A Valószínűségszámítás II. előadássorozat első előadása.
2002. szeptember 10.

A valószínűségszámítás megalapozása.

Megbeszéljük a valószínűségszámítás modelljét, legfontosabb fogalmait. Részben felidézem a szükséges mértékelméleti fogalmakat, eredményeket illetve ismertetem a *Valószínűségszámítás alaptételé*-nek nevezett eredmény bizonyítását. Fontosnak tartom annak indoklását, hogy ezek az első látásra absztraktnak tünő eredmények valójában arra is szolgálnak, hogy a heurisztikus érveléseket elmondhatóvá tegyék. Egy másik fontos, megtárgyalandó kérdés az, hogy a valószínűségszámítás itt ismertetett Kolmogorov-féle modellje nem túl megszorító-e, meg lehet-e benne fogalmazni minden értelmesnek tekinthető valószínűségekről szóló problémát. Erre a kérdésre a valószínűségszámítás alaptételének nevezett eredmény ad megnyugtató választ.

A valószínűségszámítás modellje absztrakt megfogalmazásában a következő:

Valószínűségi mező modellje. *Egy $(\Omega,\mathcal{A},P)$ hármas valószínűségi mező, ha Ω egy halmaz, $\mathcal{A}$ egy az Ω bizonyos részhalmazaiból álló σ-algebra, P az $\mathcal{A}$ σ-algebrán értelmezett egyre normált nem negatív σ-additív halmazfüggvény. Az ilyen halmazfügggvényt (egyre normált vagy valószínűségi) mértéknek is szokták hívni a $\mathcal{A}$ σ-algebrán.*

Felidézzük az itt használt fogalmakat.

Algebra és σ-algebra definiciója. *Legyen adva egy Ω halmaz, és azoknak bizonyos $A\subset\Omega$ részhalmazainak $\mathcal{A}$ rendszere. Azt mondjuk, hogy $\mathcal{A}$ algebra, ha tetszőleges $A\in\mathcal{A}$ halmazra ennek komplementere, az $\Omega\setminus A$ halmaz is eleme a $\mathcal{A}$ halmazrendszernek, azaz $\Omega\setminus A\in\mathcal{A}$, és az $\emptyset$ üres halmaz is eleme az $\mathcal{A}$ algebrának, azaz $\emptyset\in\mathcal{A}$, továbbá, ha $A\in\mathcal{A}$ és $B\in\mathcal{A}$ elemei az $\mathcal{A}$ halmazrendszernek, akkor e halmazok metszete és uniója is teljesíti az $A\cap B\in\mathcal{A}$ valamint $A\cup B\in\mathcal{A}$ feltételeket.*

Az $\mathcal{A}$ algebra akkor σ-algebra, ha ezen kívül teljesíti a következő feltételeket is: Ha $A_1,A_2,\dots,$ megszámlálható sok halmaz, melyek az $\mathcal{A}$ algebra elemei, azaz $A_n\in\mathcal{A}$ minden $n=1,2,\dots$ számra, akkor ezek metszete és uniója is benne van az $\mathcal{A}$ σ-algebrában, azaz $\bigcap\limits_{n=1}^{\infty}A_n\in\mathcal{A}$, és $\bigcup\limits_{n=1}^{\infty}A_n\in\mathcal{A}$.

Additív és σ-additív halmazfüggvény definiciója. *Legyen adva egy Ω halmaz részhalmazaiból álló $\mathcal{A}$ σ-algebra. Azt mondjuk, hogy egy $\mu(A)$, $A\in\mathcal{A}$, halmazfüggvény additív, ha bármely diszjunkt $A\in\mathcal{A}$ és $B\in\mathcal{A}$ halmazra $\mu(A\cup B)=\mu(A)+\mu(B)$. Azt mondjuk, hogy ez a μ halmazfüggvény nemcsak additív, hanem σ-additív is, ha minden diszjunkt $A_n\in\mathcal{A}$, $n=1,2,\dots,$ halmazokra $\mu\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=\sum\limits_{n=1}^{\infty}\mu(A_n)$. Ez a σ-additív halmazfüggvény nem-negatív, ha minden $A\in\mathcal{A}$ halmazra $\mu(A)\ge 0$, és egyre normált, ha $\mu(\Omega)=1$.*

Már a bevezető valószínűségszámítás előadásban láttuk, hogyan alkalmazzuk ezt a modellt valószínűségszámítási feladatok megoldására. Az $\omega\in\Omega$ elemeket (pontokat) neveztük elemi eseménynek, az összes elemi esemény unióját, Ω-t a biztos eseménynek, és kijelöltünk az Ω biztos esemény bizonyos részhalmazait, melyek egy $\mathcal{A}$ σ-algebrát alkotnak, az $A \in \mathcal{A}$ halmazoknak (eseményeknek) definiáltuk a valószínűségét, de a $\mathcal{A}$ σ-algebrába nem tartalmazó halmazoknak nem beszéltünk a valószínűségéről. Véges sok diszjunkt halmaz uniójának a valószínűsége megegyezik az egyes halmazok (események) valószínűségének az összegével, azaz a valószínűség additív, sőt ennek az állításnak igaz egy erősebb formája, végtelen sok diszjunkt esemény uniójának a valószínűsége is megegyezik az egyes események valószínűségének az összegével, a valószínűség σ-additív. Az, hogy a későbbiekben ismertetett számolásokat végre tudjuk hajtani, szükségessé teszi, hogy a valószínűség teljesítse nemcsak az additivitás, hanem a σ-additivitási tulajdonságot is.

Első pillanatban arra gondolhatnánk, hogy egyszerübb lenne Ω minden részhalmazának definiálni a valószínűségét. De, mint az már a bevezető valószínűségszámítás előadássorozatban kiderült, ez egyrészt nem lehetséges, másrészt nem is szükséges. Két olyan példát tekintettünk, ahol (bizonyítás nélkül) megtárgyaltuk, hogy egyrészt nem tudjuk Ω minden részhalmazának a valószínűségét definiálni, másrészt erre nincs is szükségünk. Csak olyan halmazok valószínűségére vagyunk kiváncsiak, melyeket egy világosan megfogalmazott szabállyal meg tudunk adni, és az ilyen halmazok nem feltétlenül terjednek ki Ω összes lehetséges részhalmazára. A két korábban tárgyalt eset, ahol ez a probléma felmerült a következő volt.

Tekintsük egy szabályos pénzdarab végtelen sok egymástól független feldobását, azaz legyen Ω az összes lehetséges végtelen hosszú fej-írás sorozat. Definiáljuk a valószínűséget az $\mathcal{A}$ σ-algebrán úgy, hogy annak az eseménynek a valószínűsége, hogy az első n dobásban egy előírt dobássorozat jelenik meg 2^{-n}, (ez jelenti azt, hogy egy szabályos pénzdarab egymástól független dobásai következtek be), és terjesszük ki ezt a definiciót értelemszerüen. Ezt a kiterjesztést (nem triviális) mértékelméleti eredmények segítségével végrehajthatjuk egy σ-algebrára, de nem tehetjük meg Ω minden részhalmazára. Ezt a feladatot természetesen lehet általánosítani más végtelen kisérletsorozatra. Ezt a gyakorlaton fogjuk részletesebben megbeszélni.

A másik ilyen típusú feladat az volt, amikor egy egységnégyzetre egyenletesen ledobtunk egy pontot. Ez azt jelenti, hogy $\Omega = \{(x,y)\colon 0 \le x \le 1,\ 0 \le y \le 1\}$, és egy $A = \{[a,b] \times [c,d]\} \in \Omega\}$ téglalapba esés valószínűsége egyenlő a téglalap területével, azaz $(b-a)(d-c)$. Itt is ki lehet terjeszteni a valószínűséget, (területet) egy elég gazdag osztályra, beszélhetünk annak az eseménynek a valószínűségéről, hogy a ledobott pont az úgynevezett Borel mérhető halmazok valamelyikébe esik, de a valószínűség nem terjeszthető ki Ω minden részhalmazára. Ezt a problémát is megbeszéljük a gyakorlaton.

Nemcsak azt kell tisztázni, hogy bizonyos halmazok (események) valószínűségét megadva létezik az általunk vizsgált σ-algebrán valószínűségi mérték, hanem azt is, hogy az egyértelmű. Ez jelenti ugyanis azt, hogy van értelme megkérdezni, mi a minket érdeklő eseménynek a valószínűsége, mely esemény a vizsgált σ-algebra eleme. Ahhoz, hogy ezt tisztázzuk, szükségünk van bizonyos mértékelméleti eredményekre. Először bizonyítás nélkül megfogalmazok egy eredményt, melyet a Carathéodory-féle kiterjesztési tételnek neveznek. (Ezt valószínűleg tanulták mértékelméletből. Az itteni tárgyalás egyik fontos célja rámutatni arra, hogy bizonyos absztrakt eredmények hasznosak gyakorlati vizsgálatok megalapozására.)

Carathéodory féle kiterjesztési tétel. *Legyen adva egy* $(\Omega, \mathcal{A}, \mu)$ *hármas, ahol* $\mathcal{A}$ *algebra (ennek definicióját felidéztem az előbbiekben). Tegyük fel továbbá, hogy* μ σ-*additív a* $\mathcal{A}$ *algebrán, azaz amennyiben* C_j, $j = 1, 2, \dots$, *diszjunkt halmazok* $\mathcal{C}_j \in \mathcal{A}$, *és ezenkívül* $C = \bigcup\limits_{j=1}^{\infty} C_j \in \mathcal{A}$, *akkor* $\mu(C) = \sum\limits_{j=1}^{\infty} \mu(C_j)$. *Legyen továbbá* $\mu(\Omega) < \infty$. *Ekkor létezik egy legkisebb* $\sigma(\mathcal{A})$ *az* $\mathcal{A}$*-t tartalmazó legszűkebb* σ*-algebra, (ezt nevezzük az* $\mathcal{A}$ *algebra által generált* σ*-algebrának), és a* μ *mérték* egyértelműen *kiterjeszthető a* $\sigma(\mathcal{A})$ σ*-algebrára.*

Megjegyzés: A Carathéodory tétel eredeti alakja általánosabb, úgynevezett (halmaz) gyűrűkre (vagy kissé még általánosabban félgyűrűkre) vonatkozik. Egy $\mathcal{A}$ halmazrendszer akkor gyűrű, ha egyrészt zárt (véges) metszet és unióképzésre, $\emptyset \in \mathcal{A}$, és ha $A \in \mathcal{A}$, $B \in \mathcal{A}$, akkor $A \setminus B \in \mathcal{A}$. Viszont nem követeljük meg, hogy egy $A \in \mathcal{A}$ halmaz komplementere is benne legyen a $\mathcal{A}$ gyűrűben. Számunkra elegendő csak az algebrákra korlátoznunk a figyelmünket. Lényeges, hogy amennyiben $\mathcal{A}$ algebra, akkor abból, hogy $A_n \in \mathcal{A}$, $n = 1, 2, \dots$, halmazok egy sorozatára, nem feltétlenül következik, hogy ezek uniója $\bigcup\limits_{n=1}^{\infty} A_n$ is eleme $\mathcal{A}$-nak. Ezt figyelembe kellett venni az $\mathcal{A}$ algebrán σ-additív mérték definiciójában.

A Carathéodory tételből az következik, hogy a valószínűségi mértéket elegendő egy szűkebb halmazrendszeren, egy olyan (halmaz)algebrán definiálni, amelyik generálja a minket érdeklő σ-algebrát. Ez segít a vizsgálatokban, de teljesen nem oldja meg a felmerülő problémákat, mert tudni kell, hogy az általunk definiált halmazfüggvény σ-additív a tekintett algebrán.

Nem kötelező feladat:

Tekintsük az r-dimenziós euklidészi térnek azon halmazait, melyek előállnak, mint véges sok olyan téglatest uniója, melyek a koordinátatengelyekkel párhuzamos balról zárt jobbról nyílt intervallumok direkt szorzatai. (Ezek az intervallumok lehetnek félegyenesek vagy egyenesek is.) Mutassuk meg, hogy az ilyen halmazok algebrát alkotnak, de nem alkotnak σ-algebrát.

Tekintsük az összes végtelen fej–írás sorozatból álló halmaz olyan részhalmazait, melyek első n koordinátája valamely n számra előírt, $n = 1, 2, \dots$, a többi koordináta pedig tetszőleges. Azon halmazok, melyek előállnak véges sok ilyen halmaz uniójaként algebrát alkotnak. Másrészt az összes olyan sorozatból álló halmaz, melyben a páros helyeken fej van, a páratlan helyeken szereplő dobások pedig tetszőlegesek, nem elemei ennek az algebrának. Ezért ez az algebra nem σ-algebra.

A továbbiak jobb megértése érdekében megbeszéljük az egyik előző félév végén szerepelt vizsgafeladat megoldását.

Feladat:

Van két egy méter hosszú botunk, melyek mindegyikét véletlenül eltörünk egymástól függetlenül úgy, hogy minden töréspont egyformán valószínű. Az első bot hosszabb és a második bot rövidebb darabját összeragasztjuk. Mi annak a valószínűsége, hogy az így összeragasztott bot hossza kisebb, mint 0.9 méter?

A feladatra két természetes megoldás adható. Az egyik az, hogy kiszámítjuk a két összeragasztandó botrész hosszának az eloszlását; ezek egyenletes eloszlásúak, az $\left[\frac{1}{2},1\right]$ illetve a $\left[0,\frac{1}{2}\right]$ intervallumban, függetlenek egymástól, ezért konvolucióval kiszámíthatjuk összegüknek, az összeragasztott bot hosszának a sűrűségfüggvényét és eloszlásfüggvényét. Ezt a megoldási módot senki sem választotta.

Mindenki a geometriai valószínűségek módszerének nevezett eljárást alkalmazta. Ez a következő észrevételen alapul. Legyen az első bot baloldali végének a hossza x, a második bot baloldali végének a hossza y. Annak a valószínűsége, hogy az (x,y) pontpár a $[0,1]\times[0,1]$ négyzet egy $[a,b]\times[c,d]$ téglalap alakú részébe esik, megegyezik e téglalap területének $(b-a)(d-c)$ területével, ezért, illetve a mérték kiterjesztésének egyértelműsége miatt annak valósszínűsége, hogy az (x,y) pontpár beleesik az egységnégyzet valamelyik szép tartományába megegyezik a tartomány területével.

Ezek után meg kell határozni, melyik tartományba kell esnie az (x,y) pontpárnak ahhoz, hogy a két összeragasztott bot összhossza kisebb legyen, mint 0.9. Azok, akik ezt jól csinálták (figyelembe vették, hogy az első botnak a hosszabb, a másodiknak pedig a rövidebb végét kell tekinteni) jó eredményt kaptak, és elfogadtam a megoldásukat. Talán a legegyszerűbb úgy számolni, hogy külön tekintjük a $\{0\le x\le\frac{1}{2},\,0\le y\le\frac{1}{2}\}$ négyzetet, ahol az $\{(x,y)\colon 1-x+y<0.9\}$ tartományt, a $\{0\le x\le\frac{1}{2},\,\frac{1}{2}\le y\le 1\}$ négyzetet, ahol az $\{(x,y)\colon x+y<0.9\}$ tartományt, a $\{\frac{1}{2}\le x\le 1,\,0\le y\le\frac{1}{2}\}$ négyzetet, ahol az $\{(x,y)\colon 1-x+1-y<0.9\}$ tartományt, és a $\{\frac{1}{2}\le x\le 1,\,\frac{1}{2}\le y\le 1\}$ négyzetet, ahol az $\{(x,y)\colon x+1-y<0.9\}$ tartományt kell tekinteni. Felrajzolva a tekintett tartományokat, kapjuk, hogy a következő két háromszög egyesítéséből álló tartomány területét kell tekinteni: Az egyik háromszüg csúcsai a $(0,\,0.1)$, $(0,\,0.9)$ és $(0.5,\,0.4)$ pontok, a másik háromszögé pedig az $(1,\,0.1)$, $(1,\,0.9)$ és $(0.5,\,0.6)$ pontok. Innen könnyű kiszámolni, hogy a keresett valószínűség 0.32.

Volt egy diák, aki meglepő érvelést alkalmazott. Azt mondta, hogy az első összeragasztandó bot hossza 0.5 és 1 között van. Törjünk le belőle 0.5-t, akkor annak hossza is 0 és 0.5 között lesz, tehát a keresett valószínűség megegyezik annak a valószínűségével, hogy a két rövidebb összeragasztott bot összhossza kisebb, mint 0.4. Ezután az (előbbinél egyszerűbben áttekinthető) geometriai meggondolással megmutatta, hogy ennek valószínűsége 0.32. Amikor részletesebb indoklást kértem, csodálkozva nézett rám.

Véletlenül kapott-e ez a diák helyes eredményt vagy jó módszert választott? A válasz az, hogy meg lehet indokolni, hogy ez a módszer jó eredményt ad, de az külön érvelést igényel. Ugyanis a hosszabb botból letörve 0.5 métert egy más hosszúságú botot kapunk, mint akkor, ha a rövidebb botot vennénk. Ez azt sugallja, hogy ebben az esetben más megoldást kapunk, híszen két különböző feladat eredményét hasonlítjuk össze.

De ha alaposabban végiggondoljuk a helyzetet, rájöhetünk arra, hogy mind a rövidebb bot hossza, mind a hosszabb bot hossza minusz 0.5 méter mennyiség ugyanolyan eloszlású, mind a kettő egyenletes eloszlású a $[0, 1/2]$ intervallumban. Ezenkívül ennek a botnak az eloszlása független attól, hogy mi a másik bot hossza. Tehát azért kaptunk helyes megoldást, mert az eredmény attól függött, hogy a két tekintett bot hosszának mi volt az eloszlása. Annak, hogy a vizsgált adott eloszlású valószínűségi változót milyen módon kaptuk meg, nincs jelentősége.

Az előbb elmondottakat érdemes részletesebben tárgyalni. Ugyanis az itt felmerült probléma sokkal általánosabb, mint ahogyan először gondolnánk. A valószínűségi változókról szóló feladatok úgy szólnak, hogy adva vannak bizonyos valószínűségi változók, melyeknek az eloszlását ismerjük, vagy bizonyos események, amelyeknek a valószínűségét ismerjük, és ebben az esetben más események vagy valószínűségi változók viselkedéséről teszünk fel kérdéseket. Viszont nem beszélünk semmit arról, hogy milyen valószínűségi mezőben dolgozunk. Annak érdekében, hogy lássuk, jogunk van így dolgozni, meg kell beszélni, hogy a valószínűségszámításban felvetett problémák függetlenek attól, hogy milyen modellben vizsgáltuk a kérdést. Ezen probléma tisztázásának érdekében vezessünk be először néhány alapvető fogalmat, és idézzünk fel néhány hasznos mértékelméleti eredményt.

Valószínűségi változó fogalma: *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező. Egy ezen a téren értelmezett valós értékű (mérhető) $\xi(\omega)$ függvényt valószínűségi változónak fogunk nevezni. Azaz ξ egy $\Omega \to R^1$ leképezés. Az a megkötés, hogy a függvény legyen mérhető, azt jelenti, hogy minden x valós számra az $\{\omega\colon \xi(\omega) < x\} \in \mathcal{A}$ feltétel teljesül.*

Azt mondjuk, hogy $\xi(\omega) = (\xi_1(\omega), \dots, \xi_n(\omega))$ n-dimenziós valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, ha $\xi(\omega)$ mérhető leképezése az $(\Omega, \mathcal{A}, P)$ térnek az $(R^n, \mathcal{B}^n)$ térre, azaz minden $(x_1, \dots, x_n) \in R^n$ pontra $\{\omega\colon \xi_1(\omega) < x_1, \dots, \xi_n(\omega) < x_n\} \in \mathcal{A}$.

Az előbbi definicióban csak speciális, $\{\omega\colon \xi_1(\omega) < x_1, \dots, \xi_n(\omega) < x_n\}$ alakú halmazokra követeltük meg, hogy a $\mathcal{A}$ σ-algebrában legyenek, azaz azt, hogy ezek olyan események, melyeknek beszélhetünk valószínűségéről. De valójában ebből már minden „értelmes esemény” mérhetősége következik, azaz minden ilyen eseménynek beszélhetünk a valószínűségéről. Ezt a tényt fogalmazza meg az alábbi mértékelméleti eredmény.

Tétel. *Legyen $\xi = (\xi_1, \dots, \xi_n)$ n-dimenziós valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ekkor az n-dimenziós tér tetszőleges B Borel mérhető halmazára $\{\omega\colon (\xi_1(\omega), \dots, \xi_n(\omega)) \in B\} \in \mathcal{A}$.*

Megjegyzés: A teljesség kedvéért felidézzük a Borel mérhető halmaz definicíóját. A mi számunkra elegendő annyit tudni, hogy az n-dimenziós tér minden „szép” halmaza, melyekkel különböző feladatok megoldása során találkozunk, Borel–mérhető. A formális definició a következő. Tekintsük az euklidészi tér részhalmazait, ezen belül az $A(x_1, \dots, x_n) = \{(u_1, \dots, u_n)\colon u_1 < x_1, \dots, u_n < x_n\}$ alakú halmazokat minden $(x_1, \dots, x_n)$ szám-n-esre. Be lehet látni, hogy létezik az R^n tér részhalmazaiból álló legszűkebb σ-algebra, és ezt nevezzük az n-dimenziós tér Borel σ-algebrájának. E Borel σ-algebra elemeit (az ebben a σ-algebrában lévő halmazokat) nevezzük Borel mérhető halmazoknak.

Valójában a valószínűségszámításban nem közvetlenül a valószínűségi változókkal dolgozunk, hanem azok eloszlásával. Ezért felidézzük az alábbi definiciót.

Valószínűségi változó eloszlásfüggvényének a definiciója. *Legyen adva egy $\xi(\omega)$ (valós értékű) valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ennek $F(x)$ eloszlásfüggvényén az $F(x) = P(\{\omega\colon \xi(\omega) < x\})$, $-\infty < x < \infty$, függvényt értjük.*

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*

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

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

Ismerve egy (egy vagy több-dimenziós) valószínűségi változó eloszlásfüggvényét, más események valószínűségét is ki lehet számítani. Ezt a tényt fejezi ki az alábbi mértékelméleti eredmény.

Eloszlások meghatározása az eloszlásfüggvény segítségével. *Legyen egy n-dimenziós $(\xi_1, \dots, \xi_n)$ valószínűségi változó eloszlásfüggvénye $F(x_1, \dots, x_n)$. Az F eloszlásfüggvény egyértelműen meghatározza a $P(\{\omega\colon (\xi_1(\omega), \dots, \xi_n(\omega)) \in B\})$ valószínűséget minden „szép" halmazra, azaz az n-dimenziós tér minden Borel mérhető B halmazára. Részletesebben kifejtve létezik olyan az F eloszlásfüggvény által meghatározott μ_F valószínűségi mérték az R^n tér $\mathcal{B}_n$ Borel σ-algebráján úgy, hogy $\mu_F((-\infty, x_1) \times \cdots \times (-\infty, x_n)) = F(x_1, \dots, x_n)$. Továbbá $P(\{\omega\colon (\xi_1(\omega), \dots, \xi_n(\omega)) \in B\}) = \mu_F(B)$ minden $B \in \mathcal{B}$ halmazra.*

Megjegyzés: A fenti eredményben felhasználtuk azt a fontos mértékelméleti tényt, hogy minden F eloszlásfüggvényhez definiálni lehet az eredményben megfogalmazott tulajdonságokkal rendelkező μ_F Lebesgue–Stieltjes mértéket. Nem térünk ki ennek az állításnak a bizonyítására, csupán megjegyezzük, hogy ezt a mértéket először egyszerű halmazokra (első megközelítésben azt mondhatjuk, hogy téglatestek véges uniójára) definiáljuk, majd a Carathéodory tétel segítségével kiterjesztjük azt a Borel mérhető halmazokra. A Carathéodory tétel alkalmazhatóságához meg kell mutatni, hogy a kiindulásként tekintett halmazokra definiált halmazokon definiált halmazfüggvény σ-additív. Ezt az euklidészi tér bizonyos szép topológiai tulajdonságainak segítségével lehet bizonyítani.

A fenti eredmény azt jelenti, hogy megadva valószínűségi változó eloszlásfüggvényét minden „vele kapcsolatos valószínűséget" meghatároztunk. Ezért természetes valószínűségi problémákban a vizsgálandó valószínűségi változók eloszlását megadni. De ahhoz, hogy ilyen módon dolgozni tudjunk, szükségünk van egy olyan eredményre, amelyik megmondja, hogy egy adott függvény mikor lehet (egy vagy több-dimenziós) eloszlásfüggvény. Erről szólnak az alábbi eredmények.

Lemma. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, és legyen $F(x) = P(\{\omega\colon \xi(\omega) < x\})$, $-\infty < x < \infty$, e valószínűségi változó eloszlásfüggvénye. Az $F(x)$ eloszlásfüggvény teljesíti a következő tulajdonságokat.*

a) *$F(x)$ monoton növekvő függvény.*

b) *$F(x)$ balról folytonos függvény, azaz $\lim\limits_{h\to 0,\, h>0} F(x-h) = F(x)$ minden $-\infty < x < \infty$ számra.*

c) $\lim\limits_{x\to\infty} F(x) = 1$.

d) $\lim\limits_{x\to-\infty} F(x) = 0$.

Tétel A1. *Ha egy $F(x)$ függvény teljesíti az előző lemmában megfogalmazott a)—d) tulajdonságokat, akkor az eloszlásfüggvény. Részletesebben megfogalmazva ez azt jelenti, hogy ha az $F(x)$ függvény teljesíti az a)—d) feltételeket, akkor létezik egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és azon olyan $\xi(\omega)$ valószínűségi változó, melynek az eloszlásfüggvénye ez az $F(x)$ függvény.*

Tétel A2. *Egy $F(u_1, \dots, u_k)$ függvény akkor és csak akkor (együttes) 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.*

Az, hogy egy (egy vagy több-dimenziós) valószínűségi változó teljesíti a fenti tételekben megfogalmazott tulajdonságokat, viszonylag egyszerűen belátható. Megjegyzem, hogy a több-dimenziós esetben megfogalmazott kissé bonyolultabban hangzó (iv) tulajdonság azt fejezi ki, hogy a $(\xi_1(\omega), \dots, \xi_n(\omega))$ véletlen vektor nem-negatív valószínűséggel esik az olyan téglatestekbe, melyeknek oldalai párhuzamosak a koordinátatengelyekkel.

Nehezebb annak a bizonyítása, hogy a tételekben felsorolt tulajdonságokkal rendelkező függvény valóban eloszlásfüggveny. Azt kell megmutatni, hogy létezik egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és azon egy (egy vagy több dimenziós) ξ vektor, amelyiknek eloszlásfüggvénye ez az F függvény. Az alábbi konstrukciót érdemes csinálni. Ha egy n-változós függvényről akarjuk belátni, hogy eloszlásfüggvény, akkor következő módon definiáljunk valószínűségi mezőt és rajta valószínűségi változót. Legyen Ω az R^n n-dimenziós tér, azaz az elemi események legyenek az $\omega = (x_1, \dots, x_n)$ szám-n-esek, legyen a $\mathcal{A}$ σ-algebra az R^n tér Borel-mérhető halmazaiból álló σ-algebra. Definiálnuk kell még a P mértéket a valószínűségi mező konstrukciójának megadása érdekében. Előbb azonban definiáljuk azt a ξ valószínűségi változót, melyről meg akarjuk mutatni, hogy F az eloszlásfüggvénye. Legyen

$$\xi(\omega) = \xi(x_1, \dots, x_n) = (\xi_1(x_1, \dots, x_n), \dots, \xi_n(x_1, \dots, x_n)) = (x_1, \dots, x_n).$$

Ezután definájluk a P mértéket az $\mathcal{A}$ σ-algebrán úgy, hogy legyen

$$P((-\infty, x_1) \times \cdots \times (-\infty, x_n)) = F(x_1, \dots, x_n).$$

Világos, hogy amennyiben belátjuk, hogy létezik ilyen tulajdonságú mérték, akkor belátjuk az állítást. A mértékelméletben belátják, hogy létezik *egyetlen* ilyen mérték. Ezt hívják az F függvény által meghatározott μ_F Stieltjes mértéknek.

A most ismertett eredmények elegendőek annak tisztázásához, hogy mikor beszélhetünk véges sok adott (együttes) eloszlású valószínűségi változóról. Ez az eredmény viszont nem elegendő annak a kérdésnek a tisztázásához, hogy tekinthetjük-e például egy szabályos pénzdarab vagy szabályos kocka *végtelen* dobássorozatának a valószínűségi modelljét. Az ilyen modellek létezésének a bizonyításához további mértékelméleti eredmények szükségesek.

Végtelen független kisérletek modelljét meg lehet konstruálni az alábbi (nem triviális) mértékelméleti eredmény segítségével.

Tétel mértékterek végtelen szorzatáról. *Legyen adva végtelen sok* $(\Omega_n, \mathcal{A}_n, \mu_n)$, $n = 1, 2, \dots$, *valószínűségi mező. Akkor létezik ezeknek végtelen*

$$(\Omega^\infty, \mathcal{A}^\infty, \mu^\infty) = \left(\prod_{n=1}^{\infty} \Omega_n, \prod_{n=1}^{\infty} \mathcal{A}_n, \prod_{n=1}^{\infty} \mu_n \right)$$

direkt szorzata. Részletesebben kifejtve a szorzat olyan valószínűségi mező, melyben $\Omega^\infty = \prod\limits_{n=1}^{\infty} \Omega_n$ *a végtelen* $(\omega_1, \omega_2, \dots)$ *sorozatokból álló tér,* $\mathcal{A}^\infty$ *az* $A_1 \times \cdots A_{n-1} \times A_n \times \Omega_{n+1} \times \cdots$, $A_j \in \mathcal{A}_j$, $1 \le j \le n$, *alakú hengerhalmazok által generált σ-algebra. A* μ_∞ *mértéket úgy definiáljuk, hogy*

$$\mu^\infty (A_1 \times \cdots \times A_{n-1} \times A_n \times \Omega_{n+1} \times \cdots) = \mu(A_1) \dots \mu(A_{n-1}) \mu_n(A_n)$$

minden $n = 1, 2, \dots$ *számra és* $A_j \in \mathcal{A}_j$, $1 \le j \le n$ *halmazra. Be lehet látni, hogy ez a halmazfüggvény egyértelműen kiterjeszthető σ-additív halmazfüggvénnyé a* $\mathcal{A}_\infty$ *σ-algebrára, és így definiáljuk a keresett* μ^∞ *mértéket.*

A valószínűségszámításban előfordul, hogy végtelen sok nem feltétlenül független valószínűségi változót kell tekinteni. Annak bizonyításához, hogy minden természetes esetben létezik megfelelő modell, ahol a vizsgálandó valószínűségi változók léteznek a fenti eredmények nem elégségesek. Most megtárgyaljuk azt az eredményt, amelyik durván szólva azt mondja ki, hogy minden értelmes, valószínűségi változókról szóló probléma tárgyalható a valószínűségszámítás klasszikus modelljén belül. Ezt az eredményt nevezzük a valószínűségszámítás (Kolmogorov–féle) alaptételének. Az eredmény megfogalmazása előtt megfogalmazunk egy hasznos mértékelméleti állítást lemma formájában.

Lemma valószínűségi változó sorozatok eloszlásának jellemzéséről. *Legyen adva valószínűségi változók valamilyen $\xi_1, \xi_2, \dots$ végtelen sorozata egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, és jelölje $F_n(x_1, \dots, x_n)$ az első n $(\xi_1, \dots, \xi_n)$ valószínűségi változó együttes eloszlásfüggvényét, $n = 1, 2, \dots,$ Az $F_n(x_1, \dots, x_n)$, $n = 1, 2, \dots,$ eloszlásfüggvények sorozata meghatározza minden ,,szép végtelen dimenziós A halmazra" a $P((\xi_1, \xi_2, \dots) \in A)$ esemény valószínűségét. Ez kissé részletesebben a következőt jelenti. Tekintsük a számegyenes végtelen sok példányának az $R^\infty = R \times R \times \cdots$ direkt szorzatát, és ennek a $B^n \times R \times R \times \dots$ alakú úgynevezett hengerhalmazait, ahol $n = 1, 2, \dots,$ tetszőleges egész szám, és B^n az n-dimenziós R^n euklidészi tér Borel mérhető részhalmaza. Az ilyen hengerhalmazokat tartalmazó legszűkebb σ-algebrát hívjuk az $\mathcal{A}^\infty$ Borel σ-algebrának az R^∞ téren. E σ-algebra elemeit nevezzük szép halmazoknak, azaz az F_n, $n = 1, 2, \dots$ eloszlásfüggvények, illetve az általuk meghatározott*

$$P(\{\omega\colon (\xi_1(\omega), \dots, \xi_n(\omega)) \in B^n\}) = \mu_{F_n}(B^n)$$

valószínűségek egyértelműen meghatározzák a $P(\{\omega\colon (\xi_1(\omega), \xi_2(\omega), \dots) \in A\})$, $A \in \mathcal{A}^\infty$ alakú események valószínűségét.

A fenti eredmény alapján természetes valószínűségi változók végtelen sorozatának az eloszlását úgy megadni, hogy tetszőleges $n = 1, 2, \dots$ számra megadjuk az első n valószínűségi változó együttes eloszlását. Kérdés, hogy bárhogyan megadhatjuk-e ezeket az eloszlásfüggvényeket. Világos, hogy amennyiben egy esemény valószínűségét két különböző F_n és F_m eloszlásfüggvény segítségével is meg tudjuk határozni, akkor e valószínúség két különböző módon meghatározott értéke meg kell, hogy egyezzen. Ez a követelmény természetessé teszi az alábbi konzisztenciatulajdonság bevezetését. A valószínűségszámítás alaptétele azt mondja ki, hogy akárhogy adjuk meg egymással konzisztens eloszlásfüggvények egy sorozatát, létezik olyan valószínűségi mező és azon olyan valószínűségi változók, melyeknek ezek a függvények az eloszlásfüggvényei. Először megfogalmazzok az említett konzisztenciatulajdonságot.

Eloszlásfüggvények konzisztenciájának a definiciója. *Legyen adva eloszlásfüggvényeknek egy $F_n(x_1, \dots, x_n)$, $n = 1, 2, \dots,$ sorozata. Azt mondjuk, hogy ezen eloszlásfüggvények konzisztensek, ha minden $n = 1, 2, \dots$ egész számra és $x_1, \dots, x_n$ számokra $F_n(x_1, \dots, x_n) = F_{n+1}(x_1, \dots, x_n, \infty)$. (Az $F_{n+1}(x_1, \dots, x_n, \infty)$ mennyiséget az $F_{n+1}(x_1, \dots, x_n, \infty) = \lim\limits_{y \to \infty} F_{n+1}(x_1, \dots, x_n, y)$ relációval definiáljuk.)*

Világos, hogy egy valószínűségi mezőn definiált valószínűségi változók sorozatának $\xi_1, \xi_2, \dots,$ első n eleme által meghatározott F_n eloszlásfüggvények teljesítik a fenti konzisztencia tulajdonságot, mert $F(x_1, \dots, x_n) = P(\{\omega\colon \xi_1(\omega) < x_1, \dots, \xi_n(\omega) < x_n\}) = P\{\omega\colon (\xi_1(\omega) < x_1, \dots, \xi_n(\omega) < x_n, \xi_{n+1}(\omega) < \infty\}) = F_{n+1}(x_1, \dots, x_n, \infty)$. Most megfogalmazzuk a valószínűségszámítás alaptételének nevezett állítást.

A valószínűségszámítás Kolmogorov-féle alaptétele. *Legyen adva $F_n(x_1, \dots, x_n)$, $n = 1, 2, \dots$ eloszlásfüggvényeknek egy konzisztens sorozata. Ekkor létezik egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező és azon $\xi_1, \xi_2, \dots,$ valószínűségi változók sorozata, melyre*

$$P\left(\{\omega\colon (\xi_1(\omega) < x_1, \dots, \xi_n(\omega) < x_n)\}\right) = F_n(x_1, \dots, x_n)$$

minden $n = 1, 2, \dots$ indexre és $x_1, \dots, x_n$ számra.

A véges dimenziós eloszlásfüggvények jellemzését leíró eredmény bizonyításához hasonlóan itt is természetes és egyszerű módon definiálhatunk egy jó valószínűségi mezőt és rajta alkalmas valószínűségi változókat, melyek alkalmas konstrukciót adnak. A fő nehézség annak a megmutatása, hogy a definiált P valószínűség valóban σ-additív, mint azt a valószínűségi mérték definiciójában megköveteltük.

Megfogalmazunk egy tisztán mértékelméleti (illetve a bizonyítás jellege alapján topológiai tételnek is tekinthető) eredményt, melyet az alaptétel mértékelméleti változatának fogunk nevezni, és megmutatjuk, hogy ebből következik a valószínűségszámítás alaptétele. (Valójában nem nehéz belátni, hogy a két eredmény ekvivalens.)

A valószínűségszámítás alaptételének mértékelméleti változata. *Tekintsük a számegyenes $R^\infty = R \times R \times \dots$ végtelen szorzatának $B^n \times R \times R \times \cdots$ alakú úgynevezett hengerhalmazait minden $n = 1, 2, \dots$ számra, ahol $B^n \in \mathcal{B}^n$, ahol $\mathcal{B}^n$ az R^n n-dimenziós tér Borel-mérhető részhalmazainak σ-algebráját jelöli. Az összes hengerhalmaz egy $\mathcal{C}$ algebrát alkot. Legyen adva a hengerhalmazok $\mathcal{C}$ algebráján egy olyan μ nem-negatív halmazfüggvény, $\mu(R^\infty) = 1$, amelyik additív és teljesíti a következő ,,gyengített σ-additívitási tulajdonságot": Minden rögzített n-re a μ halmazfüggvény megszorítása a $B^n \times R \times R \times \cdots$, $B^n \in \mathcal{B}^n$, alakú hengerhalmazokra (az ilyen alakú halmazok σ-algebrát alkotnak minden rögzített n számra.) σ-additív. Ekkor a μ halmazfüggvény σ-additív a $\mathcal{C}$ algebrán.*

Először megmutatjuk, hogy a valószínűségszámitás alaptétele következik annak mértékelméleti változatából. Először megkonstruáljuk azt az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt, ahol a kívánt eloszlású valószínűségi változókat definiálni tudjuk. Legyen $\Omega = R^\infty = R \times R \times \dots$, a számegyenes önmagával vett végtelen direkt szorzata, $\mathcal{A} = \mathcal{B}^\infty$, az R^∞ téren definiált Borel σ-algebra, (ez megegyezik a valószínűségszámítás alaptételének mértékelméleti változatában is tekintett hengerhalmazok által generált σ-algebrával), végül definiáljuk a P mértéket a következő módon: Legyen $P(B^n \times R \times \dots) = \mu_{F_n}(B^n)$, $B^n \in \mathcal{B}^n$, ahol $\mathcal{B}^n$ az R^n n-dimenziós tér Borel σ-algebrája, μ_{F_n} az $F_n(x_1, \dots, x_n)$ n-dimenziós eloszlásfüggvény által generált úgynevezett Lebesgue-Stieltjes mérték az n-dimenziós téren. (Némi munkával be lehet látni, hogy a konzisztencia feltétel biztosítja, hogy ez a definició jogos, nem definiáltuk ugyanannak az eseménynek a valószínűségét különböző módon. Meg lehet például muatatni, hogy $P(B^n \times R \times \dots) = \mu_{F_{n+1}}(B^n \times R) = \mu_{F_n}(B_n)$. Ellenőrizni lehet, hogy az így definiált P halmazfüggvény teljesíti a alaptétel mértékelméleti változatának feltételeit, ezért ezen eredmény alapján σ-additív a $\mathcal{C}$ algebrán. Ezután a Carathéodory tétel alapján ezt a mértéket ki lehet terjeszteni a $\mathcal{B}^\infty$ σ-algebrára.

A tekintett valószínűségi mező elemi eseményei az $\omega = (x_1, x_2, \dots)$ végtelen számsorozatok. Definiáljuk a $\xi_n(\omega) = \xi_n(x_1, x_2, \dots)$, $n = 1, 2, \dots$, valószínűségi változót a $\xi_n(x_1, x_2, \dots) = x_n$ képlettel. Ez a modell teljesíti a valószínűségszámítás alaptételét.

A valószínűségszámítás alaptételének egy természetes általánosítása.

Felmerülhet az itt tárgyalt probléma következő természetes általánosítása. Legyen adva egy tetszőleges számosságú T indexhalmaz. Rendeljünk hozzá a T halmaz minden véges $\{t_1, \dots, t_n\}$ részhalmazához egy $F_{t_1,\dots,t_n}(x_{t_1}, \dots, x_{t_n})$ eloszlásfüggvényt. Mikor tudunk konstruálni egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt és azon megadni ξ_t, $t \in T$, valószínűségi változóknak egy rendszerét úgy, hogy a $(\xi_{t_1}, \dots, \xi_{t_n})$ véletlen vektor eloszlásfüggvénye az $F_{t_1,\dots,t_n}(x_{t_1}, \dots, x_{t_n})$ eloszlásfüggvény legyen a T halmaz minden $\{t_1, \dots, t_n\}$ véges részhalmazára? Röviden megtárgyaljuk ezt a kérdést, de a részletek tisztázása részben a gyakorlatra marad, illetve az érdekődőknek maguknak kell azokat kidolgozni.

Természetes módon definiálható a konzisztencia feltétele ebben az esetben is. Azt a tényt kell megfogalmazni, hogy amennyiban adva van véges sok valószínűségi változó, akkor annak a valószínűsége, hogy ezek közül néhány változó értéke kisebb, mint végtelen, a többieké pedig kisebb mint bizonyos előírt számok nem változik, ha azokat az (üres) feltételeket elhagyjuk, melyek szerint bizonyos valószínűségi változók értékei végtelennél kisebbek. Ennek a tulajdonságnak formális megfogalmazását elhagyom. Maga a konstrukció elég természetes, és a konstrukció helyességének bizonyításához az alaptétel mértékelméleti változatának bizonyos általánosítására van szükség. Röviden leírom a konstrukciót, és rámutatunk, hogy a szükséges álalánosított tétel valójában következménye az alaptétel mértékelméleti változatáanak.

Az $(\Omega, \mathcal{A}, P)$ teret a következő módon definiáljuk: Az Ω halmaz úgy tekinthető, mint a számegyenes „direkt T-ik hatványa". Ez pontosabban ógy fogalmazható meg, hogy az Ω halmaz az összes lehetséges a T téren értelmezett valós értékű függvényből áll. Ez azt jelenti, hogy egy elemi esemény egy a T téren definiált $f(t)$, $t \in T$, valós szám értékű függvény. Definiálunk ezen egy az Ω halmaz részhalmazaiból álló $\mathcal{C}$ (halmaz)algebrát a következő módon. Veszünk egy tetszőleges n egész számot, egy n elemű $\{t_1, \dots, t_n\} \subset T$ halmazt és $B \in \mathcal{B}_n$ halmazt, ahol $\mathcal{B}_n$ az n-dimenziós tér Borel σ-algebráját jelöli, és ezek segítségével definiáljuk a következő $C(t_1, \dots, t_n, B) \subset \Omega$ halmazt: $C(t_1, \dots, t_n, B) = \{f(t), t \in T : (f(t_1), \dots, f(t_n)) \in B\}$. Az ilyen módon előállítható halmazok lesznek elemei a $\mathcal{C}$ algebrának. Ezen halmazok P-valószínűségét a $P(C(t_1, \dots, t_n, B)) = \mu_{F_{t_1,\dots,t_n}}(B)$ képlettel definiáljuk, ahol $\mu_{F_{t_1,\dots,t_n}}$ az $F_{t_1,\dots,t_n}$ eloszlásfüggvény által definiált Lebesgue–Stieltjes mérték. A konzisztenciafeltétel biztosítja azt, hogy ez a definició jogos, a $P(C)$, $C \in \mathcal{C}$, halmaz valószínűsége nem függ attól, hogyan adtuk meg a C halmazt. Szükségünk van az alaptétel mértékelméleti változatának arra az általánosítására, mely szerint ez a P valószínűség σ-additív a $\mathcal{C}$ algebrán. Akkor a Carathéodory tétel szerint ezt a mértéket kiterjeszthetjük a $\mathcal{C}$ által generált $\mathcal{A}$ σ-algebrára.

Az így konstruált $(\Omega, \mathcal{A}, P)$ lesz az a valószínűségi mező, amin dolgozunk, és tetszőleges $t_0 \in T$ indexre a ξ_{t_0} valószínűségi változót a $\xi_{t_0}(\omega) = f(t_0)$ képlettel definiáljuk, ha $\omega = f(t)$, $t \in T$.

Az így definiált valószínűségi változók eloszlása a kívánt lesz. Ennek igazolásában az egyetlen nehezebben bizonyítható lépés annak megmutatása, hogy valóban egy P valószínűségi mértéket definiáltunk. Viszonylag egyszerű belátni, hogy ezt a bizonyítandó állítást redukálni lehet arra az esetre, amikor a T halmaz megszámlálható. Ez a redukált feladat viszont már csak jelölésekben tér el az alaptétel mértékelméleti változatának nevezett eredménytől.

Mennyire tekinthető kielégítőnek, teljesnek a fenti eredmény a valószínűségszámítási modellek megalapozásában? Ez az eredmény leírja, hogy mikor lehet valószínűségi változók egy T paramétercsaláddal indexelt rendszerét megadni azok véges deimenziós eloszlásainak a segítségével. Ez bizonyos szempontból teljesen kielégítő eredmény. Másrészt a sztochasztikus folyamatok elméletében további problémák jelennek meg. Így például egy részecske apró véletlen ütközések során kialakult mozgásának leírására megalkották az úgynevezett Brown mozgás modelljét, mely fontos szerepet játszik a valószínűségszámításban. Felteszik, hogy a Brown mozgás trajektóriája folytonos. Annak megmagyarázása, hogy ez mit jelent, miért szabad feltenni további vizsgálatokat igényel. De mivel ez nem az itteni valószínűségszámítás II. tantárgy témája, hanem a sztochasztikus folyamatok elméletébe tartozik, ezzel a kérdéssel nem fogunk itt foglalkozni.

Kiegészítés: A valószínűségszámítási alaptétel mértékelméleti változatának bizonyítása.

Az alaptétel mértékelméleti változatának bizonyítása a következő gondolaton alapul. Ismeretes (és nem nehéz belátni), hogy egy $\mathcal{C}$ algebrán értelmezett P nem-negatív additív halmazfüggvény akkor és csak akkor σ-additív, ha minden olyan $A_n \in \mathcal{C}$ sorozatra, melyre $\bigcap\limits_{n=1}^{\infty} A_n = \emptyset$, teljesül a $\lim\limits_{n\to\infty} P\left(\bigcap\limits_{j=1}^{n} A_j\right) = 0$ reláció. Ezt a tulajdonságot fogjuk ellenőrizni. Miért érdemes ezt a tulajdonságot vizsgálni? A σ-additívitás bizonyításában a nehézséget az okozza, hogy események végtelen sorozatáról szóló tulajdonságot kell ellenőrizni. Azzal, hogy most végtelen metszetekről szóló tulajdonságot tekintünk, látszólag semmit sem nyerünk. Az, hogy mégis egyszerűbb ennek ellenőrzése, azzal függ össze, hogy bizonyos topológiai eredmények lehetővé teszik a véges metszetekről szóló eredmények felhasználását.

Először felidézzük a következő topológiai eredményt: Ha egy (szép) topologikus térben kompakt K_n, $n = 1, 2, \dots$, halmazok metszete üres, $\bigcap\limits_{n=1}^{\infty} K_n = \emptyset$, akkor létezik olyan véges N egész szám, melyre $\bigcap\limits_{n=1}^{N} K_n = \emptyset$. Nekünk nemcsak kompakt halmazokkal kell dolgoznunk, de a következő, a mértékelméletben ismert eredmény segít az általános esetet erre az egyszerűbben kezelhető esetre redukálni. Legyen adva egy (nem-negatív, véges) μ Lebesgue–Stieltjes mérték az R^n n-dimenziós tér $\mathcal{B}^n$ Borel σ-algebráján. Minden $A \in \mathcal{B}^n$ Borel mérhető halmazra és $\varepsilon > 0$ számra létezik olyan $K \subset A$ kompakt halmaz, melyre $\mu(K) > \mu(A) - \varepsilon$. Megjegyezzük, hogy hasonló gondolat szerepel a Lebesgue–Stieltjes mértékek létezésének a bizonyításában is. Ezenkívül felidézzük a topológia egyik legfontosabb eredményét, a Tyihonov tételt, amelyik szintén hasznos lesz a számunkra: Kompakt topológikus terek direkt szorzata kompakt.

A fenti kompaktsági érveket alkalmazó módszer alkalmazásában a fő nehézség, hogy a $K^n \times R \times R \times \cdots$ alakú hengerhalmazok nem lehetnek kompaktak. E nehézség leküzdésének két természetes módja van. Az egyik lehetőség abból áll, hogy nem használjuk közvetlenül a kompaktsági érveket, hanem indirekt módon megpróbáljuk bebizonyítani a kompaktsági módszerek által sugallt eredményeket. Ezt a bizonyítási módszert alkalmazza a Rényi Alfréd ,,Valószínűségszámítás" című könyvében a IV. fejezet 22. paragrafusában (Kolmogorov-féle alaptétel 235.–238. oldal) szereplő bizonyítás. Itt bizonyos átlós módszer segítségével történik annak a bizonyítása, hogy alkalmas approximáló ,,majdnem kompakt" halmazok véges metszete is üres. Mi más módszert követünk. Megmutatjuk, hogy a topológia kompaktifikációs eljárásának felhasználásával a kívánt eredmény a topológia és mértékelmélelet segítségével bebizonyítható.

Először megfogalmazzuk az alaptétel mértékelméleti változatának egy olyan módosítását, ahol a számegyenes és annak önmagával n-szer illetve végtelen sokszor vett R^n és R^∞ direkt szorzatai helyett a számegyenes egypontos kompaktifikációjának önmagával vett direkt szorzatait tekintjük, és azon fogalmazzuk meg a Tétel természetes megfelelőjét. Jelölje $\bar{R} = R \cup \{\infty\}$ a számegyenes egypontos kompaktifikációját, és legyen $\bar{R}^n = \underbrace{\bar{R} \times \cdots \times \bar{R}}_{n-\text{szer}}$, $\bar{R}^\infty = \bar{R} \times \bar{R} \times \cdots$, jelölje $\bar{\mathcal{B}}^n$ az $\bar{R}^n$ téren definiált természetes Borel σ-algebrát, amelyik például úgy definiálható, mint az $A_1 \times \cdots \times A_n$ alakú halmazok által az $\bar{R}^n$ téren generált σ-algebra, ahol olyan A_j, $1 \le j \le n$, olyan halmazok, melyekre A_j vagy eleme a számegyenes Borel σ-algebrájának vagy $A_j = A_j^0 \cup \{\infty\}$, és A_j^0 eleme a számegyenes Borel σ-algebrájának. Jelölje $\bar{\mathcal{C}}$ az összes $B^n \times \bar{R} \times \bar{R} \times \cdots$ alakú halmazból álló halmazrendszert, ahol $B^n \in \bar{\mathcal{B}}^n$, $n = 1, 2, \dots$. Nem nehéz megmutatni, hogy $\bar{\mathcal{C}}$ algebra. Definiáljuk a $\bar{\mu}$ halmazfüggvényt a $\bar{\mathcal{C}}$ algebrán a következő módon: Egy $B^n \times \bar{R} \times \bar{R} \times \cdots$, $B^n \in \bar{\mathcal{B}}^n$, $n = 1, 2, \dots$, alakú halmazon $\bar{\mu}(B^n \times \bar{R} \times \bar{R} \times \cdots) = \mu((B^n \cap R^n) \times R \times R \times \cdots)$. Nem nehéz belátni, hogy ez a definició jogos, azaz ugyanannak a halmaznak két különböző előállításához ugyanazt $\bar{\mu}$ mértéket rendeli hozzá ez a definició, $\bar{\mu}$ additív a $\bar{\mathcal{C}}$ algebrán, és ezenkívül, ha rögzítünk egy n számot, akkor a $\bar{\mu}$ halmazfüggvény megszorítása a $B^n \times \bar{R} \times \bar{R} \times \cdots$, $B^n \in \bar{\mathcal{B}}^n$ alakú halmazokra σ-addítív. Belátjuk először azt, hogy $\bar{\mu}$ a $\bar{\mathcal{C}}$ algebrán is σ-additív.

Tekintsük olyan $A_n \in \bar{\mathcal{C}}$, $n = 1, 2, \dots$, halmazok sorozatát, melyekre $\bigcap\limits_{n=1}^{\infty} A_n = \emptyset$. Be akarjuk látni, hogy ekkor $\lim\limits_{n\to\infty} \bar{\mu}\left(\bigcap\limits_{j=1}^{n} A_j\right) = 0$. Ennek érdekében megmutatjuk, hogy mindegyik A_n halmazhoz létezik olyan $K_n \subset A_n$ kompakt halmaz, melyre $\bar{\mu}(A_n \setminus$

$K_n) \le 2^{-n}$. Valóban, felírva az A_n halmazt $A_n = B^{m(n)} \times \bar{R} \times \bar{R} \times \cdots$ alakban, ahol $B^{m(m)} \in \bar{\mathcal{B}}^{m(n)}$ alkalmas $m(n)$ index-szel, találhatunk olyan kompakt $K_n^0 \in \mathcal{B}^{m(n)}$, $K_0 \subset B^{m(n)}$, halmazt, melyre $\bar{\mu}\left(\left((B^{m(n)} \setminus K_n^0) \cap R^{m(n)}\right) \times \bar{R} \times \bar{R} \times \cdots\right) \le 2^{-n}$. Itt azt használtuk ki, hogy minden az $R^{m(n)}$ euklidészi tér $\mathcal{B}^{m(n)}$ σ-algebráján definiált véges mértékre bármely Borel mérhető halmaz mértéke tetszőleges pontossággal közelíthető e halmaz által tartalmazott alkalmas kompakt halmaz mértékével. Ekkor $K_n = K_n^0 \times \bar{R} \times \bar{R} \cdots$ kompakt halmaz (mivel $\bar{R}$ is kompakt, és alkalmazhatjuk a Tyihonov tételt), $K_n \in \bar{\mathcal{C}}$, $K_n \subset A_n$ és $\bar{\mu}(A_n \setminus K_n) \le 2^{-n}$. Ekkor $\bigcap\limits_{n=1}^{\infty} K_n = \emptyset$, és mivel a K_n halmazok kompaktak, létezik olyan N index melyre $\bigcap\limits_{j=1}^{N} K_j = \emptyset$, így $\bar{\mu}\left(\bigcap\limits_{j=1}^{n} K_j\right) = 0$ minden $n \ge N$ számra. Ezért $\bar{\mu}\left(\bigcap\limits_{j=1}^{n} A_j\right) \le \sum\limits_{j=n}^{\infty} \bar{\mu}(A_j \setminus K_j) \le \sum\limits_{j=n}^{\infty} 2^{-j} \le 2^{-(n-1)}$ minden $n \ge N$ számra. Innen következik a bizonyítandó $\lim\limits_{n\to\infty} \bar{\mu}\left(\bigcap\limits_{j=1}^{n} A_j\right) = 0$ állítás.

Innen következik, hogy a $\bar{\mu}$ halmazfüggvény mérték a $\bar{\mathcal{C}}$ algebrán. Ezért a Carathéodory tétel alapján $\bar{\mu}$ egyértelműen kiterjeszthető mértékké a $\bar{\mathcal{C}}$ algebra által generált $\sigma(\bar{\mathcal{C}})$ σ-algebrárára. Nyilván $\bar{\mu}(\bar{R}^\infty) = 1$. Annak érdekében, hogy az alaptétel mértékelméleti változatát bebizonyítsuk eredeti formájában vegyük észre, hogy $\bar{\mu}(R^\infty) = 1$. Valóban, $\bar{\mu}(R^\infty) = \lim\limits_{n\to\infty} \bar{\mu}(R^n \times \bar{R} \times \bar{R} \times \cdots) = 1$. Innen következik, hogy $\mu(C) = \bar{\mu}(C)$ minden $C \in \mathcal{C}$ halmazra. Ez utóbbi reláció azért igaz, mert ha $C = B^n \times R \times R \times \cdots$ valamilyen $B \in \mathcal{B}^n$ halmazra, akkor $C = \bar{C} \cap R^\infty$, ahol $\bar{C} = B^n \times \bar{R} \times \bar{R} \times \cdots$. Innen $\bar{\mu}(C) = \bar{\mu}(\bar{C} \cap R^\infty) = \bar{\mu}(\bar{C}) = \mu(C)$. Mivel $\mathcal{C} \subset \sigma(\bar{\mathcal{C}})$ innen következik, hogy μ mérték a $\mathcal{C}$ algebrán. Ez azt jelenti, hogy bebizonyítottuk az alaptétel mértékelméleti változatát.