbf{T}$ transzformáció hatására. Ekkor a baloldali integrálnak jó közelítése a $\sum_j f(\mathbf{T}\xi^{(j)})\mathcal{J}(\xi^{(j)})\mathrm{Vol}\,(K_j) = \sum_j f(\eta^{(j)})\mathcal{J}(\xi^{(j)})\mathrm{Vol}\,(K_j)$, a jobboldali integrálnak pedig a $\sum_j f(\eta^{(j)})\mathrm{Vol}\,(\bar{K}_j)$ összeg, ahol $\mathrm{Vol}\,(K)$ a K halmaz térfogatát jelöli. Ahhoz, hogy lássuk, hogy ez a két integrálközelítő összeg közel van egymáshoz azt kell megértenünk, hogy $J(\xi^{(j)})\mathrm{Vol}\,(K_j) \sim \mathrm{Vol}\,(\bar{K}_j)$. Ez utóbbi állítást viszont a mátrixok determinánsáról mondottak tulajdonsága, illetve a $\mathbf{T}$ transzformációnak a $\xi^{(j)}$ pontok körüli természetes linearizációja segítségével lehet látni.

Ha a $\mathbf{T}$ transzormációt az $\xi^{(j)}$ pont kis környezetében tekintjük, és ott vesszük ennek közelítését a $\xi^{(j)}$ pont körüli Taylor sorával az első tagig, akkor láthatjuk, hogy miért érvényes a megfogalmazott állítás. Ez a línearizált transzformáció az $\xi^{(j)}$ pont kis környezetét alkotó K_j halmazt közelítőleg az $\eta^{(j)}$ kis környezetét alkotó $\mathbf{K}_j$ halmazba viszi. Tolvábbá e transzformáció alkalmazása egy halmaz térfogatát a $\mathbf{T}$ transzformáció $\xi^{(j)}$-beli derivált mátrixának a determinánsával, azaz az $J(\xi^{(j)})$ számmal szorozza meg. Mivel minket a tartományoknak nem az előjeles, hanem az előjel nélküli térfogata, pontosabban fogalmazva a térfogat abszolut értéke érdekel, ezért a $|J(\xi^{(j)})|$ számmal kell szoroznunk. Így jelenik meg a Jacobian az integráltranszformációs képletben.