A Valószínűségszámítás I. előadássorozat nyolcadik előadása.
2002 március 19.

Összefoglaló:

Valószínűségi változók eloszlásfüggvénye és várható értéke. *(folytatás)*

A Lebesgue integrál teljesíti a Riemann integrálra érvényes additivitási tulajdonság következő természetes általánosítását:

$$\int (c_1\xi_1(\omega) + c_2\xi_2(\omega))\, dP(\omega) = c_1 \int \xi_1(\omega) dP(\omega) + c_2 \int \xi_2(\omega) dP(\omega)$$

minden (integrálható) ξ_1 és ξ_2 valószínűségi változóra és valós c_1, c_2 számra. Ennek az azonosságnak a következménye (valójában átfogalmazása) a következő

Tétel. *Ha két ξ_1, ξ_2 valószínűségi változónak (melyek ugyanazon a valószínűségi mezőn vannak definiálva) létezik várható értéke, c_1 és c_2 két valós szám, akkor a $c_1\xi_1 + c_2\xi_2$ kifejezésnek is létezik várható értéke, és*

$$E\,(c_1\xi_1 + c_2\xi_2) = c_1 E\xi_1 + c_2 E\xi_2.$$

Megjegyzés: Ez az eredmény azt állítja, hogy az általános esetben is érvényes a diszkrét valószínűségi változókról szóló eredmény, mely véletlen összegek várható értékének additivitását fejezi ki. Ebben az esetben sem kell megkövetelni a valószínűségi változók (általános esetben még nem tárgyalt) függetlenségét. A fenti eredmény a várható értéknek eredeti definiciójából adódott, mely a várható értéket mint egy "absztrakt" Lebesgue integrált adta meg. Sokszor hasznosabb a várható érték eredeti definicója helyett azt a múlt előadáson kimondott Fontos Tételnek nevezett eredményt, illetve annak általánosítását használni, mely lehetővé teszi egy valószínűségi változó várható értékének kiszámítását akkor is, ha csak a valószínűségi változó eloszlásfüggvényét ismerjük. Mégis, mint a fenti eredmény mutatja, néha érdemes az eredeti definicióra hivatkozni.

Jegyezzük meg, hogy a várható érték (Lebesgue integrál definiciójából) könnyen látható, hogy amennyiben egy $\xi(\omega)$ valószínűségi változóra teljesül a $P(\xi(\omega) \ge 0) = 1$ feltétel, akkor $E\xi(\omega) \ge 0$. Innen következik, hogy amennyiben két $\xi(\omega)$ és $\eta(\omega)$ valószínűségi változóra teljesül, hogy $P(\xi(\omega) \ge \eta(\omega)) \ge 0$, akkor $E\xi(\omega) \ge E\eta(\omega)$. Valóban, ekkor $E\xi(\omega) - E\eta(\omega) = E\,(\xi(\omega) - \eta(\omega)) \ge 0$. Ennek az észrevételnek egyik következménye az alábbi Markov egyenlőtlenségnek nevezett reláció.

Markov egyenlőtlenség. *Ha $\xi(\omega)$ nem negatív valószínűségi változó, azaz $P(\xi(\omega) \ge 0) = 1$, akkor*

$$P(\xi(\omega) \ge x) \le \frac{E\xi(\omega)}{x} \quad \textit{minden } x > 0 \textit{ számra.}$$

Bizonyítás: Definiáljuk a ξ valószínűségi változó alábbi $\eta(\omega)$ csonkítottját:

$$\eta(\omega) = \begin{cases} x & \text{ha } \xi(\omega) \geq x \\ 0 & \text{ha } 0 \leq \xi(\omega) \leq x \end{cases}$$

Ekkor $P(\xi(\omega) \geq \eta(\omega)) = 1$, ezért $E\xi(\omega) \geq E\eta(\omega) = xP(\xi(\omega) \geq x) = xP(\eta(\omega) \geq x)$, ahonnan következik az állítás.

Definiáljuk általános esetben is valószínűségi változók szórásnégyzetét.

Szórásnégyzet definiciója. *Legyen ξ olyan valószínűségi változó, melyre $E\xi^2 < \infty$. Ekkor e valószínűségi változó szórásnégyzete*

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

(Ha $E\xi^2 = \infty$, akkor vagy nem definiáljuk a ξ valószínűségi változó szórásnégyzetét, vagy azt mondjuk, hogy $\operatorname{Var}\xi(\omega) = \infty$.

Lemma.

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

Bizonyítás:

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

Általában, a szórásnégyzet tulajdonságainak a bizonyítása csak a várható érték tulajdonságait használja. Ezért a diszkrét valószínűségi változók esetében érvényes

$$\operatorname{Var}\left(a\xi + b\right) = a^2 \operatorname{Var}\xi$$

azonosságot hasonlóan lehet bizonyítani az általános esetben.

Feladat:

Általános valószínűségi változók esetében is érvényes a

$$\inf_{-\infty < M < \infty} E(\xi - M)^2 = \operatorname{Var}\xi$$

azonosság.

Megfogalmazzuk és bebizonyítjuk az alábbi Csebisev egyenlőtlenségnek nevezett állítást.

Csebisev egyenlőtlenség. *Minden ξ valószínűségi változóra, melyre $E\xi^2 < \infty$*

$$P(|\xi - E\xi| \geq x) \leq \frac{\operatorname{Var}\xi}{x^2} \quad \textit{minden } x \geq 0 \textit{ számra.}$$

Bizonyítás:

$$P(|\xi - E\xi| \geq x) = P((\xi - E\xi)^2 \geq x^2) \leq \frac{E(\xi - E\xi)^2}{x^2} = \frac{\operatorname{Var}\xi}{x^2}$$

a Markov egyenlőtlenség alapján.

Megjegyzés: A Csebisev egyenlőtlenség azért hasznos, mert a benne szereplő szórásnégyzet sok érdekes esetben könnyen kiszámolható. Felmerülhet az a kérdés, hogy a Csebisev egyenlőtlenség mennyire éles. Erre a kérdésre később még visszatérünk.

Az előző előadáson 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. *Ha egy $F(x)$ eloszlásfüggvénynek létezik $f(u)$ sűrűségfüggvénye, akkor ez teljesíti minden $g(\cdot)$ függvényre a következő azonosságot:*

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

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.

Newton–Leibniz formula. *Legyen $F(x)$ folytonos függvény egy $[a,b]$ véges intervallumon, mely 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.

Megjegyzés 1: 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.

Megjegyzés 2: A mértékelméletben bebizonyították a Newton–Leibniz formula élesebb formáját is. Pontosan leírták azon függvények osztályát, az úgynevezett abszolut folytonos függvényeket, melyekre igaz, hogy a függvény egyenlő a deriváltjának az 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, melyre $f(u) = 1$, 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.

Megjegyzés 3: 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 pontosan jellemezni lehet a sűrűségfüggvényeket. Ezt a jellemzést adja meg az alábbi tétel.

Tétel. *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 elhagyjuk a bizonyítást. Külön tételben megfogalmazzuk azt az eredményt, mely 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. *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

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

Feladat:

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

Néhány fontos folytonos eloszlás.

a.) *Normális eloszlásfüggvény.*

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

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

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.

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

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

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

Feladat:

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 $g(u)=\dfrac{1}{\sqrt{2\pi}|\sigma|}e^{-(u-m)^2/2\sigma^2}$.

A fenti feladat eredménye alapján egy valószínűségi változót akkor nevezünk normális eloszlásúnak, ha van sűrűségfüggvénye, és az $g(u)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-(u-m)^2/2\sigma^2}$ alakban adható meg alkalmas m és $\sigma>0$ számokkal.

Feladat:

Ha ξ standard normális eloszlású valószínűségi változó, akkor$E\xi^{2k-1}=0$ $E\xi^{2k}=1\cdot3\cdots(2k-1)$ minden $k=1,2,\dots$ számra.

b.) *Egyenletes eloszlásfüggvény.*

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

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

$$\operatorname{Var}\xi = E\left(\xi - \frac{b-a}{2}\right)^2 = \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} u^2\,du$$

$$= \frac{1}{b-a}\left[\frac{u^3}{3}\right]_{-\frac{b-a}{2}}^{\frac{b-a}{2}} = \frac{2}{3(b-a)}\left(\frac{b-a}{2}\right)^3 = \frac{(b-a)^2}{12}.$$

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

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

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

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

Feladat:

Számítsuk ki egy λ paraméterű exponenciális eloszlású valószínűségi változó momentumait.

Feladat:

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 $x \ge 0$ és $y \ge 0$ számra.

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 melynek 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.

Feladat:

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.

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, melyeket a bizonyítások felhasznaálnak. 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 megfogalmaztuk az egydimenziós eloszlásfüggvények jellemzését a Tétel A-ban és az azt megelőző Lemmában. Érvényes ezeknek az természetes többdimenziós változata is. Annak érdekében, hogy megértsük hogyan szólnak ezek az eredmények azt kell tisztázni, hogy a lemmában szereplő a) feltételnek, amelyik az eloszlásfüggvény monotonítását mondja ki, hogyan szól a több-dimenziós megfelelője.

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 állítjuk, 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 több-dimenziós 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, melyek 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 több-dimenziós 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:

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.

Feladat:

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 $x \notin A$.

Be lehet látni, hogy az előző előadáson megfogalmazott Tétel A-nak és Lemmának igaz a következő több-dimenziós általánosítása.

Tétel. *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, $1 \le s \le k$, $s \ne j$ 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 $\xi_1, \dots,$ ξ_k valószínűségi változók alkalmas $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyek (együttes) eloszlásfüggvénye az $F(x_1, \dots, x_k)$ függvény. Ennek bizonyítása az előző előadás Tétel B eredményének következő állításán alapul:

Az előző előadás Tétel B eredményének általánositása. *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, mely 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, melyeknek* $F(x_1,\dots,x_k)=P(\xi_1(\omega)<x_1,\dots,\xi_k(\omega)<x_k)$, *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 egy-dimenziós esethez hasonlón a több-dimenziós 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 több-dimenziós sűrűségfüggvény definicióját és az előbb jelzett eredmény pontos megfogalmazását.

Több-dimenziós 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*

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

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 $\displaystyle\int_{g(a)}^{g(b)} f(u)\,du=\int_a^b f(g(u))g'(u)\,du$ azonosság több-dimenziós 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 $\displaystyle\int_a^b f(x)\,dx\int_c^d g(x)\,dx=\int\int_{[a,b]\times[c,d]} f(x)g(y)\,dx\,dy$. Általánosabban, ha $f(x,y)$ kétváltozós függvény az $[a,b]\times[c,d]$ téglalapon, akkor $\displaystyle\int_a^b\left(\int_c^d f(x,y)\,dy\right)dx=\int\int_{[a,b]\times[c,d]} f(x,y)\,dx\,dy$. 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 $\displaystyle\int_A f(x_1,\dots,x_k)\,dx_1\dots dx_k$ 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 analízisben nagyon fontos eredményt Fubini tételnek hívják.

Szükségünk van még a következő probléma megoldására. Ha adva van az n-dimenziós tér egy A tartományának szép, síma 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? Minket speciálisan az $n=2$ eset fog különösen érdekelni, amikor $A = (-\pi,\pi] \times (0,\infty)$, és ha $(x_1,x_2) = (r,\varphi) \in A$, akkor $y_1 = T_1(x_1,x_2) = x_1\cos x_2 = r\cos\varphi$, $y_2 = x_1\sin x_2 = T_2(x_1,x_2) = r\sin\varphi$, és $B = [-\infty,\infty] \times [-\infty,\infty] \setminus \{(0,0)\}$. Azaz a polárkoordinátarendszerre való áttérést vizsgáljuk a síkon.

Annak érdekében, hogy a számunkra érdekes eredményt meg tudjuk fogalmazni először felidézzük egy síma 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 síma 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 elmagyarázzuk.)

Integráltranszformációról szóló tétel. *Legyen adva az n-dimenziós tér egy A tartományának egy síma $y_k = T_k(x_1,\dots,x_n)$, $1 \le k \le n$, 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 az A tartományon egy (integrálható) $f(x_1,\dots,x_n)$ függvény. 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.

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 $\dfrac{\partial x}{\partial r} = \cos\varphi$, $\dfrac{\partial x}{\partial \varphi} = -r\sin\varphi$, $\dfrac{\partial y}{\partial r} = \sin\varphi$, $\dfrac{\partial y}{\partial \varphi} = r\cos\varphi$, 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 szemléletes jelentése van: 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 $\det A$ megegyezik 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. Ezért ezen eredmény szerint a $\det A$ geometriai tartalma azt fejezi ki, hogy az A mátrix hányszorosára nagyítja egy n-dimenziós vektorok által kifeszített parallelepipedon előjeles térfogatát. Ennek a geometriai ténynek fontos következményei vannak. Ez áll az előbb tárgyalt eredmény hátterében is.

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 $\bigcup K_j = A$ particióját 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 $\sum\limits_j f(\mathbf{T}\xi^{(j)})\mathcal{J}(\xi^{(j)})\mathrm{Vol}\,(K_j) = \sum\limits_j f(\eta^{(j)})\mathcal{J}(\xi^{(j)})\mathrm{Vol}\,(K_j)$, a jobboldali integrálnak pedig a $\sum\limits_j f(\eta^{(j)})\mathrm{Vol}\,(\bar K_j)$ 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 köröli természetes linearizációja segítségével lehet látni.

Ugyanis, ha tekintjük a $\mathbf{T}$ transzormációját a $\xi^{(j)}$ pont kis környezetében, 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 a $\xi^{(j)}$ pont kis környezetét alkotó K_j halmazt közelítőleg az $\eta^{(j)}$ kis környezetét alkotó $\mathbf{K}_j$ halmazba viszi. Ez a transzformáció viszont egy halmaz térfogatát a $\mathbf{T}$ transzformáció $\xi^{(j)}$-beli derivált mátrixának a determinánsával, azaz az $J(\xi^{(j)})$ számmal szorozza meg. Innen következik a kívánt állítás.