Valószínűségi mértékek és valószínűségi változók közelségének kapcsolata

Ha két valószínűségi változó közel van egymáshoz, akkor ugyanez érvényes eloszlásaikra is. Megfordítva ez természetesen nem igaz, a két valószínűségi változó lehet például független és azonos eloszlású is. Viszont a következő tipusú kérdésekre érdekes, nem triviális válasz adható. Ha adva van két egymáshoz közel lévő valószínűségi mérték μ és ν valamely (X,ρ) metrikus tér Borel σ-algebráján, akkor tudunk-e konstruálni egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőt, rajta két valószínűségi változót ξ-t és η-t úgy, hogy ξ eloszlása μ, η eloszlása ν, és a ξ és η valószínűségi változók közel vannak egymáshoz? Természetesen a közelség fogalmát pontosítani kell.

Egy természetes konstrukció a következő: Legyen $(\Omega,\mathcal{A},P)=(X\times X,\mathcal{A}\times\mathcal{A},P)$, ahol $\times$ direkt szorzatot jelöl, $\mathcal{A}$ a ρ metrika, illetve az általa meghatározott topológia által generált σ-algebra, és P alkalmas valószínűségi mérték az adott téren. Definiáljuk továbbá a $\xi(x_1,x_2)=x_1$, $\eta(x_1,x_2)=x_2$, $(x_1,x_2)\in X\times X$, valószínűségi változókat ezen a téren. Ekkor a feladat a következőképpen fogalmazható át. Konstruáljunk olyan P valószínűségi mértéket az $(X\times X,\mathcal{A}\times\mathcal{A})$ téren, melynek marginális eloszlásai μ és ν, továbbá a P mérték nagyrésze az (x,x), $x\in X$, pontokból álló átló közelébe van koncentrálva. E probléma vizsgálatában nagyon hasznos a kombinatorikában fontos König–Hall tétel, (házassági probléma), illetve ennek egy folytonos változata. Ezeket az eredményeket megfogalmazzuk, illetve egy Appendix-ben bebizonyítjuk.

König–Hall tétel, (Házassági probléma). *Legyen adva n fiú és n lány, melyek között vannak olyanok, akik ismerik egymást. (Az ismeretségek kölcsönösek.) Akkor és csak akkor tudjuk őket úgy párbaállítani (összeházasítani), hogy csak ismerősök kerüljenek párba, ha a lányok bármely csoportja együttesen legalább annyi fiút ismer, mint ennek a csoportnak a létszáma.*

Formálisan megfogalmazva: Tekintsünk egy $Y=\{y_1,\dots,y_n\}$ és $Z=\{z_1,\dots,z_n\}$ halmazból továbbá egy $d\colon Y\times Z\to\{0,1\}$ leképezésből álló páros gráfot. Itt $d(y,z)=1$, $y\in Y$, $z\in Z$ azt jelenti, hogy az y és z pontok össze vannak kötve, $d(y,z)=0$ pedig azt, hogy nincsenek összekötve. Egy $A\subset Y$ esetén definiáljuk a $B(A)\subset Z$ halmazt, mely az A-beli pontokkal összekötött pontokat jelöli, azaz

$$B(A)=\{z\colon z\in Z,\ \text{és létezik olyan } y\in A,\ \text{melyre } d(y,z)=1\}.$$

Akkor és csak akkor létezik ennek a páros gráfnak faktorizációja, azaz az Y és Z halmaz elemeinek olyan $(y_j,z_{\pi(j)})$, $j=1,2,\dots,n$, párbaállítása, melyre $d(y_j,z_{\pi(j)})=1$ minden $j=1,2,\dots,n$-re, és $\pi(j)$, $j=1,\dots,n$, az $\{1,\dots,n\}$ halmaz alkalmas permutációja, ha $|B(A)|\ge|A|$ minden $|A|\subset Z$ halmazra, ahol $|C|$ egy C halmaz számosságát jelöli.

König–Hall tétel egy folytonos változata. *Legyen adva r raktár $u_1,u_2,\dots,u_r$ nagyságú készletekkel és s üzem, $v_1\dots,v_s$ nagyságú igényekkel, $\sum\limits_{j=1}^{r}u_j=\sum\limits_{k=1}^{s}v_k$. Legyenek bizonyos raktárak és üzemek úttal összekötve. Az összes igényt a létező útakon való szállítással akkor és csak akkor elégíthetjük ki, ha üzemek tetszőleges csoportjára, az ezek valamelyikével összekötött raktárak összkapacítása nem kisebb, mint ezen üzemek összigénye.*

Formálisan megfogalmazva: Tekintsünk egy $Y = \{y_1, \dots, y_r\}$ és $Z = \{z_1, \dots, z_s\}$ halmazokból és $d\colon Y \times Z \to \{0,1\}$ leképezésből álló páros gráfot, ahol $d(y,z) = 1$, $y \in Y$, $z \in Z$ azt jelenti, hogy az y és z pontok össze vannak kötve, $d(y,z) = 0$ pedig azt, hogy nincsenek összekötve. Legyen továbbá adva két $u(y)$, $u(y) \ge 0$, $y \in Y$ és $v(z)$, $v(z) \ge 0$, $z \in Z$, a $\sum\limits_{y \in Y} u(y) = \sum\limits_{z \in Z} v(z)$ feltételt teljesítő súlyfüggvény. Minden $A \subset Y$ halmazra definiáljuk a $B(A) \subset Z$ halmazt a

$$B(A) = \{z\colon z \in Z, \text{ és létezik olyan } y \in A, \text{ melyre } d(y,z) = 1\}$$

képlettel. Akkor és csak akkor létezik olyan $w(y,z) \ge 0$ "szállítási függvény", melyre

i.) $\sum\limits_{z\colon d(y,z)=1} w(y,z) = u(y)$ *minden $y \in Y$-re,*
és $\sum\limits_{y\colon d(y,z)=1} w(y,z) = v(z)$ *minden $z \in Z$-re.*

ii.) A $w(y,z) > 0$ egyenlőtlenség csak akkor teljesül, ha $d(y,z) = 1$,

ha minden $A \in Y$-re $\sum\limits_{z \in B(A)} v(z) \ge \sum\limits_{y \in A} u(y)$.

Feladatok:

0.) Lássuk be, hogy a König–Hall tételben, illetve annak folytonos változatában a feltételek szimmetrikusak az Y és Z halmazokra, azaz, ha felcseréljük a feltételekben (és a $B(A)$ halmaz definiciójában) az Y és Z halmazt, akkor a feltételek érvényben maradnak.

1.) Legyen adva két μ és ν valószínűségi mérték egy (X, ρ) szeparábilis metrikus térnek a metrika segítségével generált toplógia által meghatározott $\mathcal{B}$ Borel σ-algebráján. Jelölje $B^\alpha = \{x\colon \rho(x, B) < \alpha\}$ egy $B \subset X$ halmaz α sugarú nyílt környezetét. Tegyük fel, hogy a μ és ν mértékek teljesítik a $\mu(B) \le \nu(B^\alpha) + \beta$ feltételt minden $B \subset X$ zárt halmazra valamilyen $\alpha > 0$ és $\beta > 0$ számokkal. Ekkor minden $\varepsilon > 0$ esetén konstruálható egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és azon olyan ξ és η X értékű valószínűségi változók, melyek eloszlása μ illetve ν, és amelyekre a

$$P(\rho(\xi, \eta) > \alpha + \varepsilon) \le \beta + \varepsilon \tag{a}$$

reláció teljesül. Megfordítva, ha a μ eloszlású ξ és ν eloszlású η valószínűségi változók teljesítik az (a) relációt, akkor $\mu(B) \le \nu(B^{\alpha+\varepsilon}) + \beta + \varepsilon$.

Ha X nemcsak szeparábilis, hanem teljes metrikus tér is, akkor az (a) reláció igaz $\varepsilon = 0$ esetén is. (Az utolsó állítás bizonyításában felhasználhatjuk azt az eredményt, mely szerint szeparábilis teljes metrikus téren μ_n valószínűségi mértékek egyenletes kompakt sorozatának létezik a mértékek gyenge konvergenciája szerint konvergens részsorozata.)

2.) Legyen adva egy (X, ρ) szeparábilis metrikus tér, azon a tér temészetes topológiája által meghatározott $\mathcal{A}$ Borel σ-algebra. Jelölje $\mathcal{M}$ az $(X, \mathcal{A})$ téren értelmezett valószínűségi mértékek terét, és a valószínűségi mértékpárokon vezessük be a következő $d(\cdot,\cdot)$ függvényt: Ha $\mu \in \mathcal{M}$ és $\nu \in \mathcal{M}$, akkor

$$d(\mu,\nu) = \inf\{\alpha\colon \mu(B) \le \nu(B^\alpha) + \alpha \text{ minden zárt } B \subset X \text{ halmazra}\},$$

ahol B^α jelentése ugyanaz, mint az előző feladatban. Lássuk be, hogy $d(\cdot,\cdot)$ metrika az $\mathcal{M}$ téren, és ez a metrika a valószínűségi mértékek terén értelmezett gyenge konvergenciáját metrizálja, azaz egy $\mu_n \in \mathcal{M}$, $n = 1,2,\dots$, sorozatra, és $\mu \in \mathcal{M}$-re $\mu_n \Rightarrow \mu$ $n \to \infty$ esetén akkor és csak akkor, ha $d(\mu_n,\mu) \to 0$, ahol $\Rightarrow$ a gyenge konvergenciát jelöli. Az $(\mathcal{M}, d)$ tér szeparábilis metrikus tér, és ha (X,ρ) teljes szeparábilis metrikus tér, akkor az $(\mathcal{M}, d)$ tér is az.

Megjegyzés: Az, hogy egy topologikus teret indukáló metrika teljes metrika-e vagy sem, nem topologikusan invariáns, azaz lehetséges, hogy egy teljes és nem teljes metrika ugyanazt a topológiát definiálja. Ezért abból az általános eredményből, hogy teljes metrikus tér valószínűségi mértékeinek terén létezik olyan teljes metrika, mely a gyenge konvergenciát indukálja nem következik, hogy a fenti $d(\cdot,\cdot)$ metrika is teljes metrikus teret indukál.

Be fogjuk látni a következő Állítás A-t:

Állítás A: *Alkalmazuk a 2. feladat jelöléseit. Ha a $\mu_n \in \mathcal{M}$, $n = 1,2,\dots$, valószínűségi mértékek sorozata és egy $\mu \in \mathcal{M}$ valószínűségi mérték teljesíti az $\mu_n \Rightarrow \mu$, ha $n \to \infty$ feltételt, akkor létezik $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és azon ξ_n, $n = 1,2,\dots$, és ξ valószínűségi változók úgy, hogy ξ_n eloszlása μ_n, $n = 1,2,\dots$, ξ eloszlása μ, és $\xi_n \to \xi$ egy valószínűséggel.*

Az Állítás A bizonyítása azon alapul, hogy alkalmas konstrukcióban különböző n indexekre azok az n indextől függő halmazok, amelyekre a ξ_n és ξ valószínűségi változók távolsága viszonylag nagy, átfedhetik egymást. Ezt elérendő, a ξ_n valószínűségi változók együttes eloszlását alkalmasan kell megválasztanunk. Ezért a fenti Állítás A-ban szereplő egy valószínűségi konvergencia nem annyira hasznos, mint azt az első pillanatban gondolnánk. A valószínűségszámítás tartalmas egy valószínűséggel érvényes állításokat tartalmazó tételei ugyanis a valószínűségi változók együttes eloszlásától függnek. Az Állítás A viszont megengedi a valószínűségi változók együttes eloszlásának megváltoztatását.

Az Állítás A bizonyítása teljes szeparábilis terekben egyszerűbb, és ebben az esetben a valószínűségi mezőt is nagyon speciálisan lehet megválasztani. Annak érdekében, hogy Állítás A-t ebben az esetben belássuk érdemes először a következő feladatot megoldani.

3.) Legyen (X,ρ) teljes szeparábilis metrikus tér, μ valószínűségi mérték ezen tér Borel σ-algebráján. Tekintsük azt a speciális $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt, melyre $\Omega = [0,1]$, $\mathcal{A}$ a Borel σ-algebra a $[0,1]$ intervallumon, P a Lebesgue mérték a $[0,1]$ intervallum Borel σ-algebráján. Konstruálható ezen a valószínűségi mezőn olyan az értékeit az X térben felvevő ξ valószínűségi változó, melynek eloszlása μ.

4.) Bizonyítsuk be Állítás A-t,ha (X,ρ) szeparábilis, teljes metrikus tér. Ebben az esetben választhatjuk a $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt, mint $\Omega = [0,1]$, $\mathcal{A}$ a Borel

σ-algebra a $[0,1]$ intervallumon, és P a Lebesgue mérték a $[0,1]$ intervallum Borel σ-algebráján.

5.) Az Állítás A (egy alkalmasan nagy $(M, \mathcal{A}, P)$ valószínűségi mezőn) érvényes tetszőleges szeparábilis, nem feltétlenül teljes (X, ρ) metrikus tér esetében.

6.) Legyenek ξ_n, $n = 1, 2, \dots$, és ξ (X, ρ) értékű valószínűségi változók, ahol (X, ρ) szeparábilis metrikus tér. Tegyük fel, hogy $\xi_n \Rightarrow \xi$, ahol $\Rightarrow$ sztochasztikus konvergenciát jelöl. Ekkor a ξ_n valószínűségi változók μ_n és a ξ valószínűségi változók μ eloszlására $\mu_n \Rightarrow \mu$, ahol $\Rightarrow$ valószínűségi mértékek gyenge konvergenciáját jelöli.

Láttuk (1. feladat), hogy az a kérdés, hogy egy (X, ρ) téren definiált μ és ν valószínűségi mértékekhez, hogyan lehet olyan μ és ν eloszlású ξ és η valószínűségi változókat úgy, hogy azok közel vannak egymáshoz a következő némileg informálisan megfogalmazható “szállítási problémához” vezet: Hogyan lehet egy μ tömegeloszlású pontrendszer pontjait viszonylag kevés mozgatással egy ν tömegeloszlású pontrendszerbe átvinni? Ha az (X, ρ) tér a számegyenes a szokásos metrikával, akkor e téren levő egyszerű rendezés miatt ez a “szállítási probléma” lényegesen egyszerűsődik. Ekkor ugyanis bizonyos természetes kiértékelések esetén (lásd a későbbi 9. feladatot) érdemes kizárni a következő lehetőséget: Léteznek olyan $x_1 < x_2$ és $x_3 < x_4$ számpárok, melyekre az x_1 pontot az x_4 és az x_2 pontot az x_3 pontba szállítjuk. Ekkor ugyanis az $x_1 \to x_3$ és $x_2 \to x_4$ szállítás gazdaságosabb. Ezt a lehetőséget zárja ki az alább ismertetendő kvantilis transzormációnak nevezett konstrukció, amelyiknek belátjuk egy optimalitási tulajdonságát.

A kvantilis transzformáció definiciójához szükségünk lesz a következő matematikai statisztikában is gyakran használt tény felidézésére. Ha ξ F eloszlású valószínűségi változó a számegyenesen, akkor alkalmas feltételek esetén $\eta = F(\xi)$ egyenletes eloszlású valószínűségi változó a $[0,1]$ intervallumon. Megfordítva, ha η egyenletes eloszlású valószínűségi változó a $[0,1]$ intervallumon, akkor $\xi = F^{-1}(\eta)$, ahol $F^{-1}(x)$ az $F(x)$ eloszlás inverze, F eloszlású valószínűségi változó. A következő feladat ennek az állításnak pontosítása, enyhe általánosítása.

7.) Legyen ξ egy $F(x) = P(\xi < x)$ eloszlású és η egy a $[0,1]$ intervallumban egyenletes eloszlású valószínűségi változó. Definiáljuk az $F(x)$ (nem feltétlenül szigorúan) monoton függvény általánosított inverzét az $F^{-1}(x) = \sup\{u\colon F(u) < x\}$ képlettel. Ekkor $\bar{\xi} = F^{-1}(\eta)$ F eloszlású valószínűségi változó. Megfordítva, legyen ε a $[0,1]$ intervallumban egyenletes eloszlású és a ξ valószínűségi változótól független valószínűségi változó. Ekkor $\bar{\eta} = \tilde{F}(\xi, \varepsilon) = F(\xi) + \varepsilon[F(\xi+0) - F(\xi)]$, ahol $F(x+0) = \lim\limits_{h>0, h\to 0} F(x+h)$, a $[0,1]$ intervallumban egyenletes eloszlású valószínűségi változó.

8.) Legyen adva két F és G eloszlásfüggvény. Ha ζ egyenletes eloszlású valószínűségi változó, akkor a $\bar{\xi} = F^{-1}(\zeta)$ és $\bar{\eta} = G^{-1}(\zeta)$ valószínűségi változók, ahol F^{-1} és G^{-1} függvények úgy vannak definiálva mint az előző feladatban, F illetve G eloszlásúak. Ha $\bar{\bar{\xi}}$ F eloszlású, és ε a $\bar{\bar{\xi}}$ valószínűségi változótól független a $[0,1]$ intervallumban egyenletes eloszlású valószínűségi változó, akkor $\bar{\bar{\eta}} = G^{-1}(\tilde{F}(\bar{\bar{\xi}}, \varepsilon))$ valószínűségi változó G eloszlású valószínűségi változó. A $(\bar{\xi}, \bar{\eta})$ és $(\bar{\bar{\xi}}, \bar{\bar{\eta}})$ véletlen vektorok eloszlása megegyezik.

A valószínűségszámításban gyakran felmerül a következő probléma. Adva van két μ és ν valószínűségi mérték a számegyenesen, és olyan (ξ, η) véletlen vektort akarunk definiálni a számegyenesen, melyre ξ eloszlása μ, η eloszlása pedig ν eloszlású. Valójában a valószínűségi problémák vizsgálatában elég megadni a (ξ, η) vektor eloszlását. Két konstrukció, mely ugyanolyan eloszlású vektort ad meg, alkalmazások szempontjából ekvivalens. Ezért az előző feladatban konstruált $(\bar{\xi}, \bar{\eta})$ és $(\bar{\bar{\xi}}, \bar{\bar{\eta}})$ véletlen vektorok konstrukcióját egyaránt kvantilis transzformációnak nevezzük. A következő feladatban a kvantilis transzformáció egy optimum tulajdonságát fogalmazzuk meg.

9.) Legyen ξ és η két valószínűségi változó $F(x)$, illetve $G(x)$ eloszlásfüggvényekkel, melyekre $E|\xi| < \infty$, $E|\eta| < \infty$, és $\Phi(x)$ egy konvex függvény a számegyenesen. Ekkor $E\Phi(\xi - \eta) \ge \int_0^1 \Phi\left(F^{-1}(x) - G^{-1}(x)\right) dx > -\infty$, ahol $F^{-1}(x)$ és $G^{-1}(x)$ megegyezik a 6. feladatban definiált inverzzel. Abban az esetben, ha a (ξ, η) vektort kvantilis transzformációval konstruáljuk, akkor a fenti egyenlőtlenség két oldala egyenlő.

Végül néhány az előzőektől kissé eltérő témájú feladatot fogalmazunk meg, melyek bizonyos vizsgálatokban hasznosak. Ezek bizonyítása használ néhány nem-triviális mértékelméleti tényt, például a feltételes eloszlás létezését vagy a Banach felbontási tételt, illetve a Radon–Nikodym tételt, melynek ez utóbbi tétel egyszerű következménye.

10.) Legyen adva három szeparábilis, teljes metrikus tér, (X_i, ρ_i), $i = 1, 2, 3$, és jelölje $\mathcal{A}_i$, $i = 1, 2, 3$, az ezen terek topológiája által indukált σ-algebrát. Legyen μ valószínűségi mérték az $(X_1 \times X_2, \mathcal{A}_1 \times \mathcal{A}_2)$ és ν valószínűségi mérték az $(X_2 \times X_3, \mathcal{A}_2 \times \mathcal{A}_3)$ téren. Ekkor létezik olyan P valószínűségi mérték az $(X_1 \times X_2 \times X_3, \mathcal{A}_1 \times \mathcal{A}_2 \times A_3)$ téren, melynek vetülete az $X_1 \times X_2$ térre a μ, az $X_2 \times X_3$ térre pedig a ν mértékkel egyezik meg.

Megjegyezzük, hogy amennyiben egy (X, ρ) tér teljes szeparábilis metrikus tér, akkor ezek végtelen direkt szorzata $X \times X \times \cdots$ szintén tekinthető mint, egy teljes szeparábilis teljes metrikus tér alkalmas $\bar{\rho}$ metrikával. Valóban, feltehetjük, hogy a ρ mérték olyan, hogy minden $x \in X$ és $\bar{x} \in X$ esetén, $\rho(x, \bar{x}) \le 1$. Ezt elérhetjük például bevezetve a $\rho'(x, \bar{x}) = \min(\rho(x, \bar{x}), 1)$ új metrikát az (X, ρ) téren, ha ez szükséges. Definiáljuk a

$$\bar{\rho}\left((x_1, x_2, \dots), (\bar{x}_1, \bar{x}_2, \dots)\right) = \sum_{k=1}^{\infty} \frac{1}{2^k} \rho(x_k, \bar{x}_k)$$

metrikát az $X \times X \times \cdots$ szorzattéren. Az $(X \times X \times \cdots, \bar{\rho})$ tér szeparábilis teljes metrikus tér ezzel a metrikával.

A 10. feladat eredményéből következik, hogy ha adva van két μ és ν valószínűségi mérték valamely (X, ρ) teljes szeparábilis metrikus téren és olyan μ eloszlású ξ és ν eloszlású η valószínűségi változókat akarunk konstruálni, melyek valamilyen (természetes) értelemben közel vannak egymáshoz, akkor ezt elérhetjük úgy, hogy olyan (ξ, ζ) és $(\eta, \bar{\zeta})$ az (X, ρ) téren értelmezett valószínűségi változókból álló párokat konstruálunk, melyek elemei közel vannak egymáshoz, ezenkívül ξ μ η pedig ν eloszlású, és a ζ illetve $\bar{\zeta}$ valószínűségi változók azonos eloszlásúak. Továbbá, a 10. feladat után tett megjegyzés alapján ezt az érvelést lehet alkalmazni akkor is ha előírt eloszlású $\xi_1, \xi_2, \dots$ sorozatot előírt eloszlású $\eta_1, \eta_2, \dots$ sorozattal akarjuk approximálni.

Az előbbi paragrafus állításait lehet némileg élesíteni. Ha adva van egy μ eloszlású ξ valószínűségi változó egy elég bő $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn továbbá egy ν valószínűségi mérték az $X \times X$ szorzattéren, ahol (X, ρ) teljes szeparábilis metrikus tér, továbbá a ν valószínűségi eloszlás vetülete μ akkor konstruálhatunk olyan η valószínűségi változót ezen az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyre teljesül az, hogy a (ξ, η) véletlen vektor eloszlása ν. Továbbá hasonló állítás érvényes, ha egy $\xi_1, \xi_2, \dots$ sorozatot olyan $\eta_1, \eta_2, \dots$ sorozattal akarunk párosítani egy elég bő valószínűségi mezőn melyek együttes eloszlása előírt. Ez utóbbi megjegyzés azt jelenti, hogy ha egy valószínűségi változónak vagy valószínűségi változók sorozatának "jó approximációját" akarjuk elérni valamely másik valószínűségi változóval vagy valószínűségi változók sorozatával, akkor elég megadni az eredeti és társított valószínűségi változók vagy sorozatok együttes eloszlását. Nem jelent valódi megszorítást, ha az approximálandó valószínűségi változó vagy sorozat rögzítve van.

A fenti állítást a következő 11. feladatban fogjuk belátni. Megjegyezzük, hogy egy valószínűségi mező "elég bő" a fenti értelemben, ha létezik benne egy a ξ valószínűségi vátozótól vagy a $\xi_1, \xi_2, \dots$ sorozattól független a $[0,1]$ intervallumban egyenletes eloszlású valószínűségi változó.

11.) Legyen adva egy ν valószínűségi mérték egy $X \times X$ szorzattéren, ahol (X, ρ) szeparábilis teljes metrikus tér. Legyen μ a ν mérték vetülete az $X \times X$ tér első koordinátájára. Ha ξ μ eloszlású valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi téren, és létezik egy χ a $[0,1]$ intervallumban egyenletes eloszlású, és a ξ-től független valószínűségi változó az $(\Omega, \mathcal{A}, P)$ téren, akkor konstruálható ezen a téren olyan η valószínűségi változó. melyre a (ξ, η) pár eloszlása ν.

12.) Ha ξ és η két valószínűségi változó egy $(X, \mathcal{A})$ mérhető téren, ξ eloszlása μ, η eloszlása ν, akkor $P(\xi \neq \eta) \geq \mathrm{Var}\,(\mu, \nu)$, ahol $\mathrm{Var}\,(\mu, \nu) = \sup\limits_{A \in \mathcal{A}} |\mu(A) - \nu(A)|$, a μ és ν mérték variációs távolsága. Tetszőleges μ-re és ν-re léteznek olyan μ és ν eloszlású ξ és η valószínűségi változók, melyekre a fenti egyenlőtlenség két oldala egyenlő.

Az 1. feladat állítását, illetve azt, hogy adott eloszlással rendelkező mértékek jó illesztése a kombinatorikában jól ismert "szállítási problémák" megoldásán alapul R. M. Dudley dolgozataiból tanultam. A 2. feladatban szereplő metrikát Prohorov metrikának hívják az irodalomban. A 4. (illetve az ennek alapjául szolgáló 3. feladat) eredménye Szkorohodtól, az 5. feladat megoldását szolgáló konstrukció pedig Dudleytól származik. A 4. feladat eredményének az 5. feladatban megfogalmazott általánosítása nem triviális. Megpróbálhatnánk az állítást visszavezetni a teljes metrikus terekről szóló állításra úgy, hogy a szeparábilis metrikus teret beágyazzuk egy teljes szeparábilis metrikus térbe. Ez mindig lehetséges, de előfordulhat, hogy a beágyazott tér a nagyobb tér egy nem mérhető részhalmaza. Ezért ilyen módon nem kapunk egyszerű bizonyítást.

Az ismertetett kvantilis transzformáció illetve annak hasznossága jól ismert az irodalomban, és ennek a módszernek a kidolgozását nehéz névhez kötni.

Megoldások

0.) Egy $B \subset Z$ halmazra definiáljuk az $A(B) = \{y \colon y \in Y,\ d(y,z) = 1\}$ halmazt. Azt kell belátnunk, hogy a König–Hall tétel felételeinek teljesülése esetén $|A(B)| \geq |B|$. Ez ekvivalens az $|Y \setminus A(B)| \leq |Z \setminus B|$ állítással. A feltételből következik, hogy $|B(Y \setminus A(B))| \geq |Y \setminus A(B)|$. Ezért elég megmutatni, hogy $B(Y \setminus A(B)) \subset Z \setminus B$. Ez igaz, mert ha $y \in B(Y \setminus A(B))$, azaz létezik olyan $z \notin A(B)$, melyre $d(y,z) = 1$, akkor az $A(B)$ halmaz definiciója alapján $y \notin B$, azaz $B(Y \setminus A(B)) \subset Z \setminus B$.

Az állítás bizonyítása a König–Hall tétel folytonos változata esetén hasonló. Ebben az esetben a $\sum\limits_{y\in Y} u(y) = \sum\limits_{z\in Z} v(z)$ reláció felhasználásával a bizonyítandó állítás a $\sum\limits_{y\in Y\setminus A(B)} u(y) \leq \sum\limits_{z\in Z\setminus B} v(z)$ állításra vezethető vissza, és ez a $B(Y \setminus A(B)) \subset Z \setminus B$ relációból következik.

1.) Definiáljuk az $(\Omega, \mathcal{A}, P)$ teret, ahol $\Omega = X \times X$, $\mathcal{A}$ az $X \times X$ tér topológiája által definiált σ-algebra, P alkalmasan definiált valószínűségi mérték az $(\Omega, \mathcal{A})$ téren. Legyen $\xi(x_1, x_2) = x_1$ és $\eta(x_1, x_2) = x_2$. A feladat állítását bebizonyítjuk, ha konstruálunk olyan P valószínűségi mértéket az $(\Omega, \mathcal{A})$ téren, melyre

a.) $P(A \times X) = \mu(A)$, $P(X \times A) = \nu(A)$ minden mérhető $A \subset X$ halmazra.

b.) $P\left(\{(x_1, x_2) \colon \rho(x_1, x_2) > \alpha + \varepsilon\}\right) \leq \beta + \varepsilon$.

Ilyen konstrukciót a König–Hall tétel folytonos változatának segítségével adunk meg.

Definiálunk egy páros gráfot alkalmas súlyfüggvényekkel. Ennek érdekében bevezetünk néhány fogalmat. Jelölje $G(x, \alpha)$ az x középpontú α sugarú gömböt. Legyen $x_1, x_2, \dots,$ egy mindenütt sűrű sorozat az X téren, és rögzítsünk egy $\varepsilon > 0$ számot. Mivel $\bigcup\limits_{n=1}^{\infty} G\left(x, \frac{\varepsilon}{5}\right) = X$ létezik olyan $N = N(\varepsilon)$ szám, melyre a $W = \bigcup\limits_{n=1}^{N(\varepsilon)} G\left(x, \frac{\varepsilon}{5}\right)$ halmazra $\mu(W) > 1 - \frac{\varepsilon}{2}$ és $\nu(W) > 1 - \frac{\varepsilon}{2}$. Definiáljuk a $V_k = G\left(x_k, \frac{\varepsilon}{5}\right) \setminus \bigcup\limits_{j=1}^{k-1} G\left(x_j, \frac{\varepsilon}{5}\right)$, $k = 1, \dots, N$ és $V_{N+1} = X \setminus W_N$ halmazokat. Ekkor V_k, $k = 1, \dots, N+1$ az X tér particiója, $d(V_k) \leq \frac{\varepsilon}{5}$, ha $1 \leq k \leq N$, ahol $d(A)$ egy $A \subset X$ halmaz átmérőjét jelöli. Továbbá $\mu(V_{N+1}) < \frac{\varepsilon}{2}$ és $\nu(V_{N+1}) < \frac{\varepsilon}{2}$. Nevezzük az x_k pontot a V_k halmaz középpontjának, $1 \leq k \leq N$.

Definiáljuk a következő páros gráfot: $Y = \{y_1, \dots, y_{N+1}\} = \{V_1, \dots, V_{N+1}\}$, $Z = \{z_1, \dots, z_{N+1}\} = \{V_1, \dots, V_{N+1}\}$, $d(y_j, z_k) = 1$, ha $\rho(x_j, x_k) \leq \alpha + \frac{\varepsilon}{2}$, $1 \leq j, k \leq N$, illetve $d(y_{N+1}, z_k) = 1$, $d(y_j, z_{N+1}) = 1$ minden $1 \leq j, k \leq N+1$ esetén. Minden egyéb esetben $d(y,z) = 0$. Vezessük be továbbá a következő súlyfüggvényeket: $u(y_j) = \mu(V_j)$, $v(z_j) = \nu(V_j)$, $j = 1, \dots, N$, $u(y_{N+1}) = \mu(V_{n+1}) + \beta$, $v(z_{N+1}) = \nu(V_{n+1}) + \beta$. Azt állítjuk, hogy ez a rendszer teljesíti a König–Hall tétel folytonos verziójának feltételeit.

A kívánt egyenlőtlenség nyílván teljesül olyan $A \subset X$ halmazra, melyre $y_{N+1} \in A$. Ekkor ugyanis $B(A) = Z$, mert az y_{N+1} a Z halmaz minden pontjával össze van kötve. Ha $y_{N+1} \notin A$ legyen $D_1 = \bigcup_{V_j \in A} V_j$ és $D_2 = \bigcup_{V_j \in B(A)} V_j$. Ekkor $\sum_{y \in A} u(y) = \mu(D_1)$ és $\sum_{z \in B(A)} v(z) = \nu(D_2) + \beta$, mivel $y_{N+1} \notin A$ és $z_{N+1} \in B(A)$. Ezért az állítás feltételei miatt elég belátni, hogy $\bar{D}_1^\alpha \subset D_2$, ahol $\bar{D}_1$ a D_1 halmaz lezártját jelöli. Viszont, ha $x \in \bar{D}_1$, akkor létezik olyan $y_j = V_j \in A$ halmaz, melynek x_j középpontjára $\rho(x, x_j) \le \frac{\varepsilon}{5}$, így $G(x, \alpha) \subset G\left(x_j, \alpha + \frac{\varepsilon}{5}\right)$. Másrészt $G\left(x_j, \alpha + \frac{\varepsilon}{5}\right) \subset D_2$. Ellenkező esetben ugyanis létezne olyan $v \in X$ pont, melyre $\rho(x_j, v) < \alpha + \frac{\varepsilon}{5}$, és $v \in V_k$ egy olyan V_k, $1 \le k \le N$ halmazra, mely nincs összekötve a V_j halmazzal, azaz $\rho(x_j, x_k) \ge \alpha + \frac{\varepsilon}{2}$. Ez azonban nem lehetséges, mert $d(V_k) \le \frac{\varepsilon}{5}$, így $\rho(x_j, x_k) \le \rho(x_j, v) + \frac{\varepsilon}{5}$. Beláttuk, hogy a König–Hall tétel folytonos változata alkalmazható erre a rendszerre.

Legyen $\bar{w}(y, z)$ egy a König–Hall tétel folytonos változatát a fenti rendszerben kielégítő függvény, és definiáljuk a $w_1(y_j, z_k)$ függvényt a következő módon:

$$\begin{aligned} w_1(y_j, z_k) &= \bar{w}(y_j, z_k) \quad \text{ha } 1 \le j, k \le N, \\ \text{és } w_1(y_j, z_k) &= 0, \quad \text{ha } j = N+1 \text{ vagy } k = N+1. \end{aligned}$$

Ez a $w_1(y, z)$ függvény teljesíti a következő tulajdonságokat:

i.) $w_1(y_j, z_k) \ge 0$, és $w_1(y_j, z_k) > 0$ csak akkor, ha $1 \le j, k \le N$ és $\rho(x_j, x_k) < \alpha + \frac{\varepsilon}{2}$

ii.) $\sum_{z \in Z} w_1(y_j, z) \le u(y_j) = \mu(V_j)$, $\sum_{y \in Y} w_1(y, z_k) \le v(z_k) = \nu(V_k)$.

iii.) $\sum_{y \in Y, z \in Z} w_1(y, z) \le 1 - \beta - \varepsilon$ mert $\sum_{y \in Y, z \in Z} w_1(y, z) \le \sum_{y \in Y, z \in Z} w(y, z) - u(y_{N+1}) - v(z_{N+1}) \le 1 + \beta - 2(\beta + \frac{\varepsilon}{2})$.

A $w_1(y, z)$ függvény tulajdonságaiból következik, hogy létezik olyan $w_2(y_j, z_k) \ge 0$, $1 \le j, k \le N+1$, függvény, melyre a $w(y, z) = w_1(y, z) + w_2(y, z)$ függvény teljesíti a

$$\sum_{z \in Z} w(y_j, z) = u(y_j) = \mu(V_j), \quad \sum_{y \in Y} w(y, z_k) = v(z_k) = \nu(V_k)$$

tulajdonságokat. Ennek a $w(y, z)$ függvénynek a segítségével definiálunk egy olyan P valószínűségi mértéket az $X \times X$ téren, mely teljesíti az a.) és b.) tulajdonságokat. Legyen

$$P(C \times D) = \frac{\mu(C)}{\mu(W_j)} \frac{\nu(C)}{\nu(W_k)} w(y_j, z_k) \quad \text{ha } C \subset W_j,\ D \subset W_k$$

valamilyen $1 \le j, k \le N+1$ indexekkel. Az általános esetben pedig legyen

$$P(C \times D) = \sum_{j=1}^{N+1} \sum_{k=1}^{N+1} P\left((C \cap V_j) \times (D \cap V_k)\right).$$

Ezután a P mértéket egyértelműen kiterjeszthetjük ezekről a téglalap halmazokról. Ez a P mérték teljesíti mind az a.) mind a b.) tulajdonságot. Ugyanis tetszőleges $A \subset V_j$, $1 \le j \le N+1$ esetén

$$P(A \times X) = \sum_{k=1}^{N+1} P(A \times V_j) = \frac{\mu(A)}{\mu(V_j)} \sum_{k=1}^{N+1} w(x_j, y_k) = \mu(A),$$

és innen következik, hogy $P(X \times A) = \nu(A)$ minden mérhető $A \subset X$ halmazra. Az a.) állítás második része hasonlóan látható. Másrészt,

$$\begin{aligned} P((x_1,x_2)\colon \rho(x_1,x_2) \le \alpha+\varepsilon) &\ge \sum_{(j,k)\colon\, \rho(x_j,x_k)\le\alpha+\frac{\varepsilon}{2}} P(V_j \times V_k) \\ &= \sum_{(j,k)\colon\, \rho(x_j,x_k)\le\alpha+\frac{\varepsilon}{2}} w(V_j \times V_k) \\ &\ge \sum_{y_j\in Y,\, z_k\in Z} w_1(V_j \times V_k) \ge 1-\beta-\varepsilon, \end{aligned}$$

és ez az állítás ekvivalens a b.) feltétellel.

Megfordítva, ha az (a) tulajdonság teljesül, akkor tetszőleges B (zárt) halmazra $\{\xi \subset A, \eta \notin A^{\alpha+\varepsilon}\} \subset \{(\xi,\eta)\colon \rho(\xi,\eta) > \alpha+\varepsilon\}$, ezért $P(\xi \in A, \eta \notin A^{\alpha+\varepsilon}) \le \beta+\varepsilon$. Mivel $\{\xi \in A\} \subset \{\xi \in A, \eta \notin A^{\alpha+\varepsilon}\} \cup \{\eta \in A^{\alpha+\varepsilon}\}$, innen következik, hogy $\mu(A) \le \beta+\varepsilon+\nu(A^{\alpha+\varepsilon})$, és ezt kellett bizonyítani.

A következő észrevételt tesszük. Ha X teljes metrikus tér és P_n, $n=1,2,\dots$, olyan valószínűségi mértékek sorozata az $X \times X$ téren, melynek marginális eloszlásai két n-től független μ és ν mérték, akkor e mértéksorozatnak létezik a gyenge konvergencia szerint konvergens részsorozata.

Ennek az állításnak a bizonyításához azt érdemes észrevenni, hogy tetszőleges $\varepsilon > 0$ számra létezik olyan $K \subset X$ kompakt halmaz, melyre $\mu(K) \ge 1-\frac{\varepsilon}{2}$ és $\nu(K) \ge 1-\frac{\varepsilon}{2}$. Ezért a $K \times K \subset X \times X$ kompakt halmazra és valószínűségi mértékek olyan P_n sorozatára az $X \times X$ téren melyek marginális eloszlásai a μ és ν mérték $P_n(K \times K) \ge 1-\varepsilon$ minden n-re, ahonnan következik, hogy a P_n mértéksorozat feszes, azaz létezik gyengén konvergens részsorozata.

Legyen P_n, $n=1,2,\dots$, olyan μ és ν marginális eloszlással rendelkező valószínűségi mértékek sorozata az $X \times X$ téren, melyek teljesítik az (a) tulajdonságot $\varepsilon = \frac{1}{n}$ számmal. Legyen P_{n_k}, $k=1,2,\dots$, ennek a sorozatnak gyengén konvergens részsorozata, és legyen P e részsorozat határértéke. A P mérték marginális eloszlásai μ és ν. Azt állítjuk, hogy egy P eloszású (ξ,η) vektor teljesíti az (a) feltételt az $\varepsilon = 0$ számmal is. Ugyanis, az $\{(x_1,x_2)\colon (x_1,x_2) \in X \times X, \rho(x_1,x_2) > \alpha+\varepsilon\}$ alakú halmazok nyíltak. Ezért

$$\beta \ge \limsup_{k\to\infty} P_{n_k}\left(\{(x_1,x_2)\colon \rho(x_1,x_2) > \alpha+\varepsilon\}\right) \ge P\left(\{(x_1,x_2)\colon \rho(x_1,x_2) > \alpha+\varepsilon\}\right)$$

minden $\varepsilon > 0$-ra, ahonnan

$$P\left(\{(x_1,x_2)\colon \rho(x_1,x_2) > \alpha\}\right) = \lim_{\varepsilon\to 0} P\left(\{(x_1,x_2)\colon \rho(x_1,x_2) > \alpha+\varepsilon\}\right) \le \beta.$$

Ezzel az 1. feladat állításait bebizonyítottuk.

2.) Először azt látjuk be, hogy $d(\cdot,\cdot)$ metrika. i.) $d(\mu,\mu) = 0$. Másrészt, belátjuk, hogy $d(\mu,\nu) = 0$ esetén $\mu = \nu$. Valóban, $d(\mu,\nu) = 0$ esetén $\mu(F) \le \nu(F)$ minden zárt $F \subset G$ halmazra, mert $F = \bigcap\limits_{\varepsilon\to 0} F^\varepsilon$, így $\mu(F) \le \liminf\limits_{\varepsilon\to 0}(\nu(F^\varepsilon)+\varepsilon) = \nu(F)$. Belátjuk, hogy valójában $\mu(F) = \nu(F)$ minden zárt halmazra. Ugyanis tudjuk, hogy $\mu(G) \ge \nu(G)$ minden nyílt G halmazra, ezért $\mu(F) = \lim\limits_{\varepsilon\to 0}\mu(F^\varepsilon) \ge \lim\limits_{\varepsilon\to 0}\nu(F^\varepsilon) = \nu(F)$. Innen következik, hogy $\mu(A) = \nu(A)$ minden zárt vagy nyílt A halmazra. Egy mértéket viszont meghatároznak a nyílt halmazokon felvett értékei, ezért $\mu = \nu$. ii.) $d(\mu,\nu) = d(\nu,\mu)$. Legyen $d(\mu,\nu) < \alpha$. Belátjuk, hogy ekkor $d(\nu,\mu) \le \alpha$. Legyen $F \subset X$ tetszőleges zárt halmaz, és definiáljuk az $F' = X \setminus F^\alpha$ halmazt. Állítjuk, hogy $(F')^\alpha \subset X \setminus F$. Valóban, ha $y \in (F')^\alpha$ akkor $d(y, X\setminus F^\alpha) < \alpha$, azaz létezik olyan $z \in X$, melyre $d(z,F) \ge \alpha$ és $d(z,y) < \alpha$. Ebből viszont következik, hogy $y \notin F$, tehát $(F')^\alpha \subset X \setminus F$. Ezt a relációt felhasználva kapjuk, hogy $1-\mu(F^\alpha) = \mu(F') \le \nu((F')^\alpha)+\alpha \le \nu(X\setminus F)+\alpha = 1-\nu(F)+\alpha$, azaz $\nu(F) \le \mu(F^\alpha)+\alpha$. Ezért $d(\nu,\mu) \le \alpha$. Ezzel beláttuk, hogy $d(\nu,\mu) \le d(\mu,\nu)$. Szimmetriaokokból $d(\mu,\nu) = d(\nu,\mu)$. iii.) $d(\mu_1,\mu_3) \le d(\mu_1,\mu_2)+d(\mu_2,\mu_3)$. Ha $d(\mu_1,\mu_2) = \alpha$, $d(\mu_2,\mu_3) = \beta$ akkor minden $\varepsilon > 0$-ra és zárt F halmazra $\mu_1(F) \le \mu_2(F^{\alpha+\varepsilon})+\alpha+\varepsilon$, $\mu_2(F^{\alpha+\varepsilon}) \le \mu_3((F^{\alpha+\varepsilon})^{\beta+\varepsilon})+\beta+\varepsilon$. Mivel $(F^{\alpha+\varepsilon})^{\beta+\varepsilon} \subset F^{\alpha+\beta+2\varepsilon}$ innen kapjuk, hogy $\mu_1(F) \le \mu_3(F^{\alpha+\beta+2\varepsilon})+\alpha+\beta+2\varepsilon$. Mivel ez az állítás érvényes minden F zárt halmazra és $\varepsilon > 0$ számra, innen következik a iii.) reláció.

$d(\mu_n,\mu) \to 0$ akkor és csak akkor, ha $\mu_n \Rightarrow \mu$. Ha $d(\mu_n,\mu) \to 0$, akkor minden zárt F halmazra, és $\varepsilon > 0$ számra $\limsup\limits_{n\to\infty}\mu_n(F) \le \mu(F^\varepsilon)+\varepsilon$. Mivel $F = \bigcap\limits_{\varepsilon\to 0} F^\varepsilon$, $\lim\limits_{\varepsilon\to 0}(\mu(F^\varepsilon)+\varepsilon) = \mu(F)$. Innen, $\limsup\limits_{n\to\infty}\mu_n(F) \le \mu(F)$ és ez az állítás ekvivalens azzal, hogy $\mu_n \Rightarrow \mu$.

A másik irányú állítás bizonyításához először azt mutatjuk meg, hogy minden $\varepsilon > 0$-ra létezik olyan $N = N(\varepsilon)$ egész szám és az X tér $V_1,\dots,V_N,V_{N+1}$ particiója, mely teljesíti az alábbi feltételeket: $\mu(\partial V_k) = 0$, $k = 1,\dots,N+1$, $\bar d(V_k) < \varepsilon^2$, $k = 1,\dots,N$, és $\mu(V_{N+1}) \le \frac{\varepsilon}{2}$, ahol ∂A az A halmaz határát és $\bar d(A)$ az A halmaz átmérőjét jelöli. Valóban, legyen $x_1,x_2,\dots$ egy mindenütt sűrű halmaz az X térben. Mindegyik x_k-re válasszunk egy olyan x_k középpontú δ_k, $\frac{\varepsilon^2}{2} < \delta_k < \varepsilon^2$ sugarú $G(x_k,\delta_k)$ gömböt az X térben, melyre $\mu\left(\partial G(x_k,\delta_k)\right) = 0$. E gömbök uniója lefedi az X teret. Válasszunk egy olyan $N = N(\varepsilon)$ számot, melyre $\mu\left(\bigcup\limits_{k=1}^{N} G(x_k,\delta_k)\right) \le \frac{\varepsilon}{2}$. Legyen $V_{N+1} = X \setminus \left(\bigcup\limits_{k=1}^{N} G(x_k,\delta_k)\right)$ és $V_k = G(x_k,\delta_k) \setminus \left(\bigcup\limits_{j=1}^{k=1} G(x_j,\delta_j)\right)$, $k = 1,\dots,N$. E halmazok teljesítik a kívánt feltételeket.

Minden $F \subset X$ zárt halmazra definiáljuk a $B(F) = \bigcup\limits_{k\colon V_k \cap F \neq \emptyset} V_k$ halmazt. Vegyük észre, hogy $F \subset B(F) \subset F^\varepsilon \cup V_{N+1}$. Továbbá $\lim\limits_{n\to\infty} \sup\limits_{F \text{ zárt halmaz}} |\mu_n(B(F) - \mu(B(F)| = 0$, mert $\lim\limits_{n\to\infty} \mu_n(V_k) = \mu(V_k)$ minden $k = 1, \dots, N+1$-re, véges sok $B(F)$ alakú halmaz van, és mindegyik véges sok V_k halmaz uniója. Ezért létezik olyan $n = n(\varepsilon)$ az F zárt halmaztól független küszöbindex, melyre

$$\mu(F^\varepsilon) - \mu_n(F) \geq \mu(B(F)) - \mu(V_{N+1}) - \mu_n(B(F)) \geq -\varepsilon$$

minden zárt F halmazra, tehát $d(\mu_n, \mu) \leq \varepsilon$, ha $n \geq n_0(\varepsilon)$. Innen következik az állítás.

Az $(\mathcal{M}, d)$ tér szeparábilis. Legyen ugyanis $x_1, x_2, \dots,$ egy mindenütt sűrű halmaz az X téren, legyen $\mathcal{M}_0$ azon diszkrét valószínűségi mértékek halmaza, melyek az $\{x_1, x_2, \dots\}$ halmaz véges részhalmazaira vannak koncentrálva, és minden pont mértéke racionális szám. Ez megszámlálható halmaz, ezért elég belátni, hogy $\mathcal{M}_0$ az $\mathcal{M}$ halmaz mindenütt sűrű részhalmaza. Ezt be lehet látni például az előző érvelés némi módosításával. Ez ugyanis mutatja, hogy tetszőleges $\mu \in \mathcal{M}$ mértékhez és $\varepsilon > 0$ számhoz választhatunk olyan $V_1, \dots, V_N$ particióját az X térnek és olyan $\eta > 0$ számot, melyre igaz, hogy ha egy $\nu \in \mathcal{M}$ mérték teljesíti a $|\mu(V_k) - \nu(V_k)| \leq \eta$ feltételt minden $k = 1, \dots, N$-re akkor ez a ν mérték teljesíti a $d(\mu, \nu) < \varepsilon$ relációt is. Mivel minden $\mu \in \mathcal{M}$ esetén $\mathcal{M}_0$ tartalmaz ilyen ν mértéket ezért $\mathcal{M}_0$ az $\mathcal{M}$ halmaz mindenütt sűrű részhalmaza.

Ha X teljes metrikus tér, akkor (X, d) is az. Elég belátni, hogy ha

$$\lim_{n\to\infty} \sup_{m\colon m \geq n} d(\mu_n, \mu_m) = 0$$

akkor a μ_n, $n = 1, 2, \dots$ sorozat, relative kompakt, azaz létezik gyengén konvergens részsorozata. Ehhez elég belátni, hogy tetszőleges $\varepsilon > 0$ számhoz létezik olyan $K \subset X$ kompakt halmaz, melyre $\mu_n(K) \geq 1 - \varepsilon$ minden n indexre.

Ezt a bizonyítandó állítást lehet gyengíteni a következő módon. Elegendő megmutatni azt, hogy tetszőleges $\varepsilon > 0$-ra létezik olyan kompakt $K = K(\varepsilon)$ halmaz, melynek K^ε ε sugarú környezete teljesíti a $\mu_n(K^\varepsilon) \geq 1 - \varepsilon$ egyenlőtlenséget minden $n = 1, 2, \dots$-ra. Tekintsünk ugyanis ilyen $K(\varepsilon 2^{-m})$ halmazokat, minden $m = 1, 2, \dots$ számra, és definiáljuk a $K = \bigcap\limits_{m=1}^{\infty} (K(\varepsilon 2^{-m})^{\varepsilon 2^{-m}}$ halmazt. Ekkor $\mu_n(K) \geq 1 - \sum\limits_{m=1}^{\infty} \varepsilon 2^{-m} = 1 - \varepsilon$. Ezért elég belátni, hogy a fenti K halmaz relative kompakt, (azaz lezártja kompakt). Ennek érdekében idézzük fel azt az eredményt, mely szerint egy teljes szeparábilis tér A részhalmaza, akkor és csak akkor relative kompakt, ha minden $\delta > 0$-ra létezik az A halmaznak véges δ hálója, azaz olyan véges halmaz, melyre tetszőleges $x \in A$ pont távolsága e véges halmaz valamelyik pontjától kisebb mint δ. Ez a feltétel teljesül a fenti K halmazra, mert minden $\delta > 0$-ra létezik olyan m szám, melyre $\delta > \varepsilon 2^{-m}$, ilyen m-re az $K(\varepsilon 2^{-m})^{\varepsilon 2^{-m}}$ halmazban van véges δ háló, és ez véges δ háló a K halmazban is.

Ezt a gyengített feltételt beláthatjuk a következő módon. Egy rögzített $\varepsilon > 0$ számhoz válasszunk olyan $n_0 = n_0(\varepsilon)$ indexet, melyre $d(\mu_{n_0}, \mu_n) \le \frac{\varepsilon}{4}$ minden $n \ge n_0$ számra, és legyen $K_0 \subset X$ olyan kompakt halmaz, melyre $\mu_{n_0}(K_0) \ge 1 - \frac{\varepsilon}{4}$. Ekkor $\mu_n(K_0^{\varepsilon/4}) \ge \mu_{n_0}(K_0) - \frac{\varepsilon}{4} \ge 1 - \frac{\varepsilon}{2}$ minden $n \ge n_0$ számra. Válasszunk továbbá olyan K_1 kompakt halmazt, melyre $\mu_n(K_1) \ge 1 - \frac{\varepsilon}{2}$ minden $n \le n_0$-ra. Ekkor a $K = K_1 \cup K_0$ halmaz kompakt, és $\mu_n(K^\varepsilon) \ge 1 - \varepsilon$ minden $n = 1, 2, \dots$ számra. Ugyanis $\mu_n(K^\varepsilon) \ge \mu_n(K_0^{\varepsilon/4} \ge 1 - \varepsilon$, ha $n \ge n_0$, és $\mu_n(K^\varepsilon) \ge \mu_n(K_1) \ge 1 - \varepsilon$, ha $n \le n_0$. A második feladat állításait beláttuk.

3.) Létezik az X térnek olyan $\mathcal{X}_1 = \{A_1, A_2, \dots\}$ particiója, melyre $\bar{d}(A_j) \le 1$, és $\mu(\partial A_j) = 0$ minden $j = 1, 2, \dots$ indexre. Itt $\bar{d}(A)$ az A halmaz átmérőjét, ∂A pedig az A halmaz határát jelöli. (Ennek az állításnak a bizonyítása az előző feladatban szereplő $\mu_n \Rightarrow \mu$ akkor $d(\mu_1, \mu_2) \to 0$ állítás bizonyításának elején megadott érvelésből következik.) A partició egyes elemeit továbbosztva meg lehet adni az X tér olyan egyre finomodó $\mathcal{X}_1 \supset \mathcal{X}_2 \supset \cdots \mathcal{X}_k \supset \cdots$ melyek $A_{j_1,\dots,j_k} \in \mathcal{X}_k$ ($A_{j_1,\dots,j_k} \subset A_{j_1,\dots,j_{k-1}}$) elemeire $\bar{d}\,(A_{j_1,\dots,j_k}) \le 2^{-k}$ és $\mu\,(\partial(A_{j_1,\dots,j_k}) = 0$. Definiáljuk a $[0,1)$ intervallumnak (a Lebesgue mértékkel ellátva) hasonló intervallumokból álló egymásba skatulyázott $\mathcal{Y}_1 \supset \mathcal{Y}_2 \supset \cdots \mathcal{Y}_k \supset \cdots$ particióit a következő módon. Legyen $\mathcal{Y}_1 = \{B_1, \dots, B_k, \dots\}$, $B_k = [b_{k-1}, b_k)$, $b_k = \sum\limits_{j=1}^{k} \mu(A_j)$, $k = 1, 2, \dots$, és ha az $\mathcal{Y}_k$ partició $B_{j_1,\dots,j_k} = [b_{j_1,\dots,j_{k-1}}, b_{j_1,\dots,j_k})$, $b_{j_1,\dots,j_k} - b_{j_1,\dots,j_{k-1}} = \mu(A_{j_1,\dots,j_k})$ halmazait már definiáltuk, akkor az $\mathcal{Y}_{k+1}$ particióját definiáljuk úgy, hogy a $B_{j_1,\dots,j_k}$ intervallumot felosztjuk egymáshoz csatlakozó $B_{j_1,\dots,j_k,s} = [b_{j_1,\dots,j_k,s-1}, b_{j_1,\dots,j_k,s})$, $\mu\,(A_{j_1,\dots,j_k,s})$ hosszúságú intervallumokra, és ez definiálja az $\mathcal{Y}_{K+1}$ particióját. Válasszunk ki minden $k \ge 1$ számhoz, és $A_{j_1,\dots,j_k}$ halmazhoz egy $x_{j_1,\dots,j_k} \in A_{j_1,\dots,j_k}$ pontot. Definiáljuk a ξ_k, $k = 1, 2, \dots$, X tér értékű valószínűségi változókat a $([0,1), \mathcal{B}, \lambda)$ valószínűségi mezőn a $\xi_k(y) = x_{j_1,\dots,j_k}$, ha $y \in B_{j_1,\dots,j_k}$ képlet segítségével. Azt állítjuk, hogy létezik a $\xi(y) = \lim\limits_{k\to\infty} \xi_k(y)$, $y \in [0,1)$ valószínűségi változó, és ξ eloszlása μ.

A fenti limesz valóban létezik, mert minden $y \in [0,1)$ ponthoz ki lehet választani egyértelműen olyan egymásba skatulyázott $B_{j_1,\dots,j_k}$ intervallumokat, melyekre $y \in B_{j_1,\dots,j_k}$, és ebből következik, hogy a $\xi_k(y) \in A_{j_1,\dots,j_k}$ pontok Cauchy sorozatot alkotnak.

Ahhoz, hogy belássuk azt, hogy a ξ valószínűségi változó μ eloszlású elég megmutatni azt, hogy $P\,(\xi \in A_{j_1,\dots,j_k}) = \mu\,(A_{j_1,\dots,j_k})$ minden $k = 1, 2, \dots$ számra és $A_{j_1,\dots,j_k}$ halmazra. A ξ valószínűségi változó konstrukciójából következik, hogy $y \in B_{j_1,\dots,j_k}$ esetén $\xi(y) \in \bar{A}_{j_1,\dots,j_k}$, ahol $\bar{A}$ az A halmaz lezártját jelöli. Először azt mutatjuk meg, hogy ebből a tényből és a később bizonyítandó

$$P\,(\xi(y) \in \partial A_{j_1,\dots,j_k}) = 0 \quad \text{minden } k = 1, 2, \dots \text{ számra és } j_1, \dots, j_k \text{ indexre.} \qquad (*)$$

állításból következik az, hogy ξ μ eloszlású. Valóban, innen következik, hogy

$$\begin{aligned} P\,(\xi(y) \in A_{j_1,\dots,j_k}) &\ge P\left(\xi(y) \in \bar{A}_{j_1,\dots,j_k}\right) - P\,(\xi(y) \in \partial A_{j_1,\dots,j_k}) \\ &\ge \lambda\,(B_{j_1,\dots,j_k}) = \mu\,(A_{j_1,\dots,j_k})\,, \end{aligned}$$

és ezeket az egyenlőtlenségeket összegezve kapjuk, hogy

$$1=\sum_{j_1,\dots,j_k} P\left(\xi(y)\in A_{j_1,\dots,j_k}\right)\ge \sum_{j_1,\dots,j_k}\mu\left(A_{j_1,\dots,j_k}\right)=1$$

Innen következik, hogy a fenti egyenlőtlenségek valójában azonosságok, ezért a ξ valószínűségi változó μ eloszlású.

A $(*)$ állítás bizonyításához vegyük észre, hogy mivel $\mu(\partial A_{j_1,\dots,j_k})=0$ minden $\varepsilon>0$ számhoz létezik olyan $\delta=\delta(\varepsilon)>0$ szám, melyre a $\partial A_{j_1,\dots,j_k}$ halmaz δ sugarú $(\partial A_{j_1,\dots,j_k})^\delta$ környezetére $\mu\left((\partial A_{j_1,\dots,j_k})^\delta\right)<\varepsilon$. Válasszunk egy olyan nagy s számot, melyre $\bar d(A_{j'_1,\dots,j'_s})<\frac{\delta}{2}$ minden $A_{j'_1,\dots,j'_s}$ halmazra. Mivel $\rho(\xi_s(y),\xi(y))\le \max\limits_{j_1,\dots,j_k}\bar d(A_{j_1,\dots,j_k})$ ezért a $\xi(y)\in\partial A_{j_1,\dots,j_k}$ esetén $\rho\left(\xi_s(y),\partial A_{j_1,\dots,j_k}\right)<\frac{\delta}{2}$. Innen kapjuk, hogy $\xi(y)\in\partial A_{j_1,\dots,j_k}$ esetén $y\in B_{j'_1,\dots,j'_s}$ olyan $j'_1,\dots,j'_s$ indexekkel, melyre $\rho\left(A_{j'_1,\dots,j'_s},\partial A_{j_1,\dots,j_k}\right)<\frac{\delta}{2}$, ezért $A_{j'_1,\dots,j'_s}\subset(\partial A_{j_1,\dots,j_k})^\delta$. Innen kapjuk, hogy

$$\begin{aligned}\lambda\left\{y\colon \xi(y)\in\partial A_{j_1,\dots,j_k}\right\}&\le \sum_{(j'_1,\dots,j'_s)\colon\ A_{j'_1,\dots,j'_s}\subset(\partial A_{j_1,\dots,j_k})^\delta}\lambda\left(B_{j'_1,\dots,j'_s}\right)\\ &=\sum_{(j'_1,\dots,j'_s)\colon\ A_{j'_1,\dots,j'_s}\subset(\partial A_{j_1,\dots,j_k})^\delta}\mu\left(A_{j'_1,\dots,j'_s}\right)\le\mu\left((\partial A_{j_1,\dots,j_k})^\delta\right)\le\varepsilon.\end{aligned}$$

Mivel ez az egyenlőtlenség minden $\varepsilon>0$-ra igaz, innen következik a $(*)$ reláció és a 3. feladat állítása.

4.) Tekintsük a μ_n, $n=1,2,\dots$, illetve μ eloszlású ξ valószínűségi változóknak a 3. feladat megoldásában megadott konstrukcióját mindegyik mérték konstrukciójában ugyanazt az $\mathcal{X}_1\subset\mathcal{X}_2\subset\cdots$ particiósorozatot használva. Az egyetlen különbség az, hogy jelen esetben azt kell biztosítani, hogy a $\mu(\partial A_{j_1,\dots,j_k})=0$ feltételek mellett a $\mu_n(\partial A_{j_1,\dots,j_k})=0$ feltételek is teljesüljenek minden $n=1,2,\dots$ indexre. Azt állítjuk, hogy a $\mu_n\Rightarrow\mu$ relációból következik, hogy $\xi_n\to\xi$ egy valószínűséggel a fenti konstrukció valószínűségi változóira.

Vegyük észre, hogy $\lim\limits_{n\to\infty}\mu_n(A_{j_1,\dots,j_k})=\mu(A_{j_1,\dots,j_k})$ minden $j_1,\dots,j_k$-ra, mert ezen halmazok határára 0 μ-mértékűek. Innen, illetve a konstrukcióból az is következik, hogy $b_{j_1,\dots,j_k}(n)\to b_{j_1,\dots,j_k}$ $n\to\infty$ esetén, ahol $b_{j_1,\dots,j_k}(n)$ jelöli a ξ_n valószínűségi változók konstrukciójában szereplő $B_{j_1,\dots,j_k}(n)$ intervallum jobboldali végpontját.

Rögzített $\varepsilon>0$ számra válasszunk olyan k számot, melyre $2^{-k}<\varepsilon$. Ekkor az $A_{j_1,\dots,j_k}$ halmazok átmérője kisebb mint ε. Válasszunk egy a $\{j_1,\dots,j_k\}$ szám k-asokból álló véges D_k halmazt, melyre $\sum\limits_{(j_1,\dots,j_k)\colon\ (j_1,\dots,j_k)\in D_k}\mu(A_{j_1,\dots,j_k})>1-\frac{\varepsilon}{2}$. Ekkor létezik olyan $n_0=n_0(\varepsilon)$ index úgy, hogy a

$$B(\varepsilon)=\bigcup_{(j_1,\dots,j_k)\in D_k}\ \bigcup_{n\ge n_0}\left(B_{j_1,\dots,j_k}\triangle B_{j_1,\dots,j_k}(n)\right)$$

halmazra $\lambda(B(\varepsilon)) < \frac{\varepsilon}{2}$, ahol $A \triangle B = (A \setminus B) \cup (B \setminus A)$ egy A és B halmaz szimmetrikus differenciáját jelöli. Ekkor a $B_1(\varepsilon) = \bigcup\limits_{(j_1,\dots,j_k)\in D_k} (B_{j_1,\dots,j_k} \setminus B(\varepsilon)$ halmazra $\lambda(B_1(\varepsilon)) > 1-\varepsilon$. A $B_2 = B_1 \cap Y_0 \cap \bigcap\limits_{n=1}^{\infty} Y_{0,n}$, ahol $Y_{0,n} = [0,1) \setminus \bigcup\limits_{k=1}^{\infty} \bigcup\limits_{j_1,\dots,j_k} \{b_{j_1,\dots,j_k}(n)\}$ szintén teljesíti a $\lambda(B_2(\varepsilon)) > 1-\varepsilon$ relációt. Továbbá $n > n_0$ és $y \in B_2(\varepsilon)$ esetén $\rho(\xi_n(y), \xi(y) \le \varepsilon$, mert ebben az esetben $\xi_n(y)$ és $\xi(y)$ ugyanannak az $A_{j_1,\dots,j_k}$ halmaznak a lezártjában van. Innen következik, hogy

$$\lambda\left(\limsup_{n\to\infty} |\xi_n - \xi| \ge \varepsilon\right) < \varepsilon.$$

Mivel ez a reláció érvényes minden $\varepsilon > 0$-ra innen következik a feladat állítása.

5.) Legyen $\bar X = X \times [0,1]$, és definiáljuk az $(\Omega, \mathcal{A})$ mértékteret mint $\Omega = \bar X \times X \times \cdots \times X \times \cdots$, $\mathcal{A}$ az $\bar X$ illetve X tereken értelmezett σ-algebrák szorzat σ-algebrája. Az $\omega = (x, u, x_1, x_2, \dots) \in \Omega$ pontban definiáljuk a ξ és ξ_n valószínűségi változókat a $\xi(\omega) = x$, $\xi_n(\omega) = x_n$, $n = 1, 2, \dots$ formulákkal. A P mérték definicióját az $\bar X$ téren bevezetett $\bar\mu = \mu \times \lambda$ mérték segítségével és alkalmasan definiált $Q_n((x,u), A)$, $(x,u) \in \bar X = X \times [0,1]$, $A \subset X$ az X halmaz mérhető részhalmaza, feltételes valószínűségi mértékek segítségével definiáljuk úgy, hogy az x_n koordináta feltételes eloszlása rögzített $(x,u) \in \bar X$ feltétel mellett $Q_n((x,u), \cdot)$ legyen, és az x_n koordináták rögzített $(x,u) \in \bar X$ feltétel mellett legyenek feltételesen függetlenek. (Az, hogy $Q_n((x,u), A)$ feltételes valószínűségi mérték azt jelenti, hogy $Q_n((x,u), \cdot)$ valószínűségi mérték az $(X, \mathcal{A})$ téren minden $(x,u) \in \bar X$ pontra, és $Q_n(\cdot, A)$ mérhető függvény minden $A \in A$ halmazra.) Formálisan megfogalmazva, legyen

$$P(A \times A_1 \times \cdots \times A_n) = \int_{(x,u)\in A} Q_1((x,u), A_1) \dots Q_n((x,u), A_n))\, d\mu(x)\, du$$
$$\text{minden } A \in \bar X,\ A_j \in X,\ j = 1, \dots, n \text{ halmazra.}$$

(A mértékelmélet nem triviális eredményeiből következik, hogy az ily módon definiált P illetve annak kiterjesztése valóban valószínűségi mértéket határoz meg. Érdemes megjegyezni, hogy ez a tény olyan eredményből következik (Tulcea–Ionescu tétel), mely nem használja a tér topológiai tulajdonságait, ezért használható nem szeparábilis metrikus terekben is.)

Minden $k = 1, 2, \dots,$ számra definiáljuk az X tér olyan $\mathcal{A}_k = \{A_{1,k}, A_{2,k}, \dots\}$ particióját, melyre $\bar d(A_{j,k}) < \frac{1}{k}$ és $\mu(\partial A_{j,k}) = 0$, $j = 1, 2, \dots,$ ahol $\bar d(A)$ jelöli az A halmaz átmérőjét ∂A pedig a határát. Minden $k = 1, 2, \dots$-ra definiáljunk egy olyan $m(k)$ indexet, melyre $\mu\left(\bigcup\limits_{j\ge m(k)} A_{j,k}\right) \le \frac{1}{k^2}$ és $\mu(A_{j,k}) > 0$ minden $1 \le j \le m(k)$-ra. (Ez utóbbi feltétel teljesítése érdekében esetleg átindexelhetjük az $A_{j,k}$ halmazokat.) Ezután tekintsünk egy olyan $1 = n_1 < n_2 < n_3 < \cdots$ számsorozatot, melyre teljesül, hogy

$$|\mu_n(A_{j,k}) - \mu(A_{j,k})| < \frac{\mu(A_{j,k})}{k^2 m(k)}, \quad \text{minden } 1 \le j < m(k) \text{ és } n_k \le n < n_{k+1} \text{számra.}$$

Ez lehetséges, mert $\lim_{n\to\infty} \mu_n(A_{j,k}) = \mu(A_{j,k})$ minden k és j számra a $\mu(\partial A_{j,k}) = 0$ reláció miatt.

Vezessük be a

$$\lambda_{j,k,n} = \frac{\min\left(\mu(A_{j,k}), \mu_n(A_{j,k})\right)}{\mu(A_{j,k})}, \quad n_k \le n < n_{k+1}, \quad j < m(k)$$

számokat. Nyílván $1 - \frac{1}{k^2} \le \lambda_{j,k,n} \le 1$. Definiáljuk a $Q_n((x,u),\cdot)$ feltételes eloszlásokat először csak bizonyos feltételeket kielégítő $(x,u) \in \bar{X}$ és n párokra. Legyen

$$Q_n((x,u),C) = \frac{\mu_n(C \cap A_{j,k})}{\mu_n(A_{j,k})}, \quad \text{ha } n_k \le n < n_{k+1},\ x \in A_{j,k},\ 1 \le j < m(k),$$
$$\text{és } 0 \le u \le \lambda_{j,k,n} \quad \text{minden } C \in \mathcal{A} \text{ halmazra.}$$

A még nem tekintett tartományon úgy akarjuk definiálni a $Q_n(\cdot,\cdot)$ feltételes eloszlásokat, hogy a P mérték vetülete az n-ik koordinátára μ_n legyen. Ennek érdekében vezessük be a

$$P_n = \left(1 - \sum_{j=1}^{m(k)-1} \min\left[\mu_n(A_{j,k}), \mu(A_{j,k})\right]\right), \quad n_k \le n < n_{k+1}$$

számokat, (ez a szám $\bar{X} = X \times [0,1]$ halmaz azon részének a μ mértéke, ahol még nem definiáltuk a $Q_n(\cdot,\cdot)$ feltételes eloszlást)) és a következő $\bar{\mu}_n$ valószínűségi mértéket az X téren:

$$\bar{\mu}_n(C) = \frac{1}{P_n}\left[\mu_n(C) - \sum_{j=1}^{m(k)-1} \min\left[\mu_n(A_{j,k}), \mu(A_{j,k})\right] \frac{\mu_n(C \cap A_{j,k})}{\mu_n(A_{j,k})}\right], \quad C \in \mathcal{A},$$

és legyen

$$Q_n((x,u),C) = \bar{\mu}_n(C), \quad \text{ha } C \in \mathcal{A}, \quad n_k \le n < n_{k+1},$$
$$\text{és } x \in A_{j,k},\ j \ge m(k) \text{ vagy } x \in A_{j,k}, j < m(k) \text{és } \lambda_{j,k,n} < u \le 1.$$

Azt állítjuk, hogy

$$\int Q_n((x,u),C)\mu(\,dx)\,du = \mu_n(C), \quad C \in \mathcal{A}. \qquad (+)$$

Valóban, ha $n_k \le n < n_{k+1}$, akkor $1 \le j < m(k)$ és $C \in \mathcal{A}$ esetén

$$\int Q_n((x,u),C \cap A_{j,k})\mu(\,dx)\,du = \mu(A_{j,k})\lambda_{j,k,n}\frac{\mu_n(C \cap A_{j,k})}{\mu_n(A_{j,k})} + \mu_n(C \cap A_{j,k})$$
$$- \min\left(\mu(A_{j,k}), \mu(A_{j,k})\right) \frac{\mu_n(C \cap A_{j,k})}{\mu_n(A_{j,k})} = \mu_n(C \cap A_{j,k}).$$

Másrészt legyen $B_k = X \setminus \left(\bigcup_{j=1}^{m(k)-1} A_{j,k} \right)$. Ekkor az $\{A_{j,k}, 1 \le j < m(k), B_k\}$ halmazrendszer az X tér egy particióját adja. Speciálisan, $A_{j,k} \cap B_k = \emptyset$ minden $1 \le j < m(k)$-ra.

$$\int Q_n((x,u), C \cap B_k)\mu(\,dx)\,du = \mu_n(C \cap B_k).$$

Ezeket az azonosságokat összeadva kapjuk a (+) relációt, ami azt jelenti, hogy a ξ_n valószínűségi változók eloszlása az előírt μ_n mérték minden $n = 1, 2, \dots$ számra.

Másrészt, $P\left(\sup_{n_k \le n < n_{k+1}} \rho(\xi_n, \xi) \ge \frac{1}{k} \middle| \bar{\xi} = (x,u) \right) = 0$, ha $(x,u) \notin X_1(k) \subset \bar{X}$, ahol $\bar{\xi}(x, u, x_1, x_2, \dots) = (x, u)$ és

$$X_1(k) = B_k \times [0,1] \cup \bigcup_{j=1}^{m(k)-1} \left\{ (x,u)\colon x \in A_{j,k}, \inf_{n_k \le n < n_{k+1}} \lambda_{j,k,n} \le u \le 1 \right\}.$$

Ezért

$$P\left(\sup_{n_k \le n < n_k} \rho(\xi_n, \xi) > \frac{1}{k} \right) \le \mu \times \lambda(X_1(k))$$
$$= \sum_{j=1}^{m(k)} \left(1 - \min_{n_k \le n < n_{k+1}} \lambda_{j,k,n} \right) \mu(A_{j,k}) + \mu(B_k) \le \frac{2}{k^2}.$$

Mivel

$$\sum_{k=1}^{\infty} P\left(\sup_{n_k \le n < n_k} \rho(\xi_n, \xi) > \frac{1}{k} \right) \le \sum_{k=1}^{\infty} \frac{2}{k^2} < \infty$$

a Borel–Cantelli lemmából következik, hogy $\xi_n \to \xi$ egy valószínűséggel, ha $n \to \infty$.

6.) Legyen F tetszőleges zárt halmaz. Mivel $F = \bigcap_{\delta \to 0} F^\delta = F$, ahol F^δ jelöli az F halmaz δ sugarú nyílt környezetét, ezért tetszőleges $\varepsilon > 0$-hoz létezik olyan $\delta > 0$, hogy $\mu(F^\delta) < \mu(F) + \varepsilon$. Mivel $\xi_n \Rightarrow \xi$ sztochasztikusan, ezért minden $\delta > 0$-hoz és $\varepsilon > 0$-hoz létezik olyan $n_0 = n_0(\varepsilon, \delta)$ index, hogy $P(\rho(\xi, \xi_n) > \delta) < \varepsilon$. Továbbá $\{\omega\colon \xi(\omega) \notin F^\delta\} \subset \{\omega\colon \xi_n(\omega) \notin F\} \cup \{\omega\colon \rho(\xi_n(\omega), \xi(\omega)) > \delta\}$, ezért $1 - \mu(F^\delta) \le 1 - \mu_n(F) + \varepsilon$, és $\mu(F) + \varepsilon \ge \mu(F^\delta) \ge \mu_n(F) - \varepsilon$, ha $n \ge n_0$. Innen $\limsup_{n \to \infty} \mu_n(F) \le \mu(F) + 2\varepsilon$. Mivel ez minden $\varepsilon > 0$-ra igaz, kapjuk, hogy $\limsup_{n \to \infty} \mu_n(F) \le \mu(F)$ minden zárt F halmazra, azaz $\mu_n \Rightarrow \mu$.

7.) Annak bizonyításához, hogy $\xi = F^{-1}(\eta)$ F eloszlású, elég azt megmutatni, hogy

$$\{\omega\colon \eta(\omega) < F(x)\} \subset \{\omega\colon F^{-1}(\eta)(\omega) < x\} \subset \{\omega\colon \eta(\omega) \le F(x)\}.$$

A jobboldali tartalmazás igaz, mert ha $F^{-1}(\eta)(\omega) < x$ akkor létezik olyan $h > 0$ szám, melyre $F^{-1}(\eta)(\omega) = \sup\{u\colon F(u) < \eta(\omega)\} < x-h$. Ebből viszont következik, hogy $\eta(\omega) \le F(x)$. Ha ugyanis $F(x) < \eta(\omega)$ volna, akkor az x szám is szerepelne azon u számok között, melyek szuprémuma meghatározza $F^{-1}(\eta)$-t, és ez ellentmond az $F^{-1}(\eta)(\omega) < x - h$ egyenlőtlenségnek.

A baloldali tartalmazás bizonyításához vegyük észre, hogy $\eta(\omega) < F(x)$ esetén $\eta(\omega) = F(x) - h$, és $F^{-1}(\eta(\omega)) = F^{-1}(F(x) - h)) = \sup\{v\colon F(v) < F(x) - h\}$ alkalmas $h > 0$ számmal. Viszont $\sup\{v\colon F(v) < F(x) - h\} < x$, mert $F(x)$ balról folytonos függvény, így $F(v) < F(x) - h$-ból következik, hogy létezik olyan $\delta = \delta(h) > 0$ szám, melyre $v < x-\delta$, tehát $F^{-1}(\eta(\omega)) = \sup\{v\colon F(v) < F(x)-h\} \le x - \delta < x$. Tehát a baloldali tartalmazás is érvényes.

Annak bizonyítása érdekében, hogy $\tilde F(\xi, \varepsilon)$ egyenletes eloszlású a $[0, 1]$ intervallumban, minden x valós számra definiáljuk a $z(x) = \sup\{y\colon F(y) < x\}$ számot. Mivel az F eloszlásfüggvény balról folytonos, ezért $F(z) \ge x$. Vizsgáljuk külön az $F(z) = x$ és $F(z) < x$ eseteket.

Ha $F(z) = x$, akkor $F(u + 0) < x$ minden $u < z$ számra, és $\{\omega\colon \tilde F(\xi(\omega), \varepsilon(\omega)) < x\} = \{\omega\colon \xi(\omega) < z\}$. Ezért $P(\tilde F(\xi, \varepsilon) < x) = P(\xi < z) = F(z) = x$.

Ha $F(z) < x$, akkor $F(z + 0) \ge x$, és

$$\begin{aligned}&\{\omega\colon \tilde F(\xi(\omega, \varepsilon(\omega)) < x\}\\ &\qquad = \{\omega\colon \xi(\omega) < z\} \cup \{\omega\colon \xi(\omega) = z, F(z) + \varepsilon(\omega)[F(z + 0) - F(z)] < x\}.\end{aligned}$$

Ezért

$$\begin{aligned}P(\tilde F(\xi, \varepsilon) < x) &= P(\xi < z) + P(\xi(\omega) = z)P(\varepsilon(\omega)[F(z + 0) - F(z)] < x)\\ &= F(z) + [F(z + 0) - F(z)]\frac{x - F(z)}{F(z + 0) - F(z)} = x.\end{aligned}$$

Innen következik a feladat második állítása.

8.) Az előző feladat eredményeiből közvetlenül következik, hogy a $\bar\xi$ valószínűségi változó F, $\bar\eta$ pedig G eloszlású. Továbbá $\tilde F(\bar{\bar\xi}, \varepsilon)$ egyenletes $\bar{\bar\eta}$ pedig G eloszlású valószínűségi változók. Annak érdekében, hogy megmutassuk azt hogy a $(\bar\xi, \bar\eta)$ és $(\bar{\bar\xi}, \bar{\bar\eta}$ vektorok eloszlása megegyezik, elég megmutatni, hogy $F^{-1}(\tilde F(\bar{\bar\xi}, \varepsilon)) = \bar{\bar\xi}$ egy valószínűséggel, mert ez azt jelenti, hogy mind a két véletlen vektor előállítható, (0 mértékű halmaz figyelmen kívül hagyásával) mint egy a $[0, 1]$ intervallumon egyenletes eloszlású valószínűségi változó transzformáltja, (mind a két koordináta ugyanannak a valószínűségi változónak a transzformáltja) és mind a két vektor előállításában ugyanazt a transzformációt alkalmazzuk. Sőt ezt a bizonyítandó állítást tovább lehet gyengíteni. Elég megmutatni azt, hogy $P(F^{-1}(\tilde F(\bar{\bar\xi}, \varepsilon)) \le \bar{\bar\xi}) = 1$, mert ha két azonos eloszlású valószínűségi változó közül az egyik egy valószínűséggel nagyobb a másiknál, akkor ezek a változók egy valószínűséggel megegyeznek. Lássuk be ezt az állítást. Ha $\bar{\bar\xi}(\omega) = x$, akkor

$$F^{-1}(\tilde F(\bar{\bar\xi}(\omega), \varepsilon(\omega))) = \sup\{u\colon F(u) < \tilde F(x, \varepsilon(\omega))\},$$

és mivel $F(v) \geq \tilde{F}(x, \varepsilon(\omega))$ ha $v > x$, innen következik, hogy $(F^{-1}(\tilde{F}(\bar{\bar{\xi}}(\omega), \varepsilon(\omega))) \leq x = \bar{\bar{\xi}}(\omega)$.

9.) Jegyezzük meg, hogy amennyiben a $\xi = F^{-1}(\zeta)$, $\eta = G^{-1}(\zeta)$, egy a $[0,1]$ intervallumban egyenletes eloszlású valószínűségi változó fenti transzformáltjai, azaz a kvantilis transzformációt alkalmazzuk, akkor a feladatban szereplő egyenlőtlenség két oldala egyenlő. A bizonyítandó egyenlőtlenséget bizonyítsuk be először abban a speciális esetben, amikor a ξ és η valószínűségi változók eloszlása a számegyenes egy véges $X = \{x_1, \dots, x_n\}$, $x_1 < x_2 < \cdots < x_n$, részhalmazába van koncentrálva. Legyen $p_j = P(\xi = x_j)$, $q_k = P(\eta = x_k)$, $r(x_j, y_k) = P(\xi = x_j, \eta = x_k)$, $1 \leq j, k \leq n$. Vezessük be a

$$
\begin{aligned}
r(x_j, x_k, y_j, y_k) &= \min\left(r(x_j, y_k), r(x_k, y_j)\right), \\
r &= \max_{\substack{\{x_j, x_k, y_j, y_k\} \in X \times X \times X \times X \\ x_j < x_k,\ y_j < y_k}} r(x_j, x_k, y_j, y_k)
\end{aligned}
$$

mennyiségeket. Vegyük észre, hogy ha a (ξ, η) vektort a kvantilis transzformáció segítségével definiáljuk, akkor a fent definiált r számra $r = 0$. Továbbá, az $r = 0$ feltétel és a ξ változó F illetve az η változó G eloszlása meghatározza a (ξ, η) vektor együttes eloszlását is. Valóban, feltehetjük az általánosság megszorítása nélkül, hogy az $(\Omega, \mathcal{A}, P)$ ahol a (ξ, η) vektor definiálva van, tér $\Omega = X \times X$ a diszkrét topológia által meghatározott $\mathcal{A}$ σ-algebrával, $\xi(x_j, x_k) = x_j$ $\eta(x_j, x_k) = x_k$, $1 \leq j, k \leq n$. Ekkor a kívánt tulajdonságú P mérték reprezentálható egy olyan szállítási probléma megoldásával, ahol az x_j pontban levő p_j tömegeket át kell szállítani úgy, hogy az x_k pontba q_k tömeget szállítunk, és ekkor $r(x_j, x_k) = P(\xi = x_j, \eta = x_k)$ az x_j pontból az x_k pontba szállított tömeg. Ekkor az $r = 0$ feltételt kielégítő szállítás olyan, hogy az x_1 pontban levő p_1 tömeget először az x_1 pontba majd, ha még maradt tömeg az x_1 pontban akkor az x_2 majd x_3 pontokba szállítjuk és így tovább, majd az x_2 pontból a legkisebb indexű még nem teljesen betöltött pontba szállítunk, utána az azt következőbe és így tovább. Ugyanezt tesszük ezután az x_3, x_4 pontokban levő tömegekkel és így tovább. Azt állítjuk, hogy

$$
\min_{\substack{\xi \text{ eloszlása } F \\ \eta \text{ eloszlása } G}} E\Phi(\xi, \eta) = E\Phi(\bar{\xi}, \bar{\eta}), \tag{b}
$$

ahol $\bar{\xi}, \bar{\eta}$ egy F és G marginális eloszlással rendelkező véletlen vektor, melyre $r = 0$.

Valóban, ha a (ξ, η) eloszlása olyan, hogy $r \neq 0$, akkor létezik olyan (x_j, x_k, y_j, y_k) számnégyes, melyre $\bar{r} = r(x_j, x_k, y_j, y_k) = \min\left(r(x_j, y_k), r(x_k, y_j)\right) > 0$. Bevezetünk egy új $\bar{r}(x_j, x_k)$, $1 \leq j, k \leq n$ együttes eloszlásokat, úgy hogy egy ezekkel az eloszlásokkal meghatározott $(\bar{\xi}, \bar{\eta})$ véletlen vektor marginális eloszlásai F és G, továbbá $E\Phi(\xi, \eta) \leq E\Phi(\tilde{\xi}, \tilde{\eta})$, és amennyiben Φ szigorúan konvex függvény, akkor szigorú egyenlőtlenség érvényes. Ezeket a valószínűségeket a következőképpen definiáljuk.

$$\begin{aligned}\bar{r}(x_j, y_j) &= r(x_j, y_j) + \bar{r}\\ \bar{r}(x_k, y_k) &= r(x_k, y_k) + \bar{r}\\ \bar{r}(x_j, y_k) &= r(x_j, y_k) - \bar{r}\\ \bar{r}(x_k, y_j) &= r(x_k, y_j) - \bar{r}\\ \bar{r}(x, y) &= r(x, y) \quad \text{egyébként.}\end{aligned}$$

Ekkor a $(\bar{\xi}, \bar{\eta})$ véletlen vektor marginális eloszlásai az előírtak, és

$$E\Phi(\tilde{\xi} - \tilde{\eta}) - E\Phi(\xi - \eta) = \bar{r}\left(\Phi(x_j - y_j) + \Phi(x_k - y_k) - \Phi(x_j - y_k) - \Phi(x_k - y_j)\right).$$

Ez a kifejezés nem negatív, és szigorúan pozitív, ha $\Phi(\cdot)$ szigorúan konvex. Ugyanis

$$x_j - y_k < \begin{matrix} x_j - y_j \\ x_k - y_k \end{matrix} < x_k - y_j,$$

és ezért a Φ függvény konvexitása miatt

$$\Phi(x_j - y_j) + \Phi(x_k - y_k) \le \Phi(x_j - y_k) + \Phi(x_k - y_j),$$

(felhasználjuk azt is, hogy $(x_j - y_j) + (x_k - y_k) = (x_j - y_k) + (x_k - y_j)$) és szigorú egyenlőtlenség érvényes akkor, ha Φ szigorúan monoton függvény.

Belátjuk ennek a relációnak a segítségével a (b) formulát. A fenti módszerrel, amennyiben $r \neq 0$, akkor olyan új a marginális eloszlásokra tett feltételeknek eleget tevő olyan új (ξ, η) véletlen vektor tudunk definiálni, melyre $E\Phi(\xi, \eta)$ nem nő, sőt amennyiben a Φ függvény szigorúan konvex, (például $\Phi(x) = x^2$) akkor ez a várható érték szigorúan monoton csökken. Továbbá minden egyes lépésben azon (x_j, y_j, x_k, y_k) négyesek halmaza, melyekre $\bar{r} = r(x_j, x_k, y_j, y_k) = \min\left(r(x_j, y_k), r(x_k, y_j)\right) > 0$ változik. Ezért véges sok lépésben eljutunk ahhoz az együttes eloszláshoz, melyre $r = 0$, és ez az eloszlás a Φ függvény esetén minimumot egy szigorúan konvex függvény esetén pedig szigorú minimumot szolgáltat.

Ha F és G két olyan eloszlásfüggvény, mely véges intervallumra van koncentrálva, akkor ezeket az eloszlásokat érdemes az $F_n(x) = F\left(\frac{[nx]}{n}\right)$ és $G_n(x) = G\left(\frac{[nx]}{n}\right)$ eloszlásokkal közelíteni. Pontosabban egy $F(x)$ és $G(y)$ marginális eloszlásokkal rendelkező $H(x, y)$ két-dimenziós eloszlást érdemes a $H_n(x, y) = H\left(\frac{[nx]}{n}, \frac{[ny]}{n}\right)$ eloszlással közelíteteni, ahol $[u]$ az u szám egész részét jelöli. Alkalmazva az állítást a már bebizonyított esetben és n-nel végtelenhez tartva megkapjuk az állítás bizonyítását az adott esetben. Az általános eset hasonlóan visszavezethető az utóbbi esetre alkalmas határátmenet segítségével, ha a valószínűségi változókat csonkítjuk valamely $\pm u$ szinten. Ezen a csonkításon azt értjük, hogy ha a (ξ, η) vektor a $[-u, u] \times [-u, u]$ négyzeten kívül veszi fel az értékét akkor a csonkított változó értéke az origó. Azután végrehajtva az $u \to \infty$ határátmenetet bebizonyíthatjuk a kívánt eredményt az általános esetben. Megjegyezzük, hogy a Φ függvény konvexitásából következik, hogy $\Phi(x) \ge Ax + B$ minden x valós számra, alkalmas A és B konstansokkal, ezért $E\Phi(\xi - \eta) \ge -|A|E(|\xi| + |\eta| - |B| > -\infty$ a feladat feltételeinek teljesülése esetén. Érdemes megjegyezni, hogy alkalmas lineáris függvényt hozzáadva a $\Phi(x)$ függvényhez redukálhatjuk a problémát arra az esetre, amikor $\Phi(x) \ge 0$ minden x számra, és $\Phi(0) = 0$. Ugyancsak feltehetjük, hogy $E\Phi(\xi - \eta) < \infty$. Ezek az észrevételek segíthetik a határátmenet végrehajtását. A bizonyítandó egyenlőtlenség baloldalán alkalmazhatjuk a mononoton konvergencia tételt a jobboldalán pedig a Fatou lemmát. A részletek kidolgozását elhagyjuk.

10.) Legyen γ a μ illetve ν mérték vetülete az X_2 térre. Legyen továbbá $P(x, A)$ az $(X_1 \times X_2, \mathcal{A}_1 \times \mathcal{A}_2, \mu)$ téren az $A \times X_2$ alakú mérhető részhalmazok feltételes eloszlása előírt $x \in X_2$, $Q(x, C)$ pedig az $(X_2 \times X_3, \mathcal{A}_2 \times A_3, \nu)$ téren az $X \times C$ alakú mérhető részhalmazok feltételes eloszlása előírt $x \in X_2$ értékek szerint. Azaz legyen $\mu(A \times B) = \int_B P(x, A)\gamma(\,dx)$ minden $A \in \mathcal{A}$ halmazra, és $\nu(B \times C) = \int_B P(x, C)\gamma(\,dx)$ minden $C \in \mathcal{A}_3$ halmazra. (Ilyen feltételes eloszlások léteznek teljes szeparábilis metrikus tereken.) Definiáljuk az $A \times B \times C$ alakú mérhető halmazok mértékét az $X_1 \times X_2 \times X_3$ téren $P(A \times B \times C) = \int_B P(x, A)Q(x, C)\gamma(\,dx)$ képlet segítségével. Ezt a mértéket kiterjeszthetjük az $(X_1 \times X_2 \times X_3, \mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{A}_3)$ térre. és ez a mérték teljesíti a kívánt feltételeket.

11.) Legyen $P(x, A)$, $x \in X$, A az X halmaz mérhető részhalmaza, a ν mérték második koordinátájának feltételes eloszlása feltéve az első koordinátát, azaz legyen $P(x, \cdot)$ valószínűségi mérték az X halmaz mérhető részhalmazain minden $x \in X$ pontra, $P(\cdot, A)$ mérhető függvény minden mérhető $A \subset X$ halmazra, és $\nu(A \times B) = \int_A P(x, B)\mu(\,dx)$ minden mérhető $A \subset X$ és $B \subset X$ halmazra. A valószínűségszámítás (mértékelmélet) egyik eredménye szerint ilyen feltételes eloszlás létezik.

Tekintsük a $(Q, \mathcal{B}, \lambda)$ valószínűségi mezőt, ahol $Q = [0, 1]$, $\mathcal{B}$ a Borel σ-algebra a $[0, 1]$ intervallumon, és λ a Lebesgue mérték a $[0, 1]$ σ-algebrán. Azt állítjuk, hogy meg lehet konstruálni $\zeta(x, u)$, $u \in Q = [0, 1]$, $x \in X$ valószínűségi változóknak az $x \in X$ paraméterrel indexezett halmaza a $(Q, \mathcal{B}, \lambda)$ valószínűségi mezőn úgy, hogy $\zeta(x, \cdot)$ $P(x, \cdot)$ eloszlású valószínűségi változó minden $x \in X$ pontra, ahol $P(x, A)$ az előző paragrafusban definiált feltételes eloszlás, és $\zeta(\cdot, u)$ mérhető függvény az (X, ρ) téren minden $u \in [0, 1]$ számra. Valójában egy ennél kissé gyengébb állítást látunk be. Csak annyit állítunk, hogy a $\zeta(x, \cdot)$ valószínűségi változó eloszlása $P(x, \cdot)$ mérték a μ mérték szerint majdnem minden x-re. Ez viszont nem jelent valódi megszorítást, mert a $P(x, \cdot)$ mértékeket egy az x pontoknak egy μ mérték szerint null mértékű halmazán megváltoztatva ismét feltételes eloszlást kapunk.

A fenti állítás bizonyítását megkapjuk a 3. feladatban megadott konstrukció természetes adaptációjának segítségével. Ott megadtunk egy lehetőséget arra, hogy hogyan lehet tetszőleges α valószínűségi mérték esetén α eloszlású valószínűségi változót konstruálni a $(Q, \mathcal{B}, \lambda)$ valószínűségi mezőn. Azt kell ellenőriznümk, hogy ezt a konstrukciót alkalmazva olyan $\mu = P(x, \cdot)$ eloszlású $\zeta(x, u)$ valószínűségi változókat konstruálhatunk szimultán módon minden $x \in X$ pontra, melyekre a $\zeta(\cdot, u$ függvények (X, ρ) mérhetőek minden $u \in [0, 1[$ pontra. Az egyetlen változtatás, melyet a 3. feladatban leírt konstrukcióhoz képest teszünk abban áll, hogy a konstukcióban felhasznált partició $A_{j_1,\dots,j_k}$ határára a $\mu(\partial A_{j_1,\dots,j_k})=0$ feltétel teljesül. Ekkor viszont a feltételes eloszlás tulajdonságai miatt a $P(x,(\partial A_{j_1,\dots,j_k}))=0$ reláció is teljesül μ majdnem minden x pontra a particiókban szerepllő összes (megszámlálható sok) $A_{j_1,\dots,j_k}$ halmazra. Alkalmazzuk a 3. feladatban leírt konstrukciót mindegyik olyan x pontra, mely teljesíti, hogy a $P(x,\cdot)$ mérték szerint a fenti particiók elemei null-mértékűek. A kivételes null-mértékű (a μ mérték szerint) halmazon legyen $\zeta(x,u)=x_0$, ahol x_0 egy rögzített az x ponttól független pont. Az így konstruált valószínűségi változók teljesítik a kívánt tulajdonságokat.

Legyen $\eta(\omega)=\zeta(\xi(\omega),\chi(\omega))$ a fent konstruált $\zeta(x,u)$ változók segítségével. Azt állítjuk, hogy ez az η változó teljesíti a feladat feltételeit, azaz $P(\xi(\omega)\in A,\eta(\omega)\in B)=\nu(A\times B)$ minden mérhető $A\subset X$ és $B\subset X$ halmazokra. Viszont mivel a $(\chi(\omega),\xi(\omega))$ vektor eloszlása $dv\,\mu(dx)$ a $[0,1]\times X$ téren, ezért

$$\begin{aligned}P(\xi(\omega)\in A,\eta(\omega)\in B)&=\int\left(\int I\left(x\in A,\zeta(x,v)\in B\right)\,dv\right)\mu(\,dx)\\&=\int I(x\in A)P(x,B)\mu(\,dx)=\int_A P(x,B)\mu(\,dx)=\nu(A\times B),\end{aligned}$$

ahol $I(C)$ a C halmaz indikátorfüggvénye. Innen következik a feladat állítása.

12.) Létezik a μ és ν mértékeknek a következő tipusú felbontása. $\mu=\gamma+\mu_1$, $\nu=\gamma+\nu_1$, ahol a γ mérték a μ és ν mértékek "közös része", a μ_1 és ν_1 mértékek pedig szingulárisak. Ez azt jelenti, hogy létezik olyan $C\in A$ halmaz, melyre a μ_1 mérték a C a ν_1 mérték pedig az $X\setminus C$ halmazra van koncentrálva, azaz $\mu_1(X\setminus C)=0$, és $\nu_1(C)=0$. Valóban, tekintsünk egy a μ és ν mértéket domináló mértéket (például választhatjuk a $\frac{\mu+\nu}{2}$ mértéket,) legyen $f(x)$ a μ és $g(x)$ a ν mérték sűrűségfüggvénye e domináló mérték szerint. Ekkor legyen γ a $\min(f(x),g(x))$, μ_1 az $f(x)-\min(f(x),g(x))$ és ν_1 a $g(x)-\min(f(x),g(x))$ sűrűségfüggvénnyel definiált mérték e domináló mérték szerint. Végül legyen C a $C=\{x\colon f(x)\ge g(x)\}$ halmaz. Ez a felbontás teljesíti a kívánt tulajdonságokat, továbbá $\mu(A)-\nu(A)\le\mu(C)-\nu(C)$ minden $A\in\mathcal{A}$ halmazra. Vegyük észre, hogy $\mu(C)-\nu(C)=\mu(C)-\gamma(C)=\mu_1(C)=\mu_1(X)=1-\gamma(X)$, ahonnan $\sup\limits_{A\in\mathcal{A}}(\mu(A)-\nu(A))=1-\gamma(X)$. Hasonlóan becsülhetőek a $\nu(A)-\mu(A)$ alakú mennyiségek. Ezért $\mathrm{Var}\,(\mu,\nu)=\sup\limits_{A\in\mathcal{A}}|\mu(A)-\nu(A)|=1-\gamma(X)$. Ha ξ és η μ illetve ν eloszlású valószínűségi változók, akkor

$$P(\xi\ne\eta)\ge P(\{\xi\in C\}\cap\{\eta\notin C\})\ge\mu(C)-\nu(C)=\mathrm{Var}\,(\mu,\nu).$$

Annak érdekében, hogy olyan konstrukciót készítsünk, melyben a fenti egyenlőtlenség két oldala megegyezik, definiáljuk a következő $(\Omega,\mathcal{B},P)$ valószínűségi mezőt. Legyen $\Omega=X\cup(X\times X)$, és $\mathcal{B}$ a természtes σ-algebra Ω-n, amelyiknek a megszorítása az X halmazra $\mathcal{A}$, az $X\times X$ halmazra pedig $\mathcal{A}\times\mathcal{A}$. Definiáljuk a P mértéket a következő módon: A P mérték megszorítása az X halmazra legyen γ, az $X\times X$ halmazra pedig $D(\mu_1\times\nu_1)$, ahol $D^{-1}=\mu_1(X)=\nu_1(X)$, azaz

D a természetes normáló faktor. Definiáljuk a ξ és η valószínűségi változókat a következő módon. Ha $\omega = x \in X$, akkor $\xi(\omega) = \eta(\omega) = x$, ha $\omega = (x_1, x_2) \in X \times X$, akkor $\xi(\omega) = x_1$, $\eta(\omega) = x_2$. Ekkor a ξ valószínűségi változó μ az η valószínűségi változó pedig ν eloszlású. Továbbá, $P(\xi = \eta) = \gamma(X) = 1 - \mathrm{Var}\,(\mu, \nu)$, mert A P mérték megszorítása az $X \times X$ halmazra a $C \times (X \setminus C)$ halmazra van koncentrálva, ahol $\xi(\omega) \neq \eta(\omega)$. Innen következik a feladat állítása.

Megjegyzés: Technikai okokból definiáltuk a szélsőértéket szolgáló valószínűségi változókat az $\Omega = X \cup (X \times X)$ és nem az $\Omega_1 = X \times X$ valószínűségi mezőn, ahol természetes lett volna dolgozni. Ez utóbbi esetben ugyanis a γ mértéket a $D = \{(x,x) \colon x \in X\}$ átlóra kellett volna koncentrálni. Viszont van olyan eset, amikor ez a D átló az $\Omega_1 = X \times X$ halmaz nem mérhető részhalmaza, és ez a tény problémát okoz. Olyan konstrukciót akartunk adni, amelyik ebben az esetben is működik.

Kiegészítés

A König–Hall tétel bizonyítása. A feltétel szükségessége nyílvánvaló. Az elégségességet az Y halmaz számossága szerinti teljes indukcióval bizonyítjuk. Ha $|Y| = 1$, akkor az állítás nyilvánvaló. Tegyük fel, hogy a feltétel elégségességét tudjuk $|Y| = k < n$ esetén, és lássuk be $|Y| = n$-re. Két esetet különböztetünk meg.

a.) Létezik olyan A halmaz, melyre $0 < |A| = k < n$ és $|B(A)| = |A|$.

b.) Minden olyan $A \subset Y$ halmazra, melyre $0 < |A| < n$ $|B(A)| > |A|$.

Az a.) esetben azt állítjuk, hogy az állítás feltételei teljesülnek mind az $\bar{Y} = A$, $\bar{Z} = B(A)$ mind az $\bar{Y} = Y \setminus A$, $\bar{Z} = Z \setminus B(A)$ halmazokra és a $d(y,z)$ függvény megszorítására ezekre az $\bar{Y} \times \bar{Z}$ halmazokra. Innen és az indukciós feltevésből következik a feltétel elégségessége az a.) esetben. Az $\bar{Y} = A$, $\bar{Z} = B(A)$ esetben ez az állítás nyilvánvaló. Ha $\bar{Y} = Y \setminus A$ és $\bar{Z} = Z \setminus B(A)$ tekintsünk egy $C \subset Y \setminus A$ halmazt. Legyen $\bar{C} = C \cup A$. Ekkor $|C| = |\bar{C}| - k$, $|B(\bar{C})| \geq |\bar{C}|$, és $B(C) \cap \bar{Z} \supset B(\bar{C}) \setminus B(A)$, ahonnan $|B(C) \cap \bar{Z}| \geq |B(\bar{C})| - k \geq |\bar{C}| - k$, és $|B(C) \cap \bar{Z}| \geq |C|$, amint állítottuk.

A b.) esetben tekintsünk egy egy $z_j \in Z$ pontot, melyre $d(y_1, z_j) = 1$. Párosítsuk az y_1 pontot az z_j ponttal. A tétel bizonyításának befejezéséhez elég belátni azt, hogy az $\bar{Y} = Y \setminus \{y_1\}$ és $\bar{Z} = Z \setminus \{z_j\}$ halmazok teljesítik a feladat feltételeit a b.) esetben. Ez azonban nyilvánvaló, mert a b.) esetben $A \subset Z$ esetén $|B(A) \cap \bar{Z}| \geq |B(A)| - 1 \geq |A|$.

A König–Hall tétel folytonos verziójának bizonyítása. A feltétel szükségessége ebben az esetben is nyílvánvaló. Az elégségesség bizonyítását visszavezetjük az eredeti König–Hall tételre.

Tekintsük először azt a speciális esetet, amikor létezik olyan N egész szám, melyre $u(y_j) = \dfrac{k_j}{N}$ minden $y_j \in Y$ és $v(z_l) = \dfrac{p_l}{N}$ minden $z_l \in Z$-ra, ahol k_j és p_l egész számok. Ekkor tekintsük a következő páros gráfot. $\bar{Y} = \{(y_j, m(j)), y_j \in Y, 1 \leq m(j) \leq k_j\}$ $\bar{Z} = \{(z_l, n(l)), z_l \in Z, 1 \leq n(l) \leq p_l\}$, $\bar{d}((y_j, m(j)), (z_l, n(l))) = d(y_j, z_l)$. Ekkor az $(\bar{Y}, \bar{Z}, \bar{d}(\cdot,\cdot))$ páros gráf teljesíti a König–Hall tétel feltételeit. Elég ugyanis a König–Hall tétel feltételeit azokra az $A \subset \bar{Y}$ halmazokra ellenőrizni, melyekre $\bar{y} = (y_j, m(j)) \in A$

esetén $(y_j,k)\in A$ minden (y_j,k), $1\le k\le k_j$ pontra, és ez a feltétel megegyezik azzal, hogy $\sum\limits_{z\in B(A)} v(z)\ge \sum\limits_{y\in A} u(y)$ minden $A\in Y$-ra.

Ezért létezik az $\bar Y$ és $\bar Z$ halmaz pontjainak egy olyan párosítása, melyben a párba állitott $\bar y$ és $\bar z$ pontokra $\bar d(\bar y,\bar z)=1$. Legyen $\bar w(\bar y,\bar z)=1$, ha $\bar y$ és $\bar z$ párba vannak állítva, és $\bar w(\bar y,\bar z)=0$ egyébként. Ekkor a $w(y_j,z_l)=\dfrac{1}{N}\sum\limits_{\substack{\bar y=(y_j,m(j)),1\le m(j)\le k_j\\ \bar z=(z_l,n(l)),1\le n_l\le p_l}}\bar w(\bar y,\bar z)$ függvény teljesíti a tétel állítását ebben a speciális esetben.

Az általános eset visszavezethető a már bebizonyított esetre alkalmas approximációval. Az általánosság megszorítása nélkül feltehetjük, hogy $\sum\limits_{y\in Y} u(y)=\sum\limits_{z\in Z} v(z)=1$. Minden $N=1,2,\dots$-re definiáljuk a következő közelítő rendszert. Legyen $\bar Y=\bar Y_N=Y\cup\{r+1\}$, $\bar Z=\bar Z_N=Z\cup\{s+1\}$, $\bar d(y,z)=\bar d_N(y,z)=d(y,z)$, ha $y\in Y$, $z\in Z$, és $\bar d(y_{r+1},z)=\bar d_N(y_{r+1},z)=\bar d(y,z_{s+1})=\bar d_N(y,z_{s+1})=1$, (azaz az y_{r+1} és z_{s+1} pontok a másik halmaz minden pontjával össze vannak kötve), $\bar u_N(y)=\dfrac{[Nu(y)]}{N}$, ha $y\in Y$, $\bar v_N(z)=\dfrac{[Nv(z)]}{N}$, ha $z\in Z$, ahol $[u]$ az u szám egész részét jelöli, továbbá $\bar u_N(y_{r+1})=\sum\limits_{y\in Y}(u(y)-\bar u_N(y))=1-\sum\limits_{y\in Y}\bar u_N(y)$, és $\bar v_N(y_{s+1})=\sum\limits_{z\in Z}(v(z)-\bar v_N(z))=1-\sum\limits_{z\in Z}\bar v_N(z)$. Vegyük észre, hogy az új rendszer is teljesíti a tétel feltételeit. Ugyanis, ha $A\subset Y$, akkor a $\bar Z_N$ halmaz A halmazzal összekötött pontjaiból álló halmaz $\bar B_N(A)=B(A)\cup\{z_{s+1}\}$, és $\sum\limits_{y\in A}\bar u_N(y)\le\sum\limits_{y\in A}u(y)\le\sum\limits_{z\in B(A)}v(z)\le\sum\limits_{z\in B(A)}\bar v_N(z)+\bar v_N(z_{s+1})\le\sum\limits_{z\in\bar B(A)}\bar v_N(z)$. Ha pedig $y_{r+1}\in A$, akkor $\bar B_N(A)=\bar Z_N$, és $\sum\limits_{y\in A}\bar u_N(y)\le 1=\sum\limits_{z\in\bar Z_N}\bar v_N(z)$. Ezért ebben az esetben alkalmazhatjuk a tételt.

Tekintsünk minden $N=1,2,\dots$-ra egy a feltételt kielégítő $\bar w_N(y,z)$, $y\in\bar Y$, $z\in\bar Z$ "szállítási függvényt". Mivel $0\le\bar w_N(y,z)\le 1$ minden $y\in\bar Y$ és $z\in\bar Z$ pontra, és $N=1,2,\dots$ számra, ezért ki lehet választani egy olyan $N_k\to\infty$ részsorozatot, melyre a $w(y,z)=\lim\limits_{k\to\infty} w_{N_k}(y,z)$ limesz minden $y\in\bar Y$ és $z\in\bar Z$ pontra létezik. Ez a $w(y,z)$ függvény teljesíti a tétel állítását.