AZ ERGOD TÉTEL

A valószínűségszámítás, az analízis és a (matematikai) fizika egyik fontos eredménye az ergod tétel. Ezt az eredményt ismertetem, illetve megmutatom, hogyan következik belőle a nagy számok erős törvénye. Célom a legfontosabb eredmények és fogalmak ismertetése. Bár a bizonyításokat is leírom, ezeket nem fogjuk kérdezni a vizsgán.

Az ergod tétel megfogalmazásában szükség van a feltételes várható érték általános, σ-algebrák mint feltételek szerinti definiciójára. Bár ez a nem triviális fogalmat később fogjuk tárgyalni, röviden ismertetem ezt a fogalmat. Ennek keretében felidézem a mértékelméletben tanult Radon–Nikodym tételt, amelynek ismerete megkerülhetetlen a feltételes várható érték általános definiciójának megértéséhez.

Az ergod tétel az úgynevezett dinamikus rendszerek viselkedésével foglalkozik. Ismertetem a dinamikus rendszer fogalmának a definicióját.

Dinamikus rendszer definiciója. *Egy olyan $(X, \mathcal{A}, \mu, T)$ rendszert nevezünk dinamikus rendszernek, amelyben $(X, \mathcal{A}, \mu)$ mértékter, T pedig mértéktartó transzformáció ezen a mértéktéren. Részletesebben megfogalmazva ez azt jelenti, hogy adva van egy X alaphalmaz, annak bizonyos 'szép' halmazainak egy $\mathcal{A}$ rendszere, amely σ-algebra, egy μ mérték ezen az $\mathcal{A}$ σ-algebrán és egy T: $X \to X$, mérhető és a μ mértékre nézve mértéktartó transzformáció az X téren. A T transzformáció mérhetősége azt jelenti, hogy minden $A \in \mathcal{A}$ (mérhető) halmaz $T^{-1}(A) = \{y\colon\ Ty \in A\}$ ősképére $T^{-1}(A) \in \mathcal{A}$. Az, hogy a T transzformáció mértéktartó a μ mértékre nézve azt jelenti, hogy $\mu(T^{-1}(A)) = \mu(A)$ minden $A \in \mathcal{A}$ halmazra.*

1. megjegyzés. Bár a definicióban nem követeltük meg, hogy a μ mérték valószínűségi (azaz véges, 1-re normált) mérték legyen, a továbbiakban mégis csak ilyen dinamikus rendszerekkel fogunk foglalkozni, és a μ mértéknek ezt a tulajdonságát mindig fel fogjuk tenni.

2. megjegyzés. Mind a mérhetőség, mind a mértéktartás definiciójában a T transzformáció T^{-1} inverze és nem maga a T transzformáció szerepelt. Mivel nem tettük fel, hogy a T transzformáció invertálható, azaz lehetséges, hogy $T(x) = T(y)$ valamely $x \neq y$ párra, ezért nem mindegy, hogy a T transzformáció vagy annak (esetleg többértékű) T^{-1} inverze szerepel-e a definicióban. Az általunk bevezetett T^{-1} transzformáció segítségével megadott definiciók bizonyultak szerencsésnek.

Legyen adva egy $U(x)$, $x \in X$, valós értékű (mérhető) függvény az $(X, \mathcal{A}, \mu, T)$ dinamikus rendszerben. Ekkor definiálhatjuk az $U(x)$ függvény $T^k U(x) = U(T^k x)$ eltoltjait minden $k = 0, 1, \dots$ számra, ahol $T^k x$ azt jelenti, hogy a T transzformációt k-szor alkalmazzuk szukcesszive az x pontra. Speciálisan $T^0 x = x$ a $k = 0$ esetben. Az ergod tétel azt mondja ki, hogy nagyon általános feltételek mellett a $\frac{1}{n} \sum\limits_{k=0}^{n-1} U(T^k x)$ átlagoknak a μ mérték szerint majdnem minden x pontban létezik határértéke, ha $n \to \infty$. Továbbá ezt a határértéket jellemezni tudjuk. Természetes lenne azt várni, hogy ez a limesz az $U(x)$ függvény $\int_X U(x)\mu(\,dx)$ integrálja azaz az $(X, \mathcal{A}, \mu)$ valószínűségi mezőn definiált $U(x)$ valószínűségi változó várható értéke. A helyzet azonban ennél bonyolultabb. Csak bizonyos plusz feltételek teljesülése esetén érvényes a

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} U(T^k x) = \int_X U(x)\mu(\,dx) \quad \text{a } \mu \text{ mérték szerint 1 valószínűséggel} \tag{1}$$

azonosság. Általános esetben is létezik a az (1) kifejezés baloldalán szereplő kifejezésnek limesze a μ mérték szerint 1 valószínűséggel, de az egy bonyolultabb kifejezés. Az (1) formula teljesülését a fizikusok úgy szokták interpretálni, hogy az időbeli átlag (limesze) egyenlő a térbeli átlaggal. Ugyanis az (1) formula baloldalán szereplő kifejezés a következő módon is megadható. Vezessük be a következő úgynevezett empirikus μ_n mértéket minden $n = 1, 2, \dots$ számra és $x \in X$ pontra. $\mu_n(A|x) = \frac{1}{n}\#\{j\colon\ 0 \le j \le n-1,\ T^j x \in A\}$ minden $A \in \mathcal{A}$ halmazra, ahol $\#(C)$ a C halmaz számosságát jelöli. Ezzel a jelöléssel $\frac{1}{n} \sum\limits_{k=0}^{n-1} U(T^k x) = \int_X U(v)\mu_n(\,dv|x)$ és az, hogy az (1) formula igaz $U(\cdot)$ függvények egy nagy családjára (minden a μ mérték szerint integrálható függvényre) úgy is interpretálható, hogy az $U(\cdot)$ függvények időbeli átlaga, azaz a $\mu_n(\cdot|x)$ empirikus mérték szerinti integrálja tart e függvények $\int U(x)\mu(\,dx)$ térbeli átlagához.

Az (1) formulában megfogalmazott (μ mérték szerinti) egy valószínűségi konvergenciát kifejező azonosságot, illetve annak általánosabb alakját, amely megadja az (1) formula határértékét az általános esetben nevezik ergod tételnek. Ennek pontos megfogalmazásához be kell vezetni a feltételes várható érték általános definicióját. A nagy számok törvénye következik az ergod tételből, de ez külön indoklást igényel.

A pontos eredmény megfogalmazásához először be kell vezetni az invariáns halmazok fogalmát, és be kell bizonyítani azok egy egyszerű tulajdonságát.

Dinamikus rendszer invariáns halmazainak a definiciója. *Egy $(X, \mathcal{A}, \mu, T)$ dinamikus rendszer valamely Egy $A \in \mathcal{A}$ halmazát e rendszer invariáns halmazának nevezünk, ha $T^{-1}(A) = A$.*

Igaz a következő (egyszerű) lemma.

Lemma az invariáns halmazok viselkedéséről. *Egy dinamikus rendszer invariáns halmazai σ-algebrát alkotnak, azaz, ha A invariáns halmaz akkor annak komplementere, $X \setminus A$ is az, és ha $A_1, A_2, \dots$ invariáns halmazok, akkor a $\bigcup\limits_{n=1}^{\infty} A_n$ és $\bigcap\limits_{n=1}^{\infty} A_n$ halmazok is invariánsak.*

Bizonyítás. Az állítás egyszerű következménye a T^{-1} inverz transzformáció tulajdonságainak.

Jelölje $\mathcal{I} \subset \mathcal{A}$ az invariáns halmazok σ-algebráját, és vezessük be a következő definiciót.

Ergodikus dinamikus rendszerek definiciója. *Egy $(X, \mathcal{A}, \mu, T)$ dinamikus rendszert ergodikusnak nevezünk, ha e rendszer invariáns halmazainak $\mathcal{I}$ σ-algebrája triviális a következő értelemben. Minden $A \in \mathcal{I}$ halmazra $\mu(A) = 0$ vagy $\mu(A) = 1$.*

Vegyük észre, hogy amennyiben $x \in A$ egy $A \in \mathcal{I}$ invariáns halmazra, akkor a $\frac{1}{n} \sum\limits_{k=0}^{n-1} U(T^k x)$ átlagokban szereplő $T^k x$ pontok mindegyike az A halmazban van. Valóban, ha A a T transzformáció invarians halmaza, akkor A a T^k, $k = 1, 2, \dots$, transzformációknak is invariáns halmaza, és ugyanez érvényes az A halmaz $X \setminus A$ komplementerére is. Ha a $T^k x \notin A$, azaz $T^k x \in X \setminus A$ reláció lenne érvényes valamely $k \ge 0$ számra, akkor abból az következne, hogy $x \in X \setminus A$, ami ellentmond az $x \in A$ feltételnek. Ez a tény azt jelzi, hogy az (1) formulában megfogalmazott azonosság nem lehet érvényes az általános esetben. Ugyanis természetesebb lenne a jobboldalon az $\int_X U(x)\mu(\,dx)$ kifejezés helyett az $\frac{1}{\mu(A)} \int_X U(x)\mu(\,dx)$ kifejezést, azaz az U függvénynek az A halmazbeli átlagát venni, ha egy olyan $x \in X$ pontot tekintünk a baloldalon, amelyre $x \in A$ valamely $A \in \mathcal{I}$ invariáns halmazzal. Sőt, mivel általában nincs egy legkisebb az x pontot tartalmazó invariáns A halmaz, a képlet még bonyolultabb. A pontos eredményt a az invariáns halmazok által generált $\mathcal{I}$ σ-algebra szerinti feltételes várható érték segítségével lehet megfogalmazni. Ez az eredmény speciálisan azt adja, hogy az (1) formula érvényes akkor, ha az $(X, \mathcal{A}, \mu, T)$ dinamikus rendszer ergodikus, és az $U(x)$ függvény integrálható a μ mérték szerint.

Egy valamely $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált ξ valószínűségi változó várható értéke szemléletesen azt fejezi ki, hogy mit várunk, mekkora lesz e változó értéke. Ez a megitélésünk változhat akkor, ha további információkkal rendelkezünk, bizonyos eseményekről tudjuk, hogy bekövetkeztek-e vagy sem. Vegyük az összes olyan eseményt, amelynek be vagy be nem következését meg tudjuk állapítani. A feltételes várható érték általános definiciójában azt akarjuk kifejezni, hogy ezen rendelkezésünkre álló plusz ismeretek alapján hogyan értékelnénk a ξ valószínűségi változó várható nagyságát. Vagyük észre, hogy az általunk megfigyelhető események együttese (azon események, amelyekről tudjuk, hogy bekövetkeztek-e vagy nem) σ-algebrát alkotnak. Valóban, ha egy esemény bekövetkeztét meg tudjuk állapítani, akkor meg tudjuk állapítani ezen esemény komplementerének a bekövetkeztét vagy be nem következtét is, ha megszámlálható sok esemény mindegyikéről külön külön el tudjuk dönteni, hogy bekövetkezett-e, akkor el tudjuk dönteni, hogy bekövetkezett-e ezen események mindegyike, vagy igaz-e az, hogy ezen események közül legalább egy bekövetkezett. Ezen tények jelentik azt, hogy a megfigyelhető események σ-algebrát alkotnak.

Ezért az általános esetben egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált ξ valószínűségi változónak egy $\mathcal{F} \subset \mathcal{A}$ σ-algebra szerinti $E(\xi|\mathcal{F})$ feltételes várható értékét definiáljuk, ahol $\mathcal{F}$ szemléletes tartalma az általunk összegyüjtött információ. Azt, hogy a feltételes várható érték ezen információktól függ a feltételes várható értéknek az a tulajdonsága fejezi ki, hogy az $E(\xi|\mathcal{F})$ feltételes várható érték egy $\mathcal{F}$ mérhető valószínűségi változó.

Az $E(\xi|\mathcal{F})$ feltételes várható érték definiciójának megadásához fel kell idéznünk a mértékelmélet egyik fontos és mély eredményét, a Radon–Nikodym tételt. Először értsük meg azt a kérdést, amelyre a Radon–Nikodym tétel adja meg a választ.

Kérdés. *Legyen adva egy μ (σ-véges) mérték egy $(\Omega, \mathcal{F})$ mérhető téren és egy az $(\Omega, \mathcal{F})$ téren definiált $\mathcal{F}$ mérhető $g(\cdot)$ függvény, amelyre $\int |g(\omega)| \, \mu(\,d\omega) < \infty$. Ekkor a $\nu(A) = \int_A g(\omega)\mu(\,d\omega)$ halmazfüggvény, $A \in \mathcal{F}$, előjeles mérték az $(\Omega, \mathcal{F})$ téren. (Azaz, ν σ-additív halmazfüggvény. Lássuk be, hogy ez következik a Lebesgue tételből.) Kérdés: Mely ν előjeles mértékek állíthatóak elő ilyen módon, azaz az $(\Omega, \mathcal{F})$ téren levő ν mértékek közül melyekhez létezik olyan $g(\omega)$ $\mathcal{F}$ mérhető függvény, amelyre $\nu(A) = \int_A g(\omega)\nu(\,d\omega)$ minden $A \in \mathcal{F}$ halmazra? Ha létezik ilyen $g(\omega)$ függvény, akkor az meg van-e határozva egyértelműen?*

Világos, hogy egy a fenti integrálelőállítás segítségével definiálható ν előjeles mérték véges, azaz $|\nu(A)| \le K$ minden $A \in \mathcal{F}$ halmazra egy alkalmas $K < \infty$ számmal. (A $K = \int |g(\omega)|\mu(\,d\omega)$ egy alkalmas választás.) Ezenkívül, ha $\mu(A) = 0$ valamely $A \in \mathcal{F}$ halmazra, akkor $\nu(A) = \int_A g(\omega)\mu(\,d\omega) = 0$. Ez az észrevétel tette természetessé a következő definició bevezetését.

Egy előjeles mérték egy mérték szerinti abszolut folytonosságának a definiciója. *Legyen μ (esetleg σ)-véges (σ-additív) mérték és ν véges (σ-additív) előjeles mérték egy Ω téren értelmezett $\mathcal{F}$ σ-algebrán. Azt mondjuk, hogy a ν előjeles mérték abszolut folytonos a μ mérték szerint, ha minden olyan $B \in \mathcal{F}$ halmazra, amelyre $\mu(B) = 0$ a $\nu(B) = 0$ reláció is teljesül.*

A Radon–Nikodym tétel azt mondja ki, hogy a ν előjeles mérték fent megfogalmazott abszolut folytonossága (és végessége) nemcsak szükséges, hanem elégséges feltétele is ν kívánt alakú előállításának. Továbbá ez az előállítás lényegében egyértelmű. Ez az egyértelműség a következőt jelenti. Világos, hogy ha a $g(\omega)$ függvényt megváltoztatjuk egy a μ mérték szerint null mértékű halmazon, akkor a $\nu(A) = \int_A g(\omega)\mu(\,d\omega)$ integrálok értékei nem változnak. Tehát ennyi szabadságunk van a $g(\cdot)$ függvény megválasztásában. A Radon–Nikodym tétel azt is kimondja, hogy több szabadságunk már nincs. Végül szeretném hangsúlyozni, hogy a Radon–Nikodym tételben megjelenő $g(\cdot)$ függvény $\mathcal{F}$ *mérhető*. Ezt azért fontos érteni, mert sok esetben olyan $(\Omega, \mathcal{F})$ térben alkalmazzuk a Radon–Nikodym tételt, ahol természetes módon megjelenik az $\mathcal{F}$ σ algebra mellett egy szintén az X téren definiált bővebb $\mathcal{A}$ σ-algebra is, amelyre $\mathcal{F} \subset \mathcal{A}$. Fontos érteni, hogy olyan $g(\cdot)$ függvényeket tekintünk ilyen esetekben is, amelyek a (szűkebb) $\mathcal{F}$ σ-algebra szerint is mérhetőek.

Radon–Nikodym tétel. *Legyen adva egy Ω tér, rajta egy $\mathcal{F}$ σ-algebra, továbbá ezen az $\mathcal{F}$ σ-algebrán egy μ (σ-véges) mérték és ν (véges) előjeles mérték. (Az, hogy egy ν előjeles mérték véges azt jelenti, hogy $|\nu(A)| < \infty$ minden $A \in \mathcal{F}$ halmazra.) Tegyük fel, hogy a ν előjeles mérték abszolut folytonos a μ mértékre nézve. Akkor létezik olyan az Ω téren definiált valós értékű $f(\omega)$ függvény, amelyre $\int |f(\omega)| \, d\mu(\omega) < \infty$, és teljesíti a következő két tulajdonságot.*

a) *$f(\omega)$ $\mathcal{F}$ mérhető.*
b) *$\nu(C) = \int_C f(\omega) \, d\mu(\omega)$ minden $C \in \mathcal{F}$ halmazra.*

Továbbá ez az $f(\omega)$ függvény egyértelműen meg van határozva a következő értelemben. Ha két $f_1(\omega)$ és $f_2(\omega)$ $\mathcal{F}$ mérhető függvény teljesíti a fenti relációt minden $C \in \mathcal{F}$

halmazra, akkor $f_1(\omega) = f_2(\omega)$ a μ mérték szerint majdnem minden $\omega \in \Omega$ pontban.

A tételben kimondott tulajdonságú $f(\omega)$ függvényt a ν előjeles mérték μ mérték szerinti Radon–Nikodym deriváltjának nevezik az irodalomban, és $\frac{d\nu}{d\mu}(\omega)$-val jelölik.

A Radon–Nikodym tétel felhasználásával tudjuk egy ξ valószínűségi változó $\mathcal{F}$ σ-algebra szerinti $E(\xi|\mathcal{F})$ feltételes várható értékét definiálni az általános esetben. Ezt a feltételes várható értéket csak akkor definiáljuk, ha $E|\xi| < \infty$.

Egy valószínűségi változó valamely σ-algebra szerinti feltételes várható értékének a definiciója. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, azon egy ξ, $E|\xi| < \infty$, valószínűségi változó és egy $\mathcal{F} \subset \mathcal{A}$ σ-algebra. Az $\eta = E(\xi|\mathcal{F})$ feltételes várható érték az az η valószínűségi változó, amelyre*

a) η $\mathcal{F}$ mérhető valószínűségi változó.

b) $\int_B \eta(\omega)P(\,d\omega) = \int_B \xi(\omega)P(\,d\omega)$ minden $B \in \mathcal{F}$ halmazra.

A következő eredmény azt mondja ki, hogy a fenti definició értelmes.

Tétel a feltételes várható értéket definiáló definició jogosságáról. *Létezik a feltételes várható érték definiciójában megfogalmazott a) és b) tulajdonságot teljesítő η valószínűségi változó, és az egy valószínűséggel meghatározott. Azaz, ha η és $\bar\eta$ két olyan valószínűségi változó, amelyek mindegyike teljesíti mind az a) mind a b) tulajdonságot, akkor $P(\eta = \bar\eta) = 1$.*

A tétel bizonyítása. Definiáljuk a $\nu(B) = \int_B \xi(\omega)P(\,d\omega)$, $B \in \mathcal{F}$, halmazfüggvényt az $\mathcal{F}$ σ-algebrán. Ekkor ν egy véges és a P mértékre abszolut folytonos előjeles mérték a $\mathcal{F}$ σ-algebrán. Ezért létezik az $\eta(\omega) = \frac{d\nu}{dP}(\omega)$ Radon–Nikodym derivált a $\mathcal{F}$ σ-algebrán, amely teljesíti mind az a) mind a b) tulajdonságot. Ha η és $\bar\eta$ teljesíti mind az a) mind a b) tulajdonságot, akkor $\eta - \bar\eta$ $\mathcal{F}$ mérhető valószínűségi változó, és $\int_B (\eta - \bar\eta)\,dP = 0$ minden $B \in \mathcal{F}$ halmazra, speciálisan a $B_1 = \{\omega\colon\ \eta(\omega) - \bar\eta(\omega) > 0\}$ $B_2 = \{\omega\colon\ \eta(\omega) - \bar\eta(\omega) < 0\}$ halmazokra is. Innen $P(B_1) = P(B_2) = 0$, azaz $P(\eta \neq \bar\eta) = 0$. Innen következik a tétel állítása.

Megjegyzés. *Legyen $(X, \mathcal{A}, \mu, T)$ egy ergodikus dinamikus rendszer, $U(\cdot)$ egy $\mathcal{A}$ mérhető függvény, amelyre $\int_X U(v)\mu(\,dv) < \infty$. Ekkor $E(U(x)|\mathcal{I}) = EU$, ahol az $U(\cdot)$ valószínűségi változó várható és feltételes várható értékét az $(X, \mathcal{A}, \mu)$ valószínűségi mezőn tekintjük, és $\mathcal{I}$ az invariáns halmazokból álló σ-algebra.*

Megfogalmazom az ergod tételt.

Ergod tétel. *Legyen $(X, \mathcal{A}, \mu, T)$ egy dinamikus rendszer, $U(x)$ egy $\mathcal{A}$ mérhető valós értékű függvény, amelyre $\int_X |U(x)|\mu(\,dx) < \infty$. Ekkor*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} U(T^j x) = E(U|\mathcal{I})(x) \quad \text{a } \mu \text{ mérték szerint majdnem minden } x \in X \text{ pontra,}$$

ahol $\mathcal{I}$ az invariáns halmazok σ-algebrája. Ha az $(X,\mathcal{A},\mu,T)$ dinamikus rendszer ergodikus, akkor $E(U|\mathcal{I})(x)=EU=\int U(x)\mu(\,dx)$.

Meg akarom mutatni, hogy a nagy számok erős törvénye következik az ergod tételből. Független, egyforma eloszlású valószínűségi változók $\xi_1,\xi_2,\dots$ sorozatának viselkedése mutat némi hasonlóságot a dinamikus rendszerekhez. Ha a $\xi_1,\xi_2,\dots$ sorozatnak megfeleltetjük a $\xi_2,\xi_3,\dots$ sorozatot akkor az hasonlít a mértéktartó shift transzformációhoz. De ahhoz, hogy az ergod tételt alkalmazni tudjuk egy olyan rendszerrel kell dolgozni, ahol valódi (mértéktartó) shift transzformáció van. A nagy számok erős törvényét vissza lehet vezetni egy ilyen rendszer vizsgálatára. Érdemes ezt a visszavezetést egy általánosabb modellben is megcsinálni. Bevezetem a következő fogalmat.

Stacionárius sorozatok definiciója. *Legyen adva $\xi_0,\xi_1,\xi_2,\dots$ valószínűségi változók egy sorozata egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Azt mondjuk, hogy ez a sorozat stacionárius, ha minden $0\le n_1<n_2<\cdots<n_k$ és $m\ge 1$ számokra a $(\xi_{n_1},\xi_{n_2},\dots,\xi_{n_k})$ és $(\xi_{n_1+m},\xi_{n_2},\dots,\xi_{n_k+m})$ véletlen vektorok eloszlása megegyezik.*

Vegyük észre, hogy független, egyforma eloszlású $\xi_0,\xi_1,\dots$ valószínűségi változók sorozata stacionárius.

Feladat.

Legyen $\xi_0,\xi_1,\dots$ valószínűségi változók egy Gauss sorozata, azaz negy olyan sorozat, amelyre $(\xi_{n_1},\dots,\xi_{n_k})$ (többváltozós) normális eloszlású véletlen vektor minden $\{n_1,\dots,n_k\}\subset\{0,1,2,\dots\}$ indexhalmazra. Ha $E\xi_n=M$ valamely M számmal, és $E\xi_i\xi_j=a(i-j)$, azaz $E\xi_i\xi_j$ csak az $i-j$ különbségtől függ, akkor a $\xi_0,\xi_1,\dots$ sorozat stacionárius.

Legyen $(X,\mathcal{A},\mu,T)$ egy dinamikus rendszer, $U(x)$ egy mérhető függvény ezen a téren. Definiáljuk a $\xi_k(x)=U(T^kx)$, $k=0,1,2,\dots$, valószínűségi változók sorozatát az $(X,\mathcal{A},\mu)$ valószínűségi mezőn. Ekkor a $\xi_0,\xi_1,\dots$ sorozat stacionárius. Megfordítva, adva egy stacionárius sorozat tudunk konstruálni egy dinamikus rendszert, és abban egy valószínűségi változót és annak eltoltjait úgy, hogy ezek hasonlítanak az eredeti stacionárius sorozathoz. Ilyen módon az ergod tételből érdekes eredményeket kaphatunk az eredeti stacionárius sorozatra.

Vezessük be az $(R^\infty,\mathcal{B}^\infty)$ mérhető teret, ahol az R^∞ halmaz pontjai az $x=(x_0,x_1,\dots)$ valós számokból álló végtelen sorozatok, és $\mathcal{B}^\infty$ a Borel σ-algebra az R^∞ téren, azaz az $\{x_0<a_0,x_1<a_1,\dots,x_k<a_k\}$ alakú halmazok által generált legszűkebb σ-algebra. Vezessük be a következő T transzformációt az R^∞ téren. $T(x_0,x_1,x_2,\dots)=(x_1,x_2,\dots)$. Adva egy $\xi_0(\omega),\xi_1(\omega),\dots$ stacionárius sorozat egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn definiáljuk a következő $Z\colon\,\Omega\to R^\infty$ leképezést. $Z(\omega)=(\xi_0(\omega),\xi_1(\omega),\dots,)$. Ekkor Z mérhető leképezés az $(\Omega,\mathcal{A})$ térből az $(R^\infty,\mathcal{B}^\infty)$ mérhető térbe. Ez természetes módon indukálja a következő a következő μ mértéket az $(R^\infty,\mathcal{B}^\infty)$ téren. $\mu(A)=P(\{\omega\colon Z(\omega)\in A\})$ minden $A\in\mathcal{B}^\infty$ halmazra. Szavakban elmondva ez azt jelenti, hogy az A halmaz μ mértéke az A halmaz ősképének a P mértéke a Z transzformációra nézve. Ez egy természetes konstrukció a mértékelméletben, és e konstrukcióban a Z transzformáció egy mértéktartó leképezés az $(\Omega,\mathcal{A},P)$ térből az $(R^\infty,\mathcal{B}^\infty,\mu)$ térbe.

Ezzel a konstrukcióval a $(R^\infty, \mathcal{B}^\infty, \mu, T)$ rendszer dinamikus rendszer. Az a tulajdonság ugyanis, hogy a $\xi_1(\omega), \xi_2(\omega), \dots$ sorozat stacionárius abban tükröződik a képtérben, hogy a T transzformáció mértéktartó. (A $\mu(T^{-1}(A)) = \mu(A)$ azonosság könnyen ellenőrizhető az $A = \{(x_0, x_1, \dots)\colon\ x_0 \in B_0, x_1 \in B_1, \dots, x_k \in B_k\}$ alakú halmazokra, ahol $B_0, \dots, B_k$ Borel mérhető halmazok a számegyenesen, és általános itt nem ismertetett mértékelméleti eredményekből innen következik, hogy T mértéktartó. Vezessük be a $\bar\xi_k(x) = x_k$, $k = 0, 1, 2, \dots$, ha $x = (x_0, x_1, \dots)$ valószínűségi változókat az $(R^\infty, \mathcal{B}^\infty, \mu)$ valószínűségi mezőn. Ekkor az $(R^\infty, \mathcal{B}^\infty, \mu, T)$ dinamikus rendszerben definiáltuk valószínűségi változók egy olyan $\bar\xi_0, \bar\xi_1, \dots$ sorozatát, amelynek ugyanolyan az eloszlása, mint a kiinduló $\xi_0, \xi_1, \dots$ stacionárius sorozatnak. Ez lehetővé teszi, hogy stacionárius sorozatok vizsgálatát visszavezessük dinamikus rendszerek vizsgálatára.

Tekintsük például független, egyforma eloszlású $\xi_0, \xi_1, \dots$ valószínűségi változók sorozatát. Definiáljuk a következő $(X, \mathcal{A}, \mu, T)$ dinamikus rendszert. Legen $X = R^\infty$, $\mathcal{A} = \mathcal{B}^\infty$, és definiáljuk a μ mértéket a következő módon. Legyen $\mu\{(x_0, x_1, \dots,)\colon\ x_0 < a_0, x_1 < a_1, \dots, x_k < a_k\} = P(\xi_0 < a_0)P(\xi_1 < a_1) \cdots P(\xi_k < a_k)$ minden k egész és $a_1, \dots, a_k$ valós számokra. Legyen a μ mérték ennek egyértelmű kiterjesztése az $(R^\infty, \mathcal{B}^\infty)$ térre. Definiáljuk a T shift transzformációt úgy, mint korábban, tehát legyen $T((x_0, x_1, \dots)) = (x_1, x_2, \dots)$. Végül legyen $\bar\xi_k(x) = x_k$, $k = 0, 1, 2, \dots$, ha $x = (x_0, x_1, \dots)$. Ekkor $(R^\infty, \mathcal{B}^\infty, \mu, T)$ egy dinamikus rendszer, tehát alkalmazható rá az ergod tétel. Továbbá a $\xi_0, \xi_1, \dots$, és $\bar\xi_0, \bar\xi_1, \dots)$ sorozatoknak ugyanaz az eloszlása, ezért egyszerre teljesítik vagy nem teljesítik a nagy számok erős törvényét. Ahhoz tehát, hogy belássuk, hogy a nagy számok erős törvénye teljesül a $\xi_0, \xi_1, \dots$ sorozatra, ha $E|\xi_0| < \infty$, elég belátni, hogy $\lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{k=0}^{n-1} \bar\xi_k = \mu(\bar\xi_0)$ $\bar\xi_0, \bar\xi_1, \dots$ sorozatra, ha $E|\xi_0| < \infty$. Ahhoz, hogy ezt levezessük az ergod tételből úgy, hogy azt az előbb konstruált $(R^\infty, \mathcal{B}^\infty, \mu, T)$ dinamikus rendszerre alkalmazzuk az $U(x_0, x_1, \dots) = x_0$ függvénnyel elég azt megmutatni, hogy ez a dinamikus rendszer ergodikus. Ez következik az alább megfogalmazott lemmából, illetve a korábban tárgyalt Kolmogorov-féle nulla egy törvényből.

Tétel az invariáns halmazok egy tulajdonságáról. *Tekintsünk egy $(R^\infty, \mathcal{B}^\infty, \mu, T)$ dinamikus rendszert, a korábban definiált R^∞ térrel, $\mathcal{B}^\infty$ σ-algebrával és $T(x_0, x_1, \dots)$ shift transzformációval. Vezessük be a $\bar\xi_k(x) = x_k$, $k = 0, 1, 2, \dots$, ha $x = (x_0, x_1, \dots)$ valószínűségi változókat és $\mathcal{F}_n = \sigma(\bar\xi_n, \bar\xi_{n+1}, \dots)$ valamint $\mathcal{F}_\infty = \bigcap\limits_{n=1}^{\infty} \mathcal{F}_n$ σ-algebrákat. Ha A invariáns halmaz, azaz $T^{-1}(A) = A$, akkor $A \in \mathcal{F}$.*

Következmény. *Ha a tételben bevezetett $\bar\xi_n$ valószínűségi változók függetlenek, akkor a tekintett dinamikus rendszer ergodikus. Ugyanis minden $A \in \mathcal{I}$ halmazra $A \in \mathcal{F}_\infty$, ezért $\mu(A) = 0$ vagy $\mu(A) = 1$ a Kolmogorov-féle nulla egy törvény alapján.*

A következmény alapján a nagy számok törvénye következik az ergod tételből, mert olyan dinamikus rendszert tekintettünk a bizonyításban, amely ergodikus.

A tétel bizonyítása. Legyen egy $A \in \mathcal{A}$ halmaz indikátor függvénye az $I_A(x_0, x_1, \dots)$ függvény. Ekkor a $T^{-n}(A)$ halmaz indikátor függvénye az $I_A(x_n, x_{n+1}, \dots)$ függvény.

Ha A invariáns halmaz akkor $A = T^{-n}(A)$ minden pozitív egész n számra, és indikátor függvénye csak az $x_n, x_{n+1}, \dots$ koordinátáktól függ. Innen következik, hogy $A \in \mathcal{F}_n$. Mivel ez minden n számra igaz, $A \in \mathcal{F}_\infty$.

Kiegészítés: Az ergod tétel bizonyítása.

A nagy számok törvényének egy bizonyítása a Kolmogorov egyenlőtlenségen alapult, ami egy maximum tipusú egyenlőtlenség. Az ergod tétel egy másik maximum ergod tételnek nevezett maximum tipusú egyenlőtlenségen alapul. Ismertetem ezt a kissé szokatlan eredményt, illetve annak legegyszerűbb ismert, Adriano Garciatól származó bizonyítását.

Maximum ergod tétel. *Legyen $(X, \mathcal{A}, \mu, T)$ egy dinamikus rendszer, $U(x)$ az X téren definiált mérhető függvény, amelyre $\int |U(x)|\mu(\,dx) < \infty$. Vezessük be az $S_k(x) = U(x) + U(Tx) + \cdots + U(T^{k-1}x)$, $k = 1, 2, \dots$, és $M_n(x) = \max\{0, S_1(x), \dots, S_n(x)\}$ mennyiségeket. Ekkor*

$$\int_{M_n(x)>0} U(x)\mu(\,dx) \ge 0.$$

Bizonyítás. Mivel $M_n(x) \ge S_k(x)$ minden $1 \le k \le n$ indexre

$$U(x) + M_n(Tx) \ge U(x) + S_k(Tx) = S_{k+1}(x) \quad \text{minden } 1 \le k \le n \text{ számra,}$$

ezért

$$U(x) \ge S_k(x) - M_n(Tx), \quad 2 \le k \le n.$$

Továbbá

$$U(x) \ge S_1(x) - M_n(Tx),$$

mert $S_1(x) = U(x)$, és $M_n(Tx) \ge 0$. Ezen egyenlőtlenségek alapján

$$U(x) \ge \max(S_1(x), \dots, S_n(x)) - M_n(Tx),$$

és mivel $M_n(x) = \max(S_1(x), \dots, S_n(x))$ az $\{x\colon\ M_n(x) > 0\}$ halmazon

$$\int_{M_n(x)>0} U(x)\mu(\,dx) \ge \int_{M_n(x)>0} [M_n(x) - M_n(Tx)]\mu(\,dx). \tag{2}$$

Továbbá

$$\int_{M_n(x)>0} M_n(x)\mu(\,dx) = \int M_n(x)\mu(\,dx),$$

és

$$\int_{M_n(x)>0} M_n(Tx)\mu(\,dx) \le \int M_n(Tx)\mu(\,dx).$$

Ezért a (2) egyenlőtlenségből következik, hogy

$$\int_{M_n(x)>0} U(x)\mu(\,dx) \geq \int M_n(x)\mu(\,dx) - \int M_n(Tx)\mu(\,dx) = 0.$$

A tételt beláttuk.

Az ergod tétel bizonyítása. Vezessük be az $\bar{U}(x) = U(x) - E(U(x)|\mathcal{I})$ függvényt. Ekkor $E(\bar{U}(x)|\mathcal{I}) = 0$, és $\frac{1}{n}\sum\limits_{k=0}^{n-1} \bar{U}(T^k x) = \frac{1}{n}\sum\limits_{k=0}^{n-1} U(T^k x) - E(\bar{U}(x)|\mathcal{I})$. (Azt a tényt használjuk ki, hogy mivel $E(\bar{U}(x)|\mathcal{I})$ $\mathcal{I}$ mérhető függvény, ezért $E(\bar{U}(T^k x)|\mathcal{I}) = E(\bar{U}(x)|\mathcal{I})$.) Innen következik, hogy elég azt belátni, hogy

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} U(T^k x) = 0 \quad \mu \text{ majdnem mindenütt,}$$

ha $E(U(x)|\mathcal{I}) = 0$ a μ mérték szerint majdnem minden $x \in X$ pontban.

Ennek bizonyítása érdekében vezessük be a $\tilde{U}(x) = \limsup\limits_{n} \frac{1}{n}\sum\limits_{k=0}^{n-1} U(T^k x)$ függvényt és valamely rögzített $\varepsilon > 0$ számra a $D = D(\varepsilon) = \{\tilde{U}(x) > \varepsilon\}$ eseményt. Ekkor $\tilde{U}(x)$ $\mathcal{I}$ mérhető függvény, és ezért $D \in \mathcal{I}$. Legyen

$$U^*(x) = [U(x) - \varepsilon] I_D(x),$$

és vezessük be a maximum ergod tétel megfogalmazásában definiált $S_k^*(x)$ és $M_n^*(x)$ függvényeket, ha $U^*(x)$ játssza az $U(x)$ függvény szerepét. Ekkor a maximum ergod tétel alapján

$$\int_{M_n^*(x)>0} U^*(x)\mu(\,dx) \geq 0. \tag{3}$$

Legyen

$$F_n = \{M_n^*(x) > 0\} = \left\{\max_{1\leq k\leq n} S_k^*(x) > 0\right\}.$$

Az F_n halmazok monoton nőnek az

$$F = \left\{\sup_{k\geq 1} S_k^*(x) > 0\right\} = \left\{\sup_{k\geq 1} \frac{S_k^*(x)}{k} > 0\right\} = \left\{\sup_{k\geq 1} \frac{S_k(x)}{k} > \varepsilon\right\} \cap D = D \tag{4}$$

halmazhoz, ahol $S_k(x) = \sum\limits_{j=0}^{k-1} U(T^j x)$, mivel D $\mathcal{I}$ invariáns halmaz, ahonnan $\frac{S_k^*(x)}{k} = I_D(x)\left(\frac{1}{k}\sum\limits_{j=0}^{k-1} U(T^j x) - \varepsilon\right) = I_D(x)\left(\frac{S_k(x)}{k} - \varepsilon\right)$ és $\sup\limits_{k\geq 1} \frac{S_k(x)}{k} \geq \tilde{U}(x)$. Továbbá, mivel $U^*(x) \leq |U(x)| + \varepsilon$, és $E|U(x)| < \infty$ a Lebesgue tétel alapján

$$\int_{F_n} U^*(x)\mu(\,dx) \to \int_F U^*(x)\mu(\,dx).$$

Ezért a (4) majd a (3) formula alapján

$$\int_D U^*(x)\mu(\,dx) = \int_F U^*(x)\mu(\,dx) = \lim_{n\to\infty} \int_{F_n} U^*(x)\mu(\,dx) \geq 0.$$

Innen

$$\begin{aligned} 0 \leq \int_D U^*(x)\mu(\,dx) &= \int_D U(x)\mu(\,dx) - \varepsilon\mu(D) \\ &= \int_D E(U(x)|\mathcal{I})\mu(\,dx) - \varepsilon\mu(D) = -\varepsilon\mu(D), \end{aligned}$$

ezért $\mu(D) = \mu(D(\varepsilon)) = 0$. (A fenti számolásban a feltételes várható értékhez való áttérésnél megint kihasználtuk, hogy $D \in \mathcal{I}$.) Mivel ez a reláció minden $\varepsilon > 0$ számra igaz, ezért a D halmaz definiciója alapján

$$\tilde{U}(x) = \limsup_{n\to\infty} \frac{S_n(x)}{n} \leq 0 \quad \mu \text{ majdnem minden } x \text{ pontra.}$$

Alkalmazva ezt az eredményt a $-U(x)$ függvényre is azt kapjuk, hogy $\lim\limits_{n\to\infty} \frac{S_n(x)}{n} = 0$ μ majdnem minden x pontra. Az ergod tételt beláttuk.