Nagy eltérések elmélete. A Szanov tétel.

Egy előző feladatsorban, a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsorban annak az eseménynek a valószínűségét vizsgáltuk, hogy független, egyforma eloszlású valószinűségi változók átlaga egy a várható értéknél nagyobb számot vesz fel. Az ottani eredményekből következik, hogy ha ξ_k, $k = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók, $S_n = \sum\limits_{k=1}^{n} \xi_k$, akkor az $\frac{S_n}{n}$ átlag $F_n(y|x) = P\left(\frac{S_n}{n} \ge y \middle| \frac{S_n}{n} \ge x\right)$ feltételes eloszlása $y > x > E\xi_1$ esetén teljesíti a $\lim\limits_{n\to\infty} F_n(y|x) = 0$ tulajdonságot. Ezt úgy is interpretálhatjuk, hogy ha az átlag nagyobb mint x, akkor az ezen feltétel mellett majdnem egy valószínűséggel az x pont kis környezetébe van koncentrálva. Tehát eme feltétel mellett az átlag eloszlása erősen lokalizált.

Egy ennél erősebb lokalizációs tulajdonságot kívánunk megfogalmazni és bebizonyítani. Ez azt a tényt fejezi ki, hogy alkalmas és nagyon általános feltételek mellett nemcsak az átlag, hanem az összeadandók, azaz a $(\xi_1, \dots, \xi_n)$ vektor együttes feltételes eloszlása is erősen lokalizált az adott feltételek mellett. Ezt az állítást akarjuk pontosan megfogalmazni, és annak néhány érdekes következményét levezetni. Ennek érdekében bevezetjük a következő fogalmat. Ha $(\xi_1, \dots, \xi_n)$ valószínűségi változók egy véges sorozata, akkor definiáljuk ezek μ_n empirikus eloszlását a következő módon: A $(\xi_1(\omega), \dots, \xi_n(\omega))$ valószínűségi változók által meghatározott $\mu_n = \mu_n(\omega)$ empririkus eloszlás (véletlen) valószínűségi mérték a számegyenesen. Tetszőleges mérhető $\mathbf{A} \subset R^1$ halmazra $\mu_n(\mathbf{A}) = \frac{1}{n} \cdot \#\{j\colon 1 \le j \le n,\ \xi_j \in \mathbf{A}\}$.

Tekintsük a $(\xi_1, \dots, \xi_n)$ vektor empirikus eloszlások a feltételes eloszlását bizonyos feltételek mellett. (Például az $S_n = \sum\limits_{k=1}^{n} \xi_k > nx$ feltétel mellett.) "Tipikus esetekben" ez a (véletlen) feltételes eloszlás egy előírt eloszlás közelében koncentrálódik nagy valószínűséggel. Ezt az eloszlást annak az észrevételnek a segítségével találhatjuk meg, hogy az adott feltétel mellett az empirikus eloszlás ennek az eloszlásnak a közelébe esik a legnagyobb valószínűséggel. Így azt meg lehet határozni egy variációs probléma megoldásának a segítségével. Ilyen módon is lehet vizsgálni a nagy eltérések elméletét, sőt bizonyos problémák vizsgálatában ez a módszer hatásosabb mint a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor módszere. Viszont e módszer alkalmazásához szükség van a (bizonyítandó) nagy eltérés tétel megfogalmazására általánosabb terekben. A nagy eltérés tételt csak olyan térben tudjuk megfogalmazni, ahol van értelme közelségről beszélni. Ezért természetes lenne metrikus terekben dolgozni. De érdemes a nagy eltérés tételt általánosabb körülmények között, topológikus terekben megfogalmazni és vizsgálni, mivel ilyen módon néhány fontos esetben sokkal tartalmasabb eredményt kapunk.

Nagy eltérés tétel megfogalmazása topológikus terekben. *Legyen adva egy $(X, \mathcal{X})$ topológikus tér, és legyen $\mathcal{A}$ a topológia által definiált σ-algebra. Tegyük fel, hogy az $(X, \mathcal{X})$ tér úgynevezett T_3 tér, azaz tetszőleges $x \in X$ pontra és olyan zárt $\mathbf{F}$ halmazra, amelyre $x \notin \mathbf{F}$ léteznek olyan diszjunkt nyílt $\mathbf{G}_1$ és $\mathbf{G}_2$ halmazok, amelyekre $x \in \mathbf{G}_1$ és $\mathbf{F} \subset \mathbf{G}_2$. Legyen μ_n, $n = 1, 2, \dots$, valószínűségi mértékek sorozata az $(X, \mathcal{A})$ téren. Azt mondjuk, hogy az $(X, \mathcal{X})$ téren értelmezett μ_n, $n = 1, 2, \dots$, valószínűségi* *mértékek sorozata teljesíti a nagy eltérés tételt egy $I(x)$, $x \in X$, függvénnyel, amelyre $I(x) \ge 0$, (és $I(x)$ felveheti a ∞ értéket is), ha*

i.) *Az $I(\cdot)$ függvény alulról félig folytonos, azaz minden $\varepsilon > 0$ számra és $x \in X$ pontra az $\{y\colon I(y) > I(x) - \varepsilon\}$ halmaz az x pont nyílt környezete.*

ii.) $\limsup\limits_{n\to\infty} -\frac{1}{n}\log \mu_n(\mathbf{G}) \le \inf\limits_{x\in\mathbf{G}} I(x)$ *minden nyílt $\mathbf{G} \subset X$ halmazra.*

iii.) $\liminf\limits_{n\to\infty} -\frac{1}{n}\log \mu_n(\mathbf{F}) \ge \inf\limits_{x\in\mathbf{F}} I(x)$ *minden zárt $\mathbf{F} \subset X$ halmazra.*

Bár a nagy eltérés tételt általános topologikus terekben fogalmaztuk meg, részletesebben csak azzal az esettel fogunk foglalkozni, amikor a tekintett tér egy mérhető téren értelmezett valószínűségi mértékek tere alkalmas topológiával. A minket leginkább érdeklő eset az, amikor valamilyen téren (például a számegyenesen) definiált független egyforma eloszlású valószínűségi változókból készített empirikus eloszlásokat tekintjük, az adott téren levő valószínűségi mértékeket ellátjuk alkalmas topológiával, és azt akarjuk belátni, hogy az empirikus eloszlások, mint ezen tér valószínűségi mértékein értelmezett véletlen mértékek, (azaz megadjuk annak a valószínűségét, hogy az empirikus eloszlás valószínűségi mértékek egy családjába esik) teljesítik a fent definiált nagy eltérés tételt. Ezenkívül bizonyos feladatok arról szólnak, hogy független valószínűségi változók részletösszegei teljesítik a nagy eltérés tételt az ebben a feladatsorban definiált értelemben is.

Az első feladatban megmutatjuk, hogy a nagy eltérés tételben szereplő $I(x)$ függvényt a μ_n mértékek egyértelműen meghatározzák.

1.) Teljesítsék a μ_n valószínűségi mértékek a nagy eltérés tételt valamilyen $I(x)$ alulról folytonos függvénnyel egy $(X,\mathcal{X})$ T_3 tulajdonságú topologikus térben. Minden $x \in X$-re teljesülnek a következő relációk:

$$I(x) = \sup_{\substack{\mathbf{G}\colon \mathbf{G} \text{ nyílt halmaz}\\ x\in\mathbf{G}}} \limsup_{n\to\infty} -\frac{1}{n}\log \mu_n(\mathbf{G})$$

$$I(x) = \sup_{\substack{\mathbf{F}\colon \mathbf{F} \text{ zárt halmaz}\\ x\in\operatorname{Int}\mathbf{F}}} \liminf_{n\to\infty} -\frac{1}{n}\log \mu_n(\mathbf{F}).$$

A következő két feladat, illetve az előttük megfogalmazott megjegyzés nem kapcsolódik szorosan a feladatsorhoz. Viszont az elmélet további, itt nem tárgyalt finomításában nagyon hasznos.

A nagy eltérés tétel megfogalmazásában természetes, hogy a zárt halmazok esetében a iii.) tulajdonságban $\limsup$ és nem $\lim$ a nyílt halmazok esetén pedig a ii.) tulajdonságban $\liminf$ és nem $\lim$ szerepel. Zárt halmazok esetén például az egy pontból álló halmazok és sűrűségfüggvénnyel rendelkező független egyforma eloszlású valószínűségi változók átlagainak eloszlásai mutatnak példát arra, hogy miért csak a iii.)-ban megfogalmazott gyengébb állítást érdemes megkövetelni. Az, hogy a nagy eltérés tétel teljesülése esetén ii.) tulajdonságban nem mindig írható limesz például a feladatsor 8. feladatának eredményéből következik.

A nagy eltérés tétel ii.) és iii.) tulajdonságában azért nem írhatunk limeszt, mert egy halmaz indikátorfüggvényének a határa erősen befolyásolhatja a halmaz μ_n mértékeinek a viselkedését. Ilyen jellegű problémák az analízis más területén is megjelennek. Ilyenkor természetes az állítás átfogalmazása halmazok indikátorfüggvényei helyett folytonos függvények segítségével. Ez történik a következő, az irodalomban Varadhan lemmának nevezett állításban. Azután belátjuk ennek az állításnak egy megfordítását. De ebben a megfordításban a nagy eltérés tételnek csak egy gyengített változatát látjuk be, ahol a iii.) tulajdonságot zárt halmazok helyett csak kompakt halmazokra fogalmazzuk meg. Annak bizonyítása, hogy konkrét esetekben a nagy eltérés tétel iii.) tulajdonsága zárt nem kompakt halmazokra is érvényes további finom érvelést igényel. Így például az ebben a feladatsorban tárgyalt Szanov tétel bizonyításában ez az egyik fő nehézség.

2.) Teljesítse egy μ_n, $n = 1, 2, \dots$, valószínűségi mértéksorozat egy $(X, \mathcal{X})$ topológikus tér Borel σ-algebráján a nagy eltérés tételt egy $I(x)$, $x \in X$, alulról félig folytonos függvénnyel. Ekkor tetszőleges folytonos, korlátos $g(x)$ függvényre

$$\lim_{n\to\infty} \frac{1}{n} \log \int e^{ng(x)} \mu_n(dx) = \sup_{x\in X} \{g(x) - I(x)\}$$

3.) Teljesítse egy μ_n, $n = 1, 2, \dots$, valószínűségi mértéksorozat a 2. feladatban felírt azonosságot valamilyen alulról folytonos, nem negatív $I(\cdot)$ függvénnyel minden folytonos és korlátos $g(\cdot)$ függvényre egy $(X, \mathcal{X})$ topológikus térben. Tegyük fel továbbá, hogy az $(X, \mathcal{X})$ tér úgynevezett T_ρ tér, azaz tetszőleges $x \in X$ pontra és $\mathbf{F} \subset X$ zárt halmazra, amelyekre $x \notin \mathbf{F}$ létezik olyan folytonos, korlátos $h(\cdot)$ függvény, amelyre $h(x) = 1$, és $h(u) = 0$ $u \in \mathbf{F}$-re. Ekkor a μ_n mértéksorozat teljesíti a nagy eltérés tétel ii.) és következő iii$'$.) tulajdonságát:

iii$'$.) $\liminf\limits_{n\to\infty} -\frac{1}{n} \log \mu_n(\mathbf{K}) \ge \inf\limits_{x\in\mathbf{K}} I(x)$ minden kompakt $\mathbf{K} \subset X$ halmazra.

A következő 4.–8. feladatokban megmutatjuk, hogy a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsorban vizsgált független egyforma eloszlású valószínűségi változók átlagainak eloszlásai teljesítik a nagy eltérés tételnek azt a formáját is, amelyet ennek a feladatsornak az elején fogalmaztunk meg.

4.) Legyen ξ olyan valószínűségi változó, amelyre $R(t) = Ee^{t\xi} < \infty$ elég kis $t > 0$-ra, és definiáljuk az $I(x) = \sup\limits_{t\ge 0}(tx - \log R(t))$ függvényt minden $x \ge E\xi$ számra. ($R(t) = \infty$, ha $e^{t\xi}$ nem integrálható.) E halmazon az $I(x)$ függvény monoton nő, azaz $I(x) \le I(y)$, ha $E\xi < y < x$, ($E\xi = -\infty$ is megengedett). Ha $I(x) < \infty$, akkor $I(y) < I(x)$, azaz ebben az esetben szigorú egyenlőtlenség is érvényes. Továbbá, ha $I(x) < \infty$ és $E\xi < x$, akkor $\lim\limits_{y<x,\, y\to x} I(y) = I(x)$. Az $I(x)$ függvény konvex. Ezért tetszőleges olyan x számra, amelyre $I(x+\varepsilon) < \infty$ elég kis $\varepsilon > 0$-ra $\lim\limits_{y>x,\, y\to x} I(y) = I(x)$.

Megjegyzés: Definiáljuk az $L = \sup\{x\colon\ I(x) < \infty\}$ számot. ($L = \infty$ lehetséges.) Az előző feladat szerint az $I(x)$ függvény az $[E\xi, L)$ intervallumon folytonos, konvex és szigorúan monoton nő, sőt mint az alábbi érvelés segítségével látható, ez az állítás érvényes ennek az intervallumnak $[E\xi, L]$ lezártján is. Továbbá, ha $x > L$, akkor $I(x) = \infty$. Könnyen bizonyítható az is, hogy $L < \infty$ akkor és csak akkor, ha a ξ valószínűségi változó egy valószínűséggel felülről korlátos. Ha ugyanis $P(\xi > x) = A(x) > 0$ minden $x > 0$ számra, akkor

$$P\left(\sum_{k=1}^{n} \xi_k > nx\right) \ge e^{n \log A(x)} \quad \text{minden } n\text{-re,}$$

ahol ξ_k, $k = 1, 2, \dots$, függetlenek és a ξ valószínűségi változókkal azonos eloszlásúak. Innen az erre az összegre kapott felső becslés alapján $I(x) \le -\log A(x) < \infty$. Ha viszont a ξ valószínűségi változó felülről korlátos, akkor az $R(t) = Ee^{t\xi}$ momentumgeneráló függvény véges minden $t \ge 0$-ra, és $\lim\limits_{t\to\infty} [\log R(t)]' < \infty$. Ekkor azonban (ld. például a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor 12. feladatának eredményét) $I(x) = \infty$ elég nagy x-re.

5.) Legyenek ξ_n, $n = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók a számegyenesen $R(t) = Ee^{t\xi_1}$ momentumgeneráló függvénnyel ($R(t) = \infty$, ha $e^{t\xi_1}$ nem integrálható), és jelölje μ_n a $\frac{S_n}{n} = \frac{\sum_{k=1}^{n} \xi_k}{n}$ valószínűségi változó eloszlását. Az egyszerűség kedvéért ebben a feladatban csak azt az esetet tekintsük, amikor az $Ee^{t\xi_1}$ momentumgeneráló függvény véges az origó egy kis környezetében. A μ_n mértékek teljesítik a nagy eltérés tételt az $I(\cdot)$ függvénnyel, ahol $I(x) = \sup\limits_{t\ge 0}(tx - \log R(t))$, ha $x \ge E\xi_1$, és $I(x) = \sup\limits_{t\le 0}(tx - \log R(t))$, ha $x \le E\xi_1$.

6.) Lássuk be a nagy eltérés tételnek az ebben a feladatsorban megfogalmazott alakját független, egyforma eloszlású valószínűségek átlagának eloszlásaira az általános esetben. Az általános esetben $I(x) = \sup\limits_{t\ge 0}(tx - \log R(t))$ ha $x \ge E\xi_1$, és $I(x) = \sup\limits_{t\le 0}(tx - \log R(t))$, ha $x \le E\xi_1$. Ha a ξ_1 valószínűségi változónak nincs várható értéke (azaz a pozitív rész várható értéke végtelen, a negatív rész várható értéke minusz végtelen), akkor $I(x) = 0$ $(= \sup\limits_{t}(tx - \log R(t)) = -R(0))$ minden x számrea.

7.) Ha a ξ valószínűségi változó $R(t) = Ee^{t\xi}$ momentumgeneráló függvénye elég kis $t > 0$-ra véges, és $E\xi_1 = -\infty$, akkor $\lim\limits_{x\to-\infty} I(x) = \lim\limits_{x\to-\infty} \sup\limits_{t\ge 0}(tx - \log R(t)) = 0$.

8.) Független, egyforma eloszlású valószínűségi változók átlagának az eloszlása teljesíti a nagy eltérés tétel ii) tulajdonságának következő élesebb formáját is:

ii′.) $\lim\limits_{n\to\infty} -\frac{1}{n} \log \mu_n(G) = \inf\limits_{x\in G} I(x)$ minden nyílt halmazra.

Adjunk példát valószínűségi mértékek μ_n sorozatára a számegyenesen, amely teljesíti a nagy eltérés tételt, de nem teljesíti a ii′.) tulajdonságot.

Az $I(x)$ függvényt $I(x) = \sup\limits_{-\infty<t<\infty}(tx - \log R(t))$ alakban is megadhatjuk a *Nagy* *eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor 18a.) feladatának eredménye alapján.

A következő feladatokban az empirikus eloszlásfüggvény viselkedéséről adunk nagy eltérés tipusú becsléseket, és ezekből vezetjük le a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsorban bebizonyított nagy eltérés tétel (nehezebben bizonyítható) alsó becslését független, egyforma eloszlású valószínűségi változók átlagára. Annak valószínűségére adunk jó becslést, hogy független F eloszlású valószínűségi változók empirikus eloszlásfüggvénye egy G, $G \neq F$ eloszlásfüggvény közelében van, és ennek segítségével adunk becslést a ξ_k valószínűségi változók átlagának eloszlására. Ennek a módszernek az az előnye, hogy nem kötődik erősen a számegyenes geometriájához. Ezért általánosabb esetekben is alkalmazható, például olyankor, amikor az összeadandók egy az Euklideszi térnél sokkal gazdagabb térben veszik fel értékeiket. Ezekben a vizsgálatokban nagyon fontos az alábbiakban bevezetett I-divergencia fogalma.

Az I-divergencia definiciója. *Legyen F és G két eloszlás a számegyenesen. (Az általános esetben legyen μ és ν két valószínűségi mérték ugyanazon az $(Y,\mathcal{B})$ téren.)*

A G eloszlásnak az F eloszlás szerinti $I(G\|F)$ I-divergenciát a következő módon definiáljuk:

$$I(G\|F)=\int_{\left\{u\colon \frac{dG}{dF}(u)>0\right\}}\log\frac{dG}{dF}(u)\,dG(u),$$

ha az G eloszlás által indukált mérték abszolút folytonos az $F(x)$ által indukált mértékre, és $I(G\|F)=\infty$, ha az nem abszolút folytonos.

Általánosabban, egy ν valószínűségi mértéknek egy μ valószínűségi mérték szerinti $I(\nu\|\mu)$ I-divergenciáját az

$$I(\nu\|\mu)=\int\log\frac{d\nu}{d\mu}(y)\,d\nu(y)$$

képlettel definiáljuk, ha ν abszolút folytonos a μ mértékre nézve, és $I(\nu\|\mu)=\infty$ ha ν nem abszolút folytonos a μ mértékre nézve. A fenti integrál úgy értendő, hogy a $\mathbf{B}=\left\{y\colon \frac{d\nu}{d\mu}(y)=0\right\}$ halmazon, ahol $\log\frac{d\nu}{d\mu}(y)=-\infty$ és $\nu(\mathbf{B})=0$, $\nu(\mathbf{B})\cdot(-\infty)=0\cdot(-\infty)=0$.

A későbbiekben tekinteni fogunk egy rögzített μ mértéket valamilyen mértéktéren vagy abban a speciális esetben, ha a számegyenesen dolgozunk egy rögzített F eloszlásfüggvényt a számegyenesen. Ha adva van egy $\xi_k(\omega)$, $k=1,2,\dots$, független μ eloszlású valószínűségi változókból álló sorozat, akkor elkészítjük e sorozat által meghatározott $\mu_n=\mu_n(\omega)$ empirikus mértékek sorozatatát. Be fogjuk látni, hogy (ha alkalmas topológiát vezetünk be a valószínűségi mértékek terén), akkor az empirikus mértékek teljesítik a nagy eltérés tételt az $I(\cdot)=I(\cdot\|\mu)$ divergencia függvénnyel.

Tekintsük azt az esetet, amikor egy F eloszlásfüggvény van adva a számegyenesen, és $\xi_1,\xi_2,\dots$ független F eloszlású valószínűségi változók sorozata, $S_n(\omega)=\sum\limits_{k=1}^{n}\xi_k(\omega)$, és $F_n(\omega, u)$ a $\xi_1(\omega), \dots, \xi_n(\omega))$ sorozat empirikus eloszlásfüggvénye. Vegyük észre, hogy $\frac{S_n(\omega)}{n} = \int u\, F_n(\omega\, du)$, tehát az

$$A_n(x) = \left\{\omega\colon \frac{S_n(\omega)}{n} > x\right\} \quad \text{és} \quad \bar{A}_n(x) = \left\{\omega\colon \int u\, F_n(\omega, du) > x\right\}$$

halmazok megegyeznek. Mivel az $x > E\xi_1$ esetben az $A_n(x)$ illetve $\bar{A}_n(x)$ események valószínűségét a nagy eltérés tétel felhasználásával az $I(\cdot) = \sup\limits_{t\ge 0}(tx - \log R(t))$, – ahol $R(t) = E^{t\xi_1}$, – illetve az $\bar{I}(\cdot) = I(\cdot\|F)$ divergencia függvény segítségével becsülhetjuk meg, ezért a fenti érvelés sugallja az alábbi, a 9. feladat (a) formulájában megfogalmazott azonosságot.

A következő 9. és 10. feladatban ezt az azonosságot látjuk be. A 9. feladatban azt az egyszerűbb esetet tekintjük, amikor a $[\log R(t)]' = x$ egyenlet megoldható. Ekkor explicit módon megadható az, hogy az azonosság két oldalán szereplő kifejezés hol veszi fel a szélsőértékét. A 10. feladat az általános esettel foglalkozik. Ekkor az I-divergencia kiszámításában hasonló problémák és gondolatok jelennek meg, mint a *Nagy eltérések elméle: Független, valós értékű valószínűségi változók* feladatsor megfelelő részében.

9.) Legyen μ és ν két valószínűségi mérték egy $(Y, \mathcal{B})$ téren. Ekkor $I(\nu\|\mu) \ge 0$, és egyenlőség csak $\mu = \nu$ esetén lehetséges.

Legyen adva egy F mérték a számegyenesen, és legyen $x \ge m = \int u\, dF(u)$. (Nincs ilyen x szám, ha $\int u\, dF(u) = \infty$, és tetszőleges x számot tekinthetünk, ha az $\int u\, dF(u)$ nem létezik.) Legyen G egy másik mérték a számegyenesen. Ha az $R(t) = \int e^{tu} F(\,du)$ momentumgeneráló függvény véges valamely $t > 0$-ra, és $\int uG(\,du) = x$, akkor $I(G\|F) \ge tx - \log R(t)$. Lássuk be ennek az egyenlőtlenségnek egy bízonyos értelmű megfordítását először abban a speciális esetben, amikor a $[\log R(t)]' = x$ egyenlet megoldható, és a megoldás az $R(t)$ függvény értelmezési tartományának belsejében van, ahol $R(t) = \int_{-\infty}^{\infty} e^{tu} F(\,du)$. Azaz, mutassuk meg, hogy e feltételek mellett

$$\inf_{G\colon \int u\, dG(u) \ge x} I(G\|F) = \sup_{t\ge 0} \left(tx - \log R(t)\right), \quad \text{ahol } R(t) = \int e^{tu} F(\,du) \tag{a}$$

Ebben az esetben az (a) formula baloldalán az infimum, a jobboldalán pedig az az szuprémum felvétetik. A baloldalon az infimum a $G(du) = F_t(u) = \frac{e^{tu}\, dF(u)}{R(t)}$ valószínűségi mértéknél, a baloldalon a szuprémum a $[\log R(t)]' = x$ megoldásánál éretik el.

10.) Lássuk be a 9. feladat (a) formuláját tetszőleges F eloszlásra és $x \ge m = \int u\, dF(u)$ számra. Azt a speciális esetet kivéve, amikor az $R(t)$ függvény minden $t \ge 0$-ra értelmezve van, és $\lim\limits_{t\to\infty} [\log R(t)]' = x$ minden $\varepsilon > 0$ számra létezik olyan G eloszlás, amelyre $x < \int u\, dG(u) < x + \varepsilon$ szigorú egyenlőtlenséggel az előző reláció mindkét oldalán, és $I(G\|F) < \sup\limits_{t\ge 0}(tx - \log R(t)) + \varepsilon$.

Legyen $F(x)$ és $G(x)$ két különböző eloszlásfüggvény a számegyenesen. Legyenek $\xi_1,\dots,\xi_n$ független $F(x)$ eloszlású valószínűségi változók, $F_n(x)=\frac{1}{n}\#\{k\colon\ \xi_k<x, 1\le k\le n\}$, a ξ_k, $k=1,\dots,n$ sorozathoz tartozó empirikus eloszlásfüggvény. A valószínűségszámítás egy klasszikus tétele, a Glivenko–Cantelli tétel szerint az $F_n(x)$ empirikus eloszlásfüggvény egy valószínűséggel konvergál szuprémum normában az $F(x)$ eloszlásfüggvényhez. Be fogjuk látni, hogy annak valószínűsége, hogy az $F_n(x)$ empirikus eloszlásfüggvény a $G(x)$ eloszlásfüggvény közelében van, $F\ne G$, (az n változóban) exponenciálisan kicsi. Pontosabban,

$$\lim_{\varepsilon\to0}\limsup_{n\to\infty}-\frac{1}{n}\log P\left(\sup_x|F_n(x)-G(x)|<\varepsilon\right)=I(G\|F), \tag{+1}$$

és

$$\lim_{\varepsilon\to0}\liminf_{n\to\infty}-\frac{1}{n}\log P\left(\sup_x|F_n(x)-G(x)|<\varepsilon\right)=I(G\|F), \tag{+2}$$

Ezt az állítást direkt módon fogjuk belátni, noha az egyszerűen következik a később tárgyalt Szanov tételnek nevezett nagy eltérés tételből.

A következő feladatban ezt az állítást abban a speciális esetben látjuk be, amikor az F és G eloszlások által meghatározott μ_F és μ_G mértékek véges sok pontba vannak koncentrálva.

11.) Legyen μ és ν két valószínűségi mérték egy véges $X=\{x_1,\dots,x_k\}$ halmazon. Vezessük be a $\mu(\{x_l\})=p_l$, $\nu(\{x_l\})=q_l$, $l=1,\dots,k$, $\sum\limits_{l=1}^k p_l=\sum\limits_{l=1}^k q_l=1$ jelölést. Tegyük fel, hogy $p_l>0$ minden $1\le l\le k$-ra. Tekintsük az X tér $X^n=\{x_1,\dots,x_k\}^n$ n-szeres direkt szorzatát, és vezessük be az $y=(y_1,\dots,y_n)\in X^n$ sorozatok terén a μ mérték $\mu^{(n)}$ direkt szorzatát, azaz legyen $\mu^{(n)}(y)=\prod\limits_{j=1}^n\mu(y_j)$ minden $y=(y_1,\dots,y_n)\in X^n$ sorozatra. Definiáljuk továbbá a

$$\chi_n=\chi_n(y,l)=\chi_n(y_1,\dots,y_n,l)=\frac{1}{n}\times\#\{s\colon\ 1\le s\le n,\ y_s=x_l\},$$

számokat minden $y=(y_1,\dots,y_n)\in X^n$ sorozatra és $l=1,\dots,k$ indexre, azaz legyen $\chi_n(y,l)$ az x_l számok relatív gyakorisága az $y=(y_1,\dots,y_n)$ sorozatban. Ekkor tetszőleges $\varepsilon>0$ számra és $n>n_0(\varepsilon,\nu,\mu)$ indexre

$$\begin{aligned}\exp&\left\{-n\left(\sum_{l=1}^k q_l\log\frac{q_l}{p_l}+C\varepsilon\right)\right\}\\ &\le\quad \mu^{(n)}\left(|\chi_n(y,l)-q_l|<\varepsilon \text{ minden } 1\le l\le k \text{ indexre}\right)\\ &\qquad\le\exp\left\{-n\left(\sum_{l=1}^k q_l\log\frac{q_l}{p_l}-C\varepsilon\right)\right\}\end{aligned} \tag{3}$$

alkalmas $C=C(\mu,\nu)$ számmal. A fenti formula úgy értendő, hogy a benne szereplő $q_l\log\frac{q_l}{p_l}$ összeadandó nullával egyenlő, ha $q_l=0$.

Bizonyítsuk be ennek az egyenlőtlenségnek a segítségével a (+1) és (+2) állítást abban a speciális esetben, ha az F és G eloszlás által meghatározott mértékek véges sok pontba vannak koncentrálva.

Be akarjuk látni a (+1) és (+2) állításokat tetszőleges F és G eloszlásokra. Ehhez először belátjuk, hogy az F és G eloszlások elég finom véges F' és G' diszkretizációjával elérhető, hogy $I(G'\|F')$ jól közelítse az $I(G\|F)$ I-divergenciát. Ezt az állítást is két részletben fogjuk bizonyítani. Először egy speciális particiót tekintünk, aztán ennek segítségével bebizonyítjuk az állítást tetszőleges finom particióra. Ezeket az állításokat általánosabb mértékterekben fogjuk bebizonyítani. Azt mondjuk, hogy halmazoknak egy $\mathcal{C}=(\mathbf{C}_1,\dots,\mathbf{C}_k)$ rendszere, ahol $\mathbf{C}_r$ az Y halmaz mérhető részhalmaza, $1\le r\le k$, egy $(Y,\mathcal{B})$ tér (véges) particiója, ha $\mathbf{C}_1,\dots,\mathbf{C}_k$ diszjunkt halmazok, és $\bigcup\limits_{r=1}^{k}\mathbf{C}_r=Y$.

Legyen adva egy $(Y,\mathcal{B})$ mérhető tér, és annak véges $\mathcal{C}$ particiója. A $\mathcal{C}$ particióhoz válasszunk egy $(x_1,\dots,x_k)$, $x_r\in\mathbf{C}_r$, $r=1,\dots,k$, (csak a $\mathcal{C}$ particiótól függő) pontrendszert. Egy az $(Y,\mathcal{B})$ téren értelmezett μ valószínűségi mérték $\mu_{\mathcal{C}}$ vetületét a $\mathcal{C}$ particióra úgy definiáljuk, mint azt a $\mu_{\mathcal{C}}$ mértéket, amely az x_r, $1\le r\le k$, pontokba van koncentrálva, és $\mu_{\mathcal{C}}(\{x_r\})=\mu(\mathbf{C}_r)$, $r=1,\dots,k$.

12.) Legyen adva két μ és ν valószínűségi mérték egy $(Y,\mathcal{B})$ téren, valamint az $(Y,\mathcal{B})$ tér egy $\mathcal{C}$ véges particiója. Mutassuk meg, hogy $I\left(\nu_{\mathcal{C}}\|\mu_{\mathcal{C}}\right)\le I(\nu\|\mu)$. Adva egy $\varepsilon>0$ konstruáljunk olyan $\mathcal{C}$ véges particiót, amelyre $I\left(\nu_{\mathcal{C}}\|\mu_{\mathcal{C}}\right)\ge I(\nu\|\mu)-\varepsilon$, ha $I(\nu\|\mu)<\infty$, és $I\left(\nu_{\mathcal{C}}\|\mu_{\mathcal{C}}\right)\ge\varepsilon^{-1}$, ha $I(\nu\|\mu)=\infty$.

Érdemes megfogalmazni a 12. feladat eredmény alábbi következményét:

$$I(\nu\|\mu)=\sup_{\mathcal{C}\text{ az }(Y,\mathcal{B})\text{ tér véges particiója}} I\left(\nu_{\mathcal{C}}\|\mu_{\mathcal{C}}\right)$$

Be akarjuk látni, hogy az $I(\nu\|\mu)$ I-divergenciát nemcsak egy speciális hanem minden elég finom partició segítségével jól lehet approximálni. Ehhez azonban bizonyos megszorítást kell tenni, nevezetesen azt, hogy a tekintett tér $\mathcal{B}$ σ-algebrája jól közelíthető véges particiókkal. Először bizonyítsunk be egy technikai jellegű mértékelméleti approximációs lemmát, amely hasznos a bizonyítandó állítás igazolásában.

13.) Legyen $(X,\mathcal{A})$ egy mérhető tér, legyenek azon adva $\mu_1,\dots,\mu_k$ valószínűségi mértékek, és legyen $\mathcal{A}_1\subset\mathcal{A}_2\subset\cdots$ egymásba skatulyázott σ-algebrák sorozata, amelyre $\mathcal{A}=\sigma\left(\bigcup\limits_{l=1}^{\infty}\mathcal{A}_l\right)$, azaz $\mathcal{A}$ az $\mathcal{A}_l$ σ-algebrák uniója által generált legszükebb σ-algebra. Ekkor minden $\mathbf{A}\in\mathcal{A}$ halmazra és $\varepsilon>0$ számra létezik olyan $n=n(\varepsilon,\mathbf{A})\ge1$ index, és $\mathbf{B}=\mathbf{B}(\varepsilon)\in\mathcal{A}_n$ halmaz, amelyekre $\mu_j(\mathbf{A}\Delta\mathbf{B})\le\varepsilon$, minden $j=1,\dots,k$ számra, ahol $\mathbf{A}\Delta\mathbf{B}=(\mathbf{A}\setminus\mathbf{B})\cup(\mathbf{B}\setminus\mathbf{A})$.

14.) Legyen adva egy $(Y,\mathcal{B})$ mérhető tér és azon egymásba skatulyázott véges elemszámú particiók $\mathcal{A}_1\subset\mathcal{A}_2\subset\cdots$ sorozata. Az $(Y,\mathcal{B})$ tér $\mathcal{A}_l$ particióját azonosítsuk azzal a (véges elemszámú) σ-algebrával amely azokból a halmazokból áll, amelyek előállíthatóak, mint a $\mathcal{A}_l$ partició bizonyos elemeinek az uniója. Teljesüljön a $\mathcal{B} = \sigma\left(\bigcup\limits_{l=1}^{\infty} \mathcal{A}_l\right)$ feltétel. (Innen következik az is, hogy az $(Y, \mathcal{B})$ tér $\mathcal{B}$ σ-algebráját megszámlálható sok halmazzal lehet generálni.) Ekkor tetszőleges az $(Y, \mathcal{B})$ téren definiált μ és ν valószínűségi mértékpárra $\lim\limits_{l\to\infty} I(\nu_{\mathcal{A}_l}\|\mu_{\mathcal{A}_l}) = I(\nu\|\mu)$.

15.) Lássuk be a 14. feladat eredményének a segítségével a (+1) és (+2) állításokat tetszőleges F és G eloszlásfüggvényekre.

Be szeretnénk látni a 9. — 15. feladatok eredményeinek segítségével az e feladatsorban megfogalmazott nagy eltérés tétel felső becslését, (azaz, a $P\left(\frac{S_n}{n} \in [a,b]\right)$ valószínűségre akarunk éles alsó becslést adni). Ez tulajdonképpen a nagy eltérés tétel állításának nehéz fele. A gondolat a következő: Ha a $\xi_1, \dots, \xi_n$ valószínűségi változók empirikus eloszlásfüggvénye $F_n(x)$, összege S_n, akkor $\frac{S_n}{n} = \int u\, dF_n(u)$. Azt várjuk, hogy ez az átlag közel van az $\int u\, dG(u)$ integrálhoz, ha az F_n empirikus eloszlásfüggvény közel van a G eloszlásfüggvényhez. Másrészt ennek az eseménynek a valószínűségét jól tudjuk becsülni. Íly módon a G eloszlásfüggvény alkalmas választásával jó alsó becslést kapunk. E becslésben az $I(G\|F)$ I-divergencia jelenik meg, de a 10. feladat eredményének segítségével meg tudjuk adni a $\sup\limits_{t\ge 0}(tx - \log R(t))$ kifejezéstől függő alsó becslést a vizsgált valószínűségre.

Ez az érv alkalmazható, de csak bizonyos finomítással. A probléma a következő: Ha néhány ξ_k valószínűségi változó abszolút értéke rendkívül nagy, akkor ezek alig módosítják az F_n empirikus eloszlásfüggvényt, de nagyon befolyásolhatják az $\frac{S_n}{n}$ átlag értékét. E nehézség leküzdése érdekében olyan G eloszlásfüggvény kis környezetének a valószínűségére adunk alsó becslést, amely egy véges intervallumba koncentrált mértéket határoz meg. Látni fogjuk, hogy a G eloszlás választására tett megszorítás nem okoz komoly problémát. A következő feladatban bebizonyítjuk a kívánt állítást a fent vázolt érvelés részletesebb kidolgozásával.

16.) Legyenek $\xi_1(\omega), \xi_2(\omega), \dots$ független F eloszlású valószínűségi változók, és $G \neq F$ egy másik eloszlásfüggvény, amelyik egy véges intervallumba van koncentrálva, azaz létezik olyan $K > 0$ szám, amelyre $G(K) = 1$, $G(-K) = 0$. Lássuk be, hogy tetszőleges $\varepsilon > 0$ számra létezik olyan $L = L(\varepsilon) > 0$ szám, amelyre $L(\varepsilon) \to \infty$, ha $\varepsilon \to 0$, és az $\bar F(x) = \bar F_{L(\varepsilon)}(x) = P(\xi_1 < x \,|\, |\xi_1| \le L)$ feltételes eloszlásfüggvény teljesíti a $\sup\limits_x |F(x) - \bar F(x)| \le \varepsilon$ és $|IG\|F) - I(G\|\bar F)| \le \varepsilon$ tulajdonságokat. Továbbá van olyan $C(\eta, \varepsilon)$ függvény, amelyre $C(\eta, \varepsilon) \to 0$, ha $\eta \to 0$ minden $\varepsilon > 0$-ra, és

$$\begin{aligned} P\left(\omega\colon \sup_x |F_n(x,\omega) - G(x)| \le \eta \,\middle|\, |\xi_k(\omega)| \le L,\ 1 \le k \le n\right) &\ge e^{-nI(G\|\bar F) - nC(\eta,\varepsilon)} \\ &\ge e^{-n(I(G\|F) + C(\eta,\varepsilon) + \varepsilon)}, \end{aligned}$$

ha $I(G\|F) < \infty$. Lássuk be e tény segítségével a nagy eltérés tétel felső becslését, azaz azt az állítást, hogy $E\xi_1 \le a < b$ esetén $\limsup\limits_{n\to\infty} -\frac{1}{n} \log P\left(\frac{S_n}{n} \in [a,b]\right) \le$

$\sup_{t\ge 0}(ta - \log R(t))$. (Ha $E\xi_1$ nem létezik, akkor bármilyen $a < b$ számpárt tekint-hetünk.)

Meg kívánjuk fogalmazni és be akarjuk bizonyítani a feladatsor elején megfogalmazott nagy eltérés tételt független valószínűségi változók empirikus eloszlásfüggvényére. Az eloszlásfüggvények terében fogunk dolgozni. A nagy eltérés tétel megfogalmazásához be kell vezetni egy topológiát az eloszlások terében. Igyekszünk a nagy eltérés tételt minél gazdagabb topológiával ellátott térben bebizonyítani, mert így több esetben tudunk alsó vagy felső becslést adni az empirikus mértéknek egy előírt halmazba esésének a valószínűségére. Bár minket elsősorban valós értékű valószínűségi változók érdekelnek, nem jelent nehézséget, sőt egyszerűbb terminológiát tesz lehetővé, ha általánosabb mérhető térbeli értéket felvevő valószínűségi változókat tekintünk. Az alábbiakban definiált, úgynevezett τ-topológiát fogjuk tekinteni.

A τ-topológia definiciója. *Legyen adva egy $(Y,\mathcal{B})$ mérhető tér, és jelölje X az ezen a téren értelmezett (előjeles) mértékek halmazát vagy annak egy részhalmazát. A τ topológiát az X téren definiáljuk. A τ-topológiában a nyílt halmazokat a $\{\mu\colon\ \mu \in X,\ \mu(\mathbf{B}_1)\in \mathbf{G}_1,\dots,\mu(\mathbf{B}_k)\in\mathbf{G}_k\}$ alakú halmazok generálják, ahol k tetszőleges természetes szám, $\mathbf{B}_1,\dots,\mathbf{B}_k$ az $(Y,\mathcal{B})$ tér mérhető halmazai, és $\mathbf{G}_1,\dots,\mathbf{G}_k$ a számegyenes nyílt részhalmazai.*

Megjegyzés: Ugyanezen τ topológia nyílt halmazainak a bázisát generálják a $\{\mu\colon\ \mu \in X,\ \mu(\mathbf{B}_1)\in \mathbf{G}_1,\dots,\mu(\mathbf{B}_k)\in\mathbf{G}_k\}$ alakú halmazok, ahol $\{\mathbf{B}_1,\dots,\mathbf{B}_k\}$ az $(Y,\mathcal{B})$ tér véges particiói, és $\mathbf{G}_1,\dots,\mathbf{G}_k$ a számegyenes nyílt részhalmazai. Ezt a tényt fel fogjuk használni néhány bizonyításban.

Nagy eltérés tétel empirikus mértékekre. *Legyen adva egy $(Y,\mathcal{B})$ mérhető tér, azon egy μ valószínűségi mérték, és legyen $\xi_1(\omega),\xi_2(\omega),\dots,$ független, $(Y,\mathcal{B})$ értékű, μ eloszlású valószínűségi változók sorozata egy $(\Omega,\mathcal{D},P)$ valószínűségi mezőn, azaz legyen $P(\xi_n\in\mathbf{B})=\mu(\mathbf{B})$ minden $\mathbf{B}\in\mathcal{B}$ halmazra, $n=1,2,\dots$. Jelölje $(X,\mathcal{X})$ az $(Y,\mathcal{B})$ téren levő valószínűségi mértékek terét az $\mathcal{X}$-topológiával. Jelölje $\mathcal{A}$ az $\mathcal{X}$ topológia által generált σ-algebrát az X téren. A $(\xi_1(\omega),\dots,\xi_n(\omega))$ valószínűségi változók által meghatározott (véletlen) $\mu_n(\omega)$ empirikus mértéket a következőképp definiáljuk:*

$$\mu_n(\omega)(\mathbf{B})=\frac{1}{n}\#\{k\colon\ 1\le k\le n,\ \xi_k(\omega)\in\mathbf{B}\}\quad \text{minden } \mathbf{B}\in\mathcal{B}\text{-re.}$$

Ekkor $\mu_n(\omega)$, $\omega\in\Omega$, tekinthető úgy, mint egy $(X,\mathcal{X})$ térbeli értéket felvevő valószínűségi változó. Jelölje $\mu^{(n)}$ a $\mu_n(\omega)$ empirikus mérték eloszlását az $(X,\mathcal{X})$ téren, azaz legyen $\mu^{(n)}(\mathbf{A})=P(\omega\colon\ \mu_n(\omega)\in\mathbf{A})$, ahol $\mathbf{A}\subset\mathcal{A}$, mérhető halmaz az $(X,\mathcal{X})$ térben. A $\mu^{(n)}$, $n=1,2,\dots$, mértékek sorozata teljesíti a nagy eltérés tételt az $(X,\mathcal{X})$ topológikus téren az $I(\nu)=I(\nu\|\mu)$, $\nu\in X$, függvénnyel, ahol $I(\nu\|\mu)$ a ν mérték μ mérték szerinti I-divergenciáját jelöli.

Be fogjuk továbbá látni a nagy eltérés tétel i.), ii.) és iii.) tulajdonságát és ezenkívül azt, hogy az előbb definiált τ topológia teljesíti a T_3 tulajdonságot, amelynek fogalmát felidéztük a nagy eltérés tétel definiciójában.

Az előbb kimondott tételnek, amelyet az irodalomban Szanov tételnek hívnak, fenti megfogalmazása kissé pontatlan. Nem teljesül ugyanis feltétlenül az a tulajdonság, hogy a μ_n empirikus mérték valószínűségi változó, azaz mérhető függvény. Sőt, a τ topológia elég gazdag ahhoz, hogy a legtermészetesebb esetekben is, például amikor számegyenesen definiált valószínűségi változókat tekintünk, nem lehet az $(X, \mathcal{X})$ tér tetszőleges nyílt vagy zárt halmaza esetén annak valószínűségéről beszélni, hogy az empirikus mérték ebbe a halmazba esik. Erre a 26. feladatban példát is fogunk látni. Viszont a Szanov tétel érvényben marad, ha a nagy eltérés ii.) tulajdonságában (annak valószínűsége, hogy az empirikus mérték nyílt halmazba esik) belső mértékről a iii.) tulajdonságában pedig (annak valószínűsége, hogy az empirikus mérték zárt halmazba esik) külső mértékről beszélünk. Így a 19. feladatban $\mu^{(n)}$ belső mértéket, a 22., 23. és 24. feladatban pedig külső mértéket jelöl. Egy halmaz belső mértéke az általa tartalmazott mérhető halmazok mértékének a szuprémuma, külső mértéke pedig az őt tartalmazó halmazok mértékének az infimuma.

A következő ok miatt érdemes bevezetni a Szanov tétel megfogalmazásában külső és belső mértékeket. Jelölje $\mathcal{A}_0$ az X halmaz $\{\mu\colon\ \mu \in X,\ \mu(\mathbf{B}_1) \in \mathbf{G}_1, \dots, \mu(\mathbf{B}_k) \in \mathbf{G}_k\} \subset X$ alakú részhalmazai által generált σ-algebrát, ahol k tetszőleges természetes szám, $\mathbf{B}_1, \dots, \mathbf{B}_k$ az $(Y, \mathcal{B})$ tér mérhető halmazai, és $\mathbf{G}_1, \dots, \mathbf{G}_k$ a számegyenes nyílt részhalmazai. Be lehet látni, hogy tetszőleges $A \in \mathcal{A}_0$ halmazra és $n = 1, 2, \dots$ számra az $\{\omega\colon\ \mu_n(\omega) \in A\}$ alakú halmazok az $(\Omega, \mathcal{D}, P)$ valószínűségi mező $\mathcal{D}$ mérhető halmazai, ezért létezik valószínűségük. A Szanov tétel azt állítja, hogy az $\mathcal{A}_0$ σ-algebrában lévő zárt és nyílt halmazokra érvényesek a nagy eltérés tétel becslései a tételben definiált $I(\cdot)$ függvénnyel. De az $(X, \mathcal{A})$ térben definiált $\mathcal{A}$ σ-algebrát nem csak az $\mathcal{A}_0$ σ-algebra definiciójában szereplő halmazok, hanem az összes nyílt halmaz generálja, ezért ez a σ-algebra sokkal nagyobb, mint $\mathcal{A}_0$. A 26. feladatban, — kihasználva, hogy kontinum sok nyílt halmaz uniója is nyílt — példát mutatunk olyan G nyílt halmazra az $(X, \mathcal{X})$ térben, amely nem eleme az $\mathcal{A}_0$ σ-algebrának, és nem lehet definiálni a $\mu_n(G) = P(\omega\colon\ \mu_n(\omega) \in G)$ valószínűséget. De lehet definiálni a $\mu_n(G)$ belső és $\mu_n(F)$ külső mértékeket nyílt G, illetve zárt F halmazokra, és a Szanov tétel érvényben marad ilyen halmazokra is, ha a fent említett módon belső és külső mértékekkel dolgozunk.

17.) Legyen $(X, \mathcal{X})$ egy $(Y, \mathcal{B})$ mérhető téren definiált valószínűségi mértékek tere a τ topológiával. Bizonyítsuk be, hogy ez a topologikus tér T_3 tér, azaz tetszőleges $\mu \in X$ pontra és zárt $\mathbf{F} \subset X$ halmazra, amelyre $\mu \notin \mathbf{F}$, létezik két diszjunkt $\mathbf{G}_1$ és $\mathbf{G}_2$ nyílt halmaz, amelyekre $\mu \in \mathbf{G}_1$ és $\mathbf{F} \subset \mathbf{G}_2$. Továbbá az $(X, \mathcal{X})$ tér teljesíti a következő T_1 tulajdonságot is: Minden egy elemű, $\{\mu\}$, $\mu \in X$ halmaz zárt.

18.) Legyen $(Y, \mathcal{B})$ mérhető tér, $(X, \mathcal{X})$ pedig az $(Y, \mathcal{B})$ téren definiált valószínűségi mértékek tere a τ topológiával ellátva. Rögzítsünk egy $\mu \in X$ valószínűségi mértéket. Ekkor az $I(\nu\|\mu)$ I-divergencia mint a $\nu \in X$ változó függvénye alulról félig folytonos.

Tekintsük azt a speciális esetet, amikor $(Y, \mathcal{B}) = (R^1, \mathcal{B})$, a számegyenes a Borel σ-algebrával. Vegyük az $(R^1, \mathcal{B})$ téren definiált valószínűségi mértékek X halmazán azt a (τ-topológiánál durvább) τ_0 topológiát, amelynek bázisát alkotják a $\{\mu\colon\ \mu \in X,\ \mu(\mathbf{B}_1) \in \mathbf{G}_1, \dots, \mu(\mathbf{B}_k) \in \mathbf{G}_k\}$ alakú halmazok, ahol $\mathbf{B}_1, \dots, \mathbf{B}_k$ szakaszok vagy félegyenesek a számegyenesen, és $\mathbf{G}_1, \ldots \mathbf{G}_k$ nyílt halmazok. Az $I(\nu\|\mu)$ I-divergencia mint a $\nu \in X$ változó függvénye egy rögzített μ mértékkal a τ_0 topológia szerint is alulról félig folytonos.

19.) Legyen adva egy $(Y, \mathcal{B})$ mérhető tér, és legyen $(X, \mathcal{A})$ az $(Y, \mathcal{B})$ téren definiált valószínűségi mértékek tere a τ topológiával illetve az általa definiált σ-algebrával. Rögzítsünk egy $\mu \in X$ valószínűségi mértéket, és legyenek $\xi_1, \xi_2, \ldots,$ független egyforma eloszlású $(Y, \mathcal{B})$ értékű valószínűségi változók μ eloszlással. Jelölje $\mu^{(n)}$, $n = 1, 2, \ldots,$ a $\xi_1(\omega), \ldots, \xi_n(\omega)$ valószínűségi változók empirikus eloszlásának (mint $(X, \mathcal{A})$ értékű valószínűségi változónak) az eloszlását az $(X, \mathcal{A})$ téren. Ekkor a $\mu^{(n)}$ mértékek teljesítik a nagy eltérés tétel ii.) tulajdonságát is az $I(\cdot\|\mu)$ I-divergenciával, azaz tetszőleges nyílt $\mathbf{G} \subset X$ halmazra

$$\limsup_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{G}) \le \inf_{\nu \in \mathbf{G}} I(\nu\|\mu).$$

A nagy eltérés utolsó feltételének a bizonyítása, — zárt halmazok valószínűségének a felső becslése, — finomabb érvelést igényel. A problémát az okozza, hogy hiába tudunk a ,,nagyon nem tipikus" pontok kis környezeteinek a valószínűségére jó becslést adni, ezek uniójának a valószínűségére az ilyen becslések önmagukban nem elegendőek. Természetes gondolat valamilyen jó approximációval és véges fedéssel megpróbálni ezt a problémát megoldani. Ez a módszer akkor működik, ha a becsülendő halmaz nemcsak zárt, hanem kompakt is. A 20. feladatban elégséges feltételt adunk arra, hogy egy zárt halmaz kompakt legyen. Ezután bebizonyítjuk a nagy eltérés tétel iii.) tulajdonságát kompakt halmazokra, illetve megmutatjuk, hogy általános zárt halmazok valószínűségének a becslése visszavezethető egy minimax tételre. Végül ezt a minimax tételt bizonyítjuk be.

20.) Legyen $(Y, \mathcal{B})$ mérhető tér, $(X, \mathcal{X})$ az $(Y, \mathcal{B})$ téren definiált valószínűségi mértékek tere a τ topológiával. Legyen $\mu \in X$ tetszőleges valószínűségi mérték, $R > 0$ pozitív valós szám. Ekkor a $\{\nu\colon\ \nu \in X,\ I(\nu\|\mu) \le R\}$ halmaz kompakt.

21.) Az $(X, \mathcal{X})$ tér a τ topológiával T_ρ tér. (E fogalom definicióját felidéztük a 3. feladat definiciójában.)

A 21. feladat állítására nem lesz a továbbiakban szükségünk. E feladat célja az volt, hogy megmutassuk: esetünkben is alkalmazható a 3. feladat eredménye. A bizonyítás néhány alapvető topológiai eredményből és egy az előző feladatban is alkalmazott konstrukcióból következik.

A következő feladatban a nagy eltérés tétel iii) tulajdonságának egy gyengített változatát bizonyítjuk be. Ebben zárt halmazok helyett csak kompakt halmazokat tekintünk. Először egy általánosabb állítást bizonyítunk be arról, hogy ez a reláció teljesül $\mathbf{Q}_n$ mértékek egy sorozatára az $(X, \mathcal{A})$ téren egy $I(\nu)$, $\nu \in X$, függvénnyel alkalmas feltételek teljesülés esetén. Aztán megmutatjuk, hogy ezek a feltételek teljesülnek a mi esetünkben is.

22.) Legyen $(Y, \mathcal{B})$ egy mérhető tér, $(X, \mathcal{X})$ az $(Y, \mathcal{B})$ téren értelmezett valószínűségi mértékek tere a τ topológiával, $\mathcal{A}$ az $\mathcal{X}$ topológia által generált σ-algebra az X téren. Legyen $\mathbf{Q}_n$, $n = 1, 2, \dots$, valószínűségi mértékek egy sorozata az $(X, \mathcal{A})$ téren, és $I(\cdot)$ alkalmasan választott nem negatív függvény $(X, \mathcal{X})$-en, amely a $+\infty$ értéket is felveheti. Tegyük fel, hogy minden $\nu \in X$ mértékre és $\varepsilon > 0$ számra létezik a ν pontnak olyan $\mathbf{G}(\nu, \varepsilon)$ nyílt környezete, amelyre $\liminf\limits_{n\to\infty} -\frac{1}{n} \log \mathbf{Q}_n(\mathbf{G}(\nu, \varepsilon)) > I(\nu) - \varepsilon$, ha $I(\nu) < \infty$ és $\liminf\limits_{n\to\infty} -\frac{1}{n} \log \mathbf{Q}_n(\mathbf{G}(\nu, \varepsilon)) > \frac{1}{\varepsilon}$, ha $I(\nu) = \infty$. Ekkor a $\mathbf{Q}_n$ mértéksorozat és az $I(\cdot)$ függvény teljesíti a nagy eltérés tétel iii.) tulajdonságának következő gyengített alakját:

iii′.) $\liminf\limits_{n\to\infty} -\frac{1}{n} \log \mathbf{Q}_n(\mathbf{K}) \ge \inf\limits_{\nu \in \mathbf{K}} I(\nu)$ minden kompakt $\mathbf{K} \subset X$ halmazra.

Legyen $\mathbf{Q}_n$ n független μ eloszlású valószinűségi változóból készített $\mu_n = \mu_n(\omega)$ empirikus mérték $\mu^{(n)}$ eloszlása az $(X, \mathcal{A})$ téren, és $I(\cdot) = I(\cdot \| \mu)$. Ezzel a választással alkalmazhatjuk a feladat első részének az eredményét, mert teljesülnek az ott előírt feltételek.

A nagy eltérés tétel bizonyítása zárt, de nem kompakt halmazokra nehezebb. Ez a probléma a nagy eltérés tétel bizonyításának egyik legnehezebb pontja az általános (nem feltétlenül független minta empirikus eloszlását vizsgáló) esetben is. A következő két feladat célja ennek a problémának a megoldása. Az első feladat egy viszonylag egyszerű, természetes módon bizonyítható becslést ad a minket érdeklő kifejezésre alkalmas "diszkretizáció" segítségével. Közvetlenül nem annak a valószínűségét vizsgáljuk, hogy a tekintett empirikus eloszlás beleesik-e egy (esetleg bonyolult) zárt halmazba. Ehelyett az empirikus mérték vetületeit tekintjük véges particiókra, és azt becsüljük, hogy ezek a vetületek milyen valószínűséggel esnek bizonyos a feladat jellegéből kifolyólag természetes módon definiált halmazokba. Ez lényegesen egyszerűbb feladat, (valójában a 12. feladat eredményét kell alkalmaznunk), és ez az eljárás vezet a 23. feladatban megadott becsléshez. Annak megmutatása, hogy az így kapott becslés lehetővé teszi a nagy eltérés tételben szereplő iii.) tulajdonság bizonyítását lényegesen nehezebb probléma. Ez egy minimax jellegű eredményből következik, amelyiket a 24. feladat megoldásában bizonyítunk be.

23.) Legyen $(Y, \mathcal{B})$ mérhető tér, $(X, \mathcal{X})$ az $(Y, \mathcal{B})$ téren értelmezett valószínűségi mértékek tere a τ topológiával, $\mathcal{A}$ az ezen topológia által generált σ-algebra az X téren. Legyen $\mu \in X$ valószínűségi mérték, $\mu^{(n)}$ e mérték n-szeres direkt szorzata önmagával, $\mu_n = \mu_n(\omega) \in X$, $n = 1, 2, \dots$, n darab független μ eloszlású valószínűségi változó empirikus eloszlása, $\mathbf{F} \subset X$ az $(X, \mathcal{A})$ tér zárt részhalmaza. Ekkor

$$\liminf_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mu_n \in \mathbf{F}) \ge \sup_{\mathcal{C} \in \mathcal{C}^*} \inf_{\nu_{\mathcal{C}} \in [\mathbf{F}_{\mathcal{C}}]} I\left(\nu_{\mathcal{C}} \| \mu_{\mathcal{C}}\right),$$

ahol $\mathcal{C}^*$ jelöli az $(Y, \mathcal{B})$ tér összes $\mathcal{C} = \{\mathbf{C}_1, \dots, \mathbf{C}_k\}$ véges (mérhető) particióját. A $\nu_{\mathcal{C}}$ illetve a $\mu_{\mathcal{C}}$ mérték egy $\nu \in X$ illetve a kezdetben rögzített $\mu \in X$ mérték és egy $\mathcal{C} = \{\mathbf{C}_1, \dots, \mathbf{C}_k\} \in \mathcal{C}^*$ partició esetén az a mérték (egy véges $\{x_1 \dots, x_k\}$ halmazon), amelyre $\nu_{\mathcal{C}}(\{x_j\}) = \nu(\mathbf{C}_j)$, illetve $\mu_{\mathcal{C}}(\{x_j\}) = \mu(\mathbf{C}_j)$, és $x_j \in \mathbf{C}_j$,

$j = 1, \dots, k$, egy a $\mathcal{C}$ particióhoz rögzített pontrendszer. Egy $\mathbf{F} \subset X$ halmaz $\mathbf{F}_{\mathcal{C}}$ vetülete egy $\mathcal{C}$ particióra a $\nu_{\mathcal{C}}$ mértékekekből álló halmaz az $Y_{\mathcal{C}} = \{x_1, \dots, x_k\}$ (véges) halmazon, ahol $\nu \in \mathbf{F}$, és $x_j \in \mathbf{C}_j$, $j = 1, \dots, k$, a $\mathcal{C} = \{\mathbf{C}_1, \dots, \mathbf{C}_k\}$ halmazhoz kijelölt pontrendszer. Az $[\mathbf{F}_{\mathcal{C}}]$ halmaz az $\mathbf{F}_{\mathcal{C}}$ halmaz lezártja az $(X_{\mathcal{C}}, \mathcal{X}_{\mathcal{C}})$ téren. Az $(X_{\mathcal{C}}, \mathcal{X}_{\mathcal{C}})$ tér az $(Y_{\mathcal{C}}, \mathcal{B}_{\mathcal{C}})$, ($\mathcal{B}_{\mathcal{C}}$ jelöli a diszkrét topológiát) mérhető téren értelmezett valószínűségi mértékek tere a τ topológiával.

24.) Az előző feladat jelölésével

$$\sup_{\mathcal{C} \in \mathcal{C}^*} \inf_{\nu_{\mathcal{C}} \in [\mathbf{F}_{\mathcal{C}}]} I\left(\nu_{\mathcal{C}} \| \mu_{\mathcal{C}}\right) \geq \inf_{\nu \in \mathbf{F}} \sup_{\mathcal{C} \in \mathcal{C}^*} I\left(\nu_{\mathcal{C}} \| \mu_{\mathcal{C}}\right) = \inf_{\nu \in \mathbf{F}} I(\nu \| \mu). \tag{4}$$

Bizonyítsuk be e formula segítségével a Szanov tételben megfogalmazott nagy eltérés tétel iii.) tulajdonságát.

Az előző feladatokban egy zárt $\mathbf{F}$ halmaz $\mathbf{F}_{\mathcal{C}}$ vetületét vettük egy $\mathcal{C}$ particióra, majd annak $[\mathbf{F}_{\mathcal{C}}]$ lezártját tekintettük. Bár erre a bizonyítás helyességének igazolásához nincs szükség, érdemes példát látni arra, hogy ez a lezárás valódi operáció, azaz lehetséges, hogy $\mathbf{F}$ zárt, de $\mathbf{F}_{\mathcal{C}}$ nem zárt halmaz.

25.) Konstruáljunk $(Y, \mathcal{B})$ mérhető teret, az $(Y, \mathcal{B})$ téren értelmezett valószínűségi mértékek $(X, \mathcal{X})$ terén a τ topológiával egy zárt $\mathbf{F}$ halmazt és egy $\mathcal{C}$ véges particiót az $(Y, \mathcal{B})$ téren, amelyre az $\mathbf{F}_{\mathcal{C}}$ halmaz nem zárt. (A 23. feladatban bevezetett fogalmakat használjuk.) Mutassuk meg, hogy a következő példa teljesíti ezeket a feltételeket: $(Y, \mathcal{B})$ a $[0, 1]$ intervallum a Borel σ-algebrával, $\mathbf{F} = \{\nu_k,\ k = 2, 3, \dots\}$, ahol a ν_k mérték a következő: $\nu_k\left(\left\{\frac{1}{k}\right\}\right) = 1 - \frac{1}{k}$, $\nu_k\left(\left\{\frac{3}{4}\right\}\right) = \frac{1}{k}$; $\mathcal{C} = \{\mathbf{C}_1, \mathbf{C}_2\}$, $\mathbf{C}_1 = \left[0, \frac{1}{2}\right]$, $\mathbf{C}_2 = \left(\frac{1}{2}, 1\right]$ választással.

26.) Legyen $(Y, \mathcal{B})$ mérhető tér, $(X, \mathcal{X})$ az ezen a téren értelmezett valószínűségi mértékek tere a τ topológiával. Legyenek $\xi_1(\omega), \dots, \xi_n(\omega)$ független, egyforma eloszlású Y értékű valószínűségi változók egy $(\Omega, \mathcal{D}, P)$ valószínűség mezőn. Ha

$$\{\xi_1(\omega), \dots, \xi_n(\omega)\} = \{y_1, \dots, y_n\} \in Y^n,$$

akkor jelölje $\mu_n(\omega) = \mu_n(y_1, \dots, y_n)$ ezeknek a valószínűségi változóknak az empirikus eloszlásfüggvényét az $\omega \in \Omega$ pontban. Tetszőleges $\mathbf{B} \subset Y$ (nem feltétlenül mérhető) halmaz és n pozitív egész szám esetén lehet definiálni olyan $\mathbf{G} = \mathbf{G}(\mathbf{B}, n)$ nyílt halmazt az $(X, \mathcal{X})$ térben, amelyre az $\{\omega\colon \mu_n(\omega) \in \mathbf{G}\}$ és az

$$\{\omega\colon \xi_k(\omega) \in \mathbf{B},\ \text{minden } k = 1, \dots, n\}$$

események megegyeznek. Ennek az észrevételnek a segítségével mutassunk példát arra, hogy az $\{\omega\colon \mu_n(\omega) \in \mathbf{G}\}$ eseménynek nem minden (a τ topológia szerint) nyílt $\mathbf{G}$ halmazra van valószínűsége.

27.) Bizonyítsuk be a 15. feladat állítását (a (+1) és (+2) relációkat) a Szanov tétel segítségével.

Megoldások.

1.) Tekintsünk először egy olyan $x \in X$ pontot, amelyre $I(x) < \infty$. Az $I(\cdot)$ alulról félig folytonossága miatt $\sup\limits_{\mathbf{G}\colon x\in\mathbf{G},\ \mathbf{G}\text{ nyílt}} \inf\limits_{y\in\mathbf{G}} I(y) = I(x)$. Továbbá azt állítom, hogy mivel a topológia T_3 tipusú, ezért $\sup\limits_{\mathbf{F}\colon x\in\operatorname{Int}\mathbf{F},\ \mathbf{F}\text{zárt}} \inf\limits_{y\in\mathbf{F}} I(y) = I(x)$. Valóban, tetszőleges $\varepsilon > 0$ számhoz létezik olyan nyílt $\mathbf{G}$, $x \in \mathbf{G}$, halmaz, amelyre $\inf\limits_{y\in\mathbf{G}} I(y) \ge I(x) - \varepsilon$. Továbbá a topológia T_3 tulajdonsága miatt létezik olyan zárt $\mathbf{F}_0$ halmaz, amelyre $x \in \operatorname{Int}\mathbf{F}_0 \subset \mathbf{F}_0 \subset \mathbf{G}$. Valóban, a T_3 tulajdonság miatt léteznek olyan diszjunkt, nyílt $\mathbf{G}_1$ és $\mathbf{G}_2$ halmazok, amelyekre $x \in \mathbf{G}_1$, és $X \setminus \mathbf{G} \subset \mathbf{G}_2$. Ekkor $\mathbf{F}_0 = X \setminus \mathbf{G}_2$ teljesíti a kívánt tulajdonságokat. Ezért $\sup\limits_{\mathbf{F}\colon x\in\operatorname{Int}\mathbf{F},\ \mathbf{F}\text{ zárt}} \inf\limits_{y\in\mathbf{F}} I(y) \ge \inf\limits_{y\in\mathbf{F}_0} I(x) \ge I(x) - \varepsilon$. A másik írányú becslés nyílvánvaló.

Ezekből a relációkból és a nagy eltérés tételből következik, hogy

$$\sup_{\substack{\mathbf{G}\colon \mathbf{G}\text{ nyílt halmaz}\\ x\in\mathbf{G}}} \limsup_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{G}) \le I(x)$$

$$\sup_{\substack{\mathbf{F}\colon \mathbf{F}\text{ zárt halmaz}\\ x\in\operatorname{Int}\mathbf{F}}} \liminf_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{F}) \ge I(x).$$

Rögzítsünk egy kis $\varepsilon > 0$ számot. Legyen $\mathbf{F}$, $x \in \operatorname{Int}\mathbf{F}$, olyan zárt halmaz, amelyre $\liminf\limits_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{F}) \ge I(x) - \varepsilon$. Ekkor egy $x \in \mathbf{G} \subset \operatorname{int}\mathbf{F}$ halmazra $\limsup\limits_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{G}) \ge \liminf\limits_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{F}) \ge I(x)-\varepsilon$. Mivel ez az állítás tetszőleges $\varepsilon > 0$ számra és alkalmas $\mathbf{G}$ nyílt halmazra igaz, a fenti becslésekből következik a bizonyítandó állítások első fele. Az állítás másik fele hasonlóan bizonyítható. Létezik olyan $\mathbf{G}$, $x \in \mathbf{G}$ nyílt halmaz, amelyre $\limsup\limits_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{G}) \le I(x) + \varepsilon$. Legyen $\mathbf{F}$ a $\mathbf{G}$ halmaz lezártja. Ekkor $x \in \operatorname{Int}\mathbf{F}$, és

$$\liminf_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{F}) \le \limsup_{n\to\infty} -\frac{1}{n}\log\mu_n(\mathbf{F}) \le I(x) + \varepsilon.$$

Innen következik a feladat másik állítása.

Az $I(x) = \infty$ eset az előző esethez hasonlóan néhány természetes változtatással bizonyítható.

2.) Lássuk be először az alsó becslést. Rögzített $\varepsilon > 0$ számhoz válasszunk olyan $x_0 \in X$ pontot, amelyre $g(x_0) - I(x_0) > \sup\limits_{x\in X}\{g(x) - I(x)\} - \varepsilon$. Mivel a $g(x)$ függvény folytonos, létezik az x_0 pontnak olyan $\mathbf{G} \subset X$ nyílt környezete, amelyre $g(x) > g(x_0) - \varepsilon$, ha $x \in \mathbf{G}$. Ezen egyenőtlenségből, az $x_0 \in \mathbf{G}$ relációból és a nagy eltérés tétel definiciójának ii) relációjából következik, hogy minden $\varepsilon > 0$ számhoz létezik olyan $n_0 = n_0(\varepsilon)$ küszöbindex, amelyre $\mu_n(\mathbf{G}) > \sup\limits_{x\in\mathbf{G}} e^{-n(I(x)-\varepsilon)} >$

$e^{-n(I(x_0)-\varepsilon)}$, és

$$\int e^{ng(x)}\,d\mu_n(x) \ge \int_{\mathbf{G}} e^{ng(x)}\,d\mu_n(x) \ge e^{n(g(x_0)-\varepsilon)}\mu_n(\mathbf{G})$$
$$\ge e^{n(g(x_0)-I(x_0)-2\varepsilon)} \ge \sup_{x\in X} e^{n(g(x)-I(x)-3\varepsilon)},$$

ha $n \ge n_0(\varepsilon)$. Mivel $\varepsilon > 0$ tetszőleges, innen következik, hogy

$$\liminf \frac{1}{n}\log \int e^{ng(x)}\,d\mu_n(x) \ge \sup_{x\in X}\{g(x)-I(x)\}.$$

A felső becslés bizonyítása érdekében rögzítsünk egy elég kis $\varepsilon > 0$ számot, és definiáljuk az $\mathbf{F}(k) = \mathbf{F}(k,\varepsilon)$

$$\mathbf{F}(k) = \{x\colon\, x \in X,\ k\varepsilon \le g(x) \le (k+1)\varepsilon\}$$

halmazokat, $-\frac{1}{\varepsilon}\sup|g(x)| \le k < \frac{1}{\varepsilon}\sup|g(x)|$ indexekre. Az $\mathbf{F}(k)$ halmazok zártak, és $x \in \mathbf{F}_k$ esetén $-I(x) = g(x) - I(x) - g(x) \le \sup\limits_{x\in X}\{g(x)-I(x)\} - k\varepsilon$. Ezért elég nagy n-re

$$\int_{\mathbf{F}(k)} e^{ng(x)}\,d\mu_n(x) \le e^{n(k+1)\varepsilon}\mu_n(\mathbf{F}(k))$$
$$\le \exp\left\{n\left[(k+1)\varepsilon + \sup_{x\in X}\{g(x)-I(x)\} - k\varepsilon + \varepsilon\right]\right\}$$
$$= \exp\left\{n\left[\sup_{x\in X}\{g(x)-I(x)\} + 2\varepsilon\right]\right\}.$$

Ezeket az egyenlőtlenségeket összeadva az összes (véges sok) k-ra, és kihasználva, hogy $\varepsilon > 0$ tetszőleges, megkapjuk a kívánt felső becslést.

3.) A ii.) tulajdonság belátása érdekében egy $\mathbf{G} \subset X$ nyílt halmaz esetében tetszőleges $\varepsilon > 0$-hoz tekintsünk olyan $x \in \mathbf{G}$ pontot, amelyre $I(x) < \inf\limits_{u\in\mathbf{G}} I(u) + \varepsilon$, és legyen $g(\cdot)$ olyan folytonos függvény az X téren, amelyre $g(x) = 0$, $g(u) = -\frac{1}{\varepsilon}$, ha $u \notin \mathbf{G}$, és $g(u) \le 0$ minden $u \in X$-re. Ilyen $g(\cdot)$ függvény valóban létezik, mert véve egy $h(\cdot)$ függvényt, amelyre $h(x) = 0$, $h(u) \ge 0$ minden $u \in X$ pontban, $h(u) = 1$, ha $u \notin \mathbf{G}$, akkor definiáljuk a $h_1(u) = \min\{h(u), 1\}$, $g(u) = -\frac{1}{\varepsilon}h_1(u)$ függvényeket. Ez a $g(\cdot)$ függvény teljesíti a kívánt feltételeket. Ekkor

$$\mu_n(\mathbf{G}) \ge \int_X e^{ng(u)}\mu_n(\,du) - \int_{X\setminus\mathbf{G}} e^{ng(u)}\mu_n(\,du) \ge \int_X e^{ng(u)}\mu_n(\,du) - e^{-n/\varepsilon},$$

ezért

$$-\frac{1}{n}\log\mu_n(\mathbf{G}) \le -\frac{1}{n}\log\left[\int e^{ng(u)}\,\mu_n(du) - e^{-n/\varepsilon}\right].$$

Legyen $A = \inf_{u\in G} I(u)$. Ha $A=\infty$, akkor $\limsup -\frac{1}{n}\log\mu_n(\mathbf{G}) \le \infty = \inf_{x\in G} I(x)$. Ha $A<\infty$, akkor

$$\int e^{ng(u)}\,\mu_n(du) \ge \exp\left\{n\left(\sup_{u\in\mathbf{G}}(g(u)-I(u))-\varepsilon\right)\right\} \ge e^{n(g(x)-I(x)-\varepsilon)} \ge e^{-n(A+2\varepsilon)}$$

elég nagy n-re. Ezért

$$\begin{aligned}\limsup -\frac{1}{n}\log\mu_n(\mathbf{G}) &\le \limsup -\frac{1}{n}\log\{e^{-n(A+2\varepsilon)}-e^{n/\varepsilon}\}\\ &\le \limsup -\frac{1}{n}\log e^{-n(A+2\varepsilon)} + \limsup -\frac{1}{n}\log(1-e^{-n/\varepsilon+n(A+2\varepsilon)}) = A+2\varepsilon\end{aligned}$$

minden elég kis $\varepsilon>0$ számra. Innen $\limsup -\frac{1}{n}\log\mu_n(\mathbf{G}) \le A = \inf_{x\in G} I(x)$.

A iii$'$.) tulajdonság bizonyítása érdekében minden $x\in\mathbf{K}$ ponthoz és $\varepsilon>0$ számhoz definiálunk egy alkalmas $g_{x,\varepsilon}=g_{x,\varepsilon}(u)$ folytonos és korlátos függvényt az $(X,\mathcal{X})$ téren. Minden $x\in\mathbf{K}$ pontra definiáljuk az $\mathbf{F}(x)=\{z\colon\ I(z)\le I(x)-\varepsilon\}$, ha $I(x)<\infty$, és $\mathbf{F}(x)=\{z\colon\ I(z)\le\varepsilon^{-1}\}$, ha $I(x)=\infty$ halmazokat. Mivel az $I(\cdot)$ függvény alulról félig folytonos, ezért az $\mathbf{F}(x)$ halmazok zártak. Ezért létezik olyan folytonos $g_{x,\varepsilon}$ függvény az $(X,\ \mathcal{X})$ téren, amelyre $g_{x,\varepsilon}(x)=\varepsilon$, $g_{x,\varepsilon}(u)=-I(x)$ az $u\in\mathbf{F}(x)$ pontokban, ha $I(x)<\infty$, és az $I(x)=\infty$ esetben $g_{x,\varepsilon}(x)=\varepsilon$, $g_{x,\varepsilon}(u)=-\frac{2}{\varepsilon}$, ha $u\in\mathbf{F}(x)$. Továbbá $g_{x,\varepsilon}(u)\le\varepsilon$ minden $u\in X$-re. (Ilyen függvények konstrukciója azért lehetséges, mert az $(X,\mathcal{X})$ tér T_ρ tér.) Definiáljuk az x pont $\mathbf{U}(x)=\mathbf{U}(x,\varepsilon)=\{z\colon\ g_{x,\varepsilon}(z)>0\}$ nyílt környezetét. Mivel az $\mathbf{U}(x)$ halmazok a $\mathbf{K}$ kompakt halmaz fedését adják, létezik véges sok $x_1,\dots,x_k$ pont, amelyekre a $\mathbf{U}(x_j)$, $j=1,\dots,k$ halmazok lefedik a $\mathbf{K}$ halmazt. Legyen $g_j(u)=g_{x_j,\varepsilon}(u)$, $j=1,\dots,k$. Ekkor

$$\mu_n(\mathbf{K}) \le \sum_{j=1}^k \mu_n(\mathbf{U}(x_j)) \le \sum_{j=1}^k \int_{\mathbf{U}(x_j)} e^{ng_j(u)}\,\mu_n(du) \le \sum_{j=1}^k \int e^{ng_j(u)}\,\mu_n(du).$$

Továbbá elég nagy n-re

$$\begin{aligned}\int e^{ng_j(u)}\,\mu_n(du) &\le \exp\left\{n\left(\sup_{u\in X}[g_j(u)-I(u)]+\varepsilon\right)\right\}\\ &\le e^{n(3\varepsilon-I(x_j))}\quad \text{ha } I(x_j)<\infty,\end{aligned}$$

mert $u\in\mathbf{F}(x_j)$-re $g_j(u)-I(u)\le -I(x_j)-I(u)\le -I(x_j)$, és $u\notin\mathbf{F}(x_j)$-re $g_j(u)-I(u)\le g_j(u)-[I(x_j)-\varepsilon]\le 2\varepsilon-I(x_j)$. Hasonlóan,

$$\int e^{ng_j(u)}\,\mu_n(du) \le e^{n(2\varepsilon-1/\varepsilon)},\quad \text{ha } I(x_j)=\infty.$$

Ezért

$$\frac{1}{n}\log\mu_n(\mathbf{K}) \le \frac{\text{const.}}{n} - \min\left(\frac{1}{\varepsilon},\ \min_{I(x_j)<\infty} I(x_j)\right)+3\varepsilon \le -\min\left(\frac{1}{\varepsilon},\ \inf_{u\in\mathbf{K}} I(u)\right)+4\varepsilon$$

elég nagy n-re. Innen következik a iii$'$.) tulajdonság.

4.) Az $I(x)$ függvény definiciójából következik, hogy az monoton nő. Továbbá, ha $I(z) < \infty$, $E\xi < z$, akkor minden kis $\varepsilon > 0$ számra létezik olyan $t > 0$ szám, amelyre $I(z) < tz - \log R(t) + \varepsilon$, és létezik egy ε-tól független $\delta = \delta(z)$ szám úgy, hogy az előző egyenlőtlenséget kielégítő (ε, t) párra $t > \delta$. Valóban, mivel $z > E\xi_1$, $I(z) > \eta$ alkalmas $\eta > 0$ számra, és ha $\varepsilon < \frac{\eta}{2}$, akkor a $tz - \log R(t) > \frac{\eta}{2}$ feltételnek teljesülnie kell. De mivel $\lim\limits_{t\to 0}(tz - \log R(t)) = 0$ ez csak úgy lehetséges, ha $t > \delta$ alkalmas $\delta = \delta(z)$ számmal. Adva egy (y, x) pár, amelyre $E\xi < y < x$ válasszunk egy $\varepsilon < (x-y)\delta$, $\delta = \delta(y)$ számot, amelyre létezik olyan $t > \delta$ szám, hogy $I(y) < ty - \log R(t) + \varepsilon$. Ekkor $I(y) < tx - \log R(t) - t(x-y) + \varepsilon \le I(x) + \varepsilon - \delta(x-y) < I(x)$. Az $I(\cdot)$ függvény baloldali folytonosságának bizonyítása érdekében egy olyan x számra, amelyre $I(x) < \infty$, $x > E\xi$ egy elég kis $\varepsilon > 0$-ra válasszunk olyan t számot, melyre $I(x) < tx - \log R(t) + \varepsilon$. Legyen y olyan szám, amelyre $y < x$ és $x - y < \frac{\varepsilon}{t}$. Ekkor $I(x) \ge I(y) \ge t(y-x) + tx - \log R(t) \ge I(x) - \varepsilon - t(x-y) \ge I(x) - 2\varepsilon$. Innen következik a kívánt folytonosság. Végül, mivel a $tx - \log R(t)$ függvény minden rögzített t-re konvex (sőt lineáris) függvény, ezért $I(x) = \sup\limits_{t\ge 0}(tx - \log R(t))$ is konvex függvény. Innen következik az $I(x)$ függvény jobbról való folytonossága az adott feltételek mellett. Ha ugyanis $I(x) < I(x+0) < \infty$ lenne valamely $x \ge E\xi$-re, akkor a szigorú $\lim\limits_{\varepsilon\to 0} I(x+\varepsilon) > \lim\limits_{\varepsilon\to 0} \frac{I(x)+I(x+2\varepsilon)}{2}$ egyenlőtlenség teljesülne, és ez ellentmond az $I(\cdot)$ függvény konkvexitásának.

5.) A nagy eltérés tétel i.) tulajdonsága, az hogy az $I(\cdot)$ függvény alulról félig folytonos közvetlenül kiolvasható az előző feladat eredményéből.

Ha G nyílt halmaz, válasszunk tetszőleges $\varepsilon > 0$-ra olyan $z \in G$ pontot, amelyre $I(z) < \inf\limits_{x\in G} I(x) + \varepsilon$, és legyen $h > 0$ olyan szám, amelyre $[z, z+h] \in G$ ha $z \ge E\xi_1$, és $[z-h, z] \in G$, ha $z < E\xi_1$. Ekkor

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in G\right) \ge \limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in [z, z \pm h]\right) = I(z).$$

Az utolsó azonosság belátásához elég bebizonyítani azt, hogy $I(z) < I(z+h)$, ha $z \ge E\xi_1$ és $I(z) < \infty$, és $I(z) < I(z-h)$, ha $z < E\xi_1$ és $I(z) < \infty$. Ezt az állítást viszont az $x > E\xi_1$ beláttuk a 4. feladatban. Az $x < E\xi_1$ hasonlóan bizonyítható, vagy következik az előző esetből, azt alkalmazva a $-\xi_1$ valószínűségi változóra. Ezért igaz a nagy eltérés tételben szereplő ii.) tulajdonság.

Egy F zárt halmazt írjunk fel $F = F_1 \cup F_2$ alakban, ahol $F_1 = F \cap \{x \ge E\xi_1\}$, $F_2 = F \cap \{x \le E\xi_1\}$.

A iii.) állítást elég belátni az F_1 és F_2 halmazokra. Legyen $z = \min\{x\colon x \in F_1\}$. Ekkor a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor 19. feladata alapján

$$\liminf_{n\to\infty} -\frac{1}{n} \log P\left(\frac{S_n}{n} \in F_1\right) \ge \liminf_{n\to\infty} -\frac{1}{n} \log P\left(S_n \ge z\right) = I(z).$$

Viszont $I(z) = \inf_{x\in F} I(x)$, mert az $I(x)$ függvény monoton nő.

6.) Tekintsük először csak azt az esetet, amikor a ξ_1 valószínűségi változónak létezik várható értéke. Először azt látjuk be, hogy az $I(x)$ függvény alulról félig folytonos. Ha $Ee^{t\xi} < \infty$ alkalmas $t > 0$-val, akkor a 4. feladat eredményéből következik ez az állítás $x \ge E\xi_1$-re. Továbbá $\lim_{x\to E\xi_1} I(x) = 0$. Alkalmazva ezt az eredményt a $-\xi_1$ valószínűségi változóra, kapjuk az analóg állítást $x \le E\xi_1$ esetén, ha $Ee^{t\xi_1} < \infty$ alkalmas $t < 0$-ra. Mivel a többi esetben $I(x) = 0$, ezekből az állításokból következik a nagy eltérés tétel i.) tulajdonsága.

Tekintsük a ii.) tulajdonságot először olyan G nyílt halmazokra, amelyek tartalmazzák a (véges) $E\xi_1$ várható értéket. Ebben az esetben a nagy számok törvénye alapján $P\left(\frac{S_n}{n} \in G\right) \to 1$, ezért a ii.) tulajdonság teljesül. Ha $E\xi \notin G$, akkor $\varepsilon > 0$-ra létezik olyan $u \in G$ és $[a,b] \subset G$ intervallum, amelyre $I(u) \le \inf_{x\in G} I(x) + \varepsilon$, $u \in [a,b]$ és $[a,b] \subset (E\xi_1, \infty)$ vagy $[a,b] \subset (-\infty, E\xi_1)$. Elég az első esetet vizsgálni, mert a második eset a $-\xi_k$ valószínűségi változók bevezetésével visszavezethető erre az esetre. Ha $Ee^{t\xi_1} < \infty$ alkalmas $t > 0$-ra, akkor ez következik az $I(x)$ függvény 4. feladatban belátott szigorú monotonításából. Ha pedig ez a tulajdonság nem teljesül, akkor a a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor 15. feladatának eredménye alapján $\frac{1}{n} \log P\left(\frac{S_n}{n} \in [a,b]\right) \to 0$, ha $n \to \infty$, és innen következik az állítás az adott esetben.

Hasonlóan az előző feladat érveléséhez, elég a iii.) tulajdonságot abban az esetben belátni, ha az F zárt halmaz teljesíti az $F \subset [E\xi_1, \infty)$ vagy az analóg módon tárgyalható $F \subset (-\infty, E\xi_1]$ tulajdonságot. Ha $Ee^{t\xi_1} < \infty$ alkalmas $t > 0$-ra, akkor ez az állítás az előző feladat bizonyításából kiolvasható, viszont ha ez a tulajdonság nem teljesül, akkor az állítás nyílvánvaló, mert ekkor $I(x) = 0$ minden $x \ge E\xi_1$-re.

Végül, ha a ξ_1 valószínűségi változónak nincs várható értéke, akkor $I(x) = 0$ minden x-re. Ezért a i.) és iii.) tulajdonság ebben az esetben nyílvánvaló. A ii.) tulajdonság pedig következik a *Nagy eltérések elmélete; Független. valós értékű valószínűségi változók* feladatsor 15. feladatának eredményéből.

7.) Az adott feltételek mellett a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor 8. feladata alapján $[\log R(0)]' = -\infty$. Továbbá a $\log R(t)$ konvexitása és folytonossága valamint a $\log R(0) = 0$ reláció miatt minden $\varepsilon > 0$-ra létezik olyan $t_0 = \delta(\varepsilon)$.) szám, hogy $\log R(t) > -\varepsilon$, ha $0 \le t_0 \le t_0$, és $\log R(t) \ge \log R(t_0) + (t - t_0)[\log R(t_0)]' \ge -\varepsilon - K(\varepsilon)t$, ha $t \ge t_0$ alkalmas $K(\varepsilon)$ számmal. Ezért, ha $x < -K(\varepsilon)$, akkor $tx - \log R(t) < \varepsilon$ minden $t \ge 0$ esetén. Mivel ez az egyenlőtlenség igaz minden $\varepsilon > 0$-ra, és $\log R(0) = 0$, így $t = 0$-ra $tx - \log R(t) = \log R(0)$ minden x számra innen következik a feladat állítása.

8.) A 6. feladat eredményének felhasználása után azt kell belátni, hogy

$$\liminf_{n\to\infty} -\frac{1}{n} \log \mu_n(G) \ge \inf_{x\in G} I(x)$$

minden nyílt G halmazra, ahol μ_n egy ξ_k, $k = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változókból álló sorozat első n eleméből készített átlag eloszlása, és

$I(\cdot)$ a μ_n, $n=1,2,\dots$, mértéksorozatra érvényes nagy eltérés tétel $I(\cdot)$ függvénye. Ha $E\xi_1\in G$, akkor a bizonyítandó egyenlőtlenség érvényes, mert ekkor $\inf\limits_{x\in G} I(x)=0$. Ezért elég azt a két esetet nézni, amikor $G\in(-\infty,E\xi_1)$ vagy $G\in(\xi_1,\infty)$. Mivel az $I(\cdot)$ függvény mind a két félegyenesen monoton függvény, ezért mind a két esetben teljesülnek a $G\subset G'=(x',x'')$ és $\inf\limits_{x\in G} I(x)=\inf\limits_{x\in G'} I(x)$ relációk, ahol $x'=\inf\{x\colon\ x\in G\}$ és $x''=\sup\{x\colon\ x\in G\}$. Ezért a kívánt állítást elég ilyen G' nyílt intervallumra belátni (esetleg $\pm\infty$ végpontokkal). Ez viszont kiolvasható a nagy eltérés tétel alakjából független egyforma eloszlású valószínűségi változók összegére.

Definiáljuk az ellenpéldát a következő módon:

$$\mu_n(0)=2^{-1},\qquad \mu_n\left(j2^{-n^2}\right)=2^{-(n+1)},\quad j=1,2\dots,2^n,$$

és legyen $I(0)=0$, és $I(x)=\infty$, ha $x\neq 0$. Ekkor $I(x)$ alulról félig folytonos függvény, tehát teljesül az i.) reláció. A iii.) reláció is teljesül, mert egy zárt F halmazra vagy $0\in F$, amely esetben $\limsup\limits_{n\to\infty} -\frac1n\log\mu_n(F)=0$ vagy F nem tartalmazza a 0 egy kis környezetét, és ekkor $\mu_n(F)=0$ elég nagy n-re. A ii) reláció teljesül tetszőleges (nem feltétlenül nyílt) G halmazra, mert vagy $0\in G$, amely esetben $\mu_n(G)\ge\frac12$ minden n-re vagy $0\notin G$, és ekkor $\inf\limits_{x\in G} I(x)=\infty$. Végül a ii.) relációban nem lehet azonosságot írni, mert a nyílt $G=R^1-\{0\}$ halmazra $\mu_n(G)=\frac12$, tehát $\liminf\limits_{n\to\infty} -\frac1n\log\mu_n(G)=0$.

9.) Mivel $\log(1+v)\le v$, $0\le v\le 1$ esetben, és egyenlőség csak $v=0$ esetén van,

$$\begin{aligned}I(\nu\|\mu)&=-\int\log\frac{d\mu}{d\nu}(u)\,d\nu(u)\ge-\int\frac{d(\mu-\nu)}{d\nu}(u)\,d\nu(u)\\&=\int d\mu(u)-\int d\nu(u)=0,\end{aligned}$$

ha ν abszolút folytonos a μ mértékre nézve, és egyenlőség csak $\nu=\mu$ esetben áll fenn. Ha ν nem abszolút folytonos, akkor ez az állítás nyílvánvaló.

Ha az $R(t)=\int e^{tu}\,dF(u)$ integrál véges, akkor definiáljuk az $F_t(du)=\frac{e^{tu}F(du)}{R(t)}$ valószínűségi mértéket. Ekkor $\int u\,dG(u)=x$ esetén

$$I(G\|F)=\int\log\frac{dF_t(u)}{dF(u)}\,dG(u)+\int\log\frac{dG(u)}{dF_t(u)}\,dG(u)=tx-\log R(t)+I(G\|F_t),$$

ahonnan $I(G\|F)\ge tx-\log R(t)$, és egyenlőség csak $G=F_t$ esetén áll fenn. Innen következik, hogy a bizonyítandó (a) azonosság baloldala mindig nagyobb vagy egyenlő mint a jobboldal. Továbbá, mivel $\int u\,dF_t(u)=[\log R(t)]'$ a $[\log R(t)]'=x$ megoldhatósága esetén mind a bal mind a jobboldal felveszi a szélsőértéket a feladatban leírt esetben és a két kifejezés egyenlő.

10.) Be kell látni az (a) azonosság még hiányzó felét, azaz az $\inf\limits_{G\colon \int u\,dG(u)\ge x} I(G\|F) \le \sup\limits_{t\ge 0}(tx-\log R(t))+\varepsilon$ egyenlőtlenséget minden $\varepsilon>0$ számra abban az esetben, ha a $[\log R(t)]'=x$ egyenlet nem oldható meg. Az $R(t)$ momentumgeneráló függvény vizsgálatában, (a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor (8., 12., 14. és 15. feladatainak) eredményeiből következik, hogy a következő eseteket kell vizsgálni: a.) Az $R(t)$ momentumgeneráló függvény egy véges jobbról zárt $[0,T]$ intervallumban van definiálva, és $[\log R(T)]'=m_T\le x$, b.) Az $R(t)$ függvény értelmezve van az egész $[0,\infty]$ félegyenesen, de $\lim\limits_{t\to\infty}[\log R(t)]'=m\le x$. c.) Minden $t>0$ esetén $R(t)=\infty$.

Az a.) esetben az (a) azonosság jobboldala $t=T$ pontban veszi fel a szuprémumát. Elég belátni, hogy ha $x>m_T$, akkor tetszőleges $\varepsilon>0$-ra létezik olyan G eloszlás, amelyre $x+\varepsilon\ge\int u\,dG(u)\ge x$, és $I(G\|F_T)<\varepsilon$, ahol $m_T=\int u\,dF_T(u)$. Ekkor ugyanis $I(G\|F)=T\int u\,dG(u)-\log R(T)+I(G\|F_T)\le T(x+\varepsilon)-\log R(T)+\varepsilon$. Teljesül továbbá az $\int e^{tu}dF_T(u)=\infty$ feltétel minden $t>0$ esetén, és a G eloszlásfüggvény konstrukciójában az F_T eloszlásnak ezt a tulajdonságát fogjuk használni. Ezért az a.) és c.) esetet, ahol ugyanezt az egyenlőtlenséget kell bizonyítani $T=0$ választással, egyszerre vizsgálhatjuk. Feltehetjük, hogy a szigorú egyenlőtlenség $x>m_T$, illetve $x>m$ teljesül, ($m_T=\int u\,dF_T(u)$, $m=\int u\,dF(u)$,) mert $x=m_T$, illetve $x=m$ esetén $G=F_T$ és $G=F$ kielégíti a kívánt egyenlőtlenséget. A konstrukció megadása érdekében az $x>m_T$ és $x>m$ esetben vegyük észre, hogy az a.) esetben minden elég kis $\varepsilon'>0$ számra léteznek olyan $K_1>0$ és $K_2>0$ számok, amelyekre $\mu_1=\mu([-K_1,K_2])=F_T(K_2)-F_T(-K_1)>1-\varepsilon'$, $\int_{-K_1}^{K_2}u\,dF_T(u)<m_T<x\mu_1$, és hasonló állítás érvényes a c.) esetben is. Valóban, $K_1>0$-t és $K_2>0$-t elég nagynak választva teljesíthetjük az első feltételt, és mivel $1-F_T(K_2)>0$, a $K_1>0$ konstansot még nagyobbra választva, ha szükséges az $\int u\,dF_T(u)=m_T<x$ feltétel miatt a második feltételt is teljesíteni tudjuk. Továbbá, mivel az F_T eloszláshoz tartozó $R_T(t)$ momentumgeneráló függvény végtelen tetszőleges $t>0$ esetén, tetszőleges $\delta>0$ számhoz végtelen sok pozitív p egész szám van, amelyekre, $\mu_2=\mu_2^{(p)}=F_T(2^{p+1})-F_T(2^p)>e^{-\delta 2^p}$. Hasonló állítás érvényes a c.) esetben is.

Definiáljuk a $\bar\mu_1$ és $\bar\mu_2$ valószínűségi mértékeket, mint

$$\bar\mu_1(\mathbf{A})=\frac{1}{\mu_1}\int_{\mathbf{A}\cap[-K_1,K_2]}dF_T(u),$$
$$\bar\mu_2(\mathbf{A})=\bar\mu_2^{(p)}(\mathbf{A})=\frac{1}{\mu_2^{(p)}}\int_{\mathbf{A}\cap[2^p,2^{p+1}]}dF_T(u),$$

ahol μ_1, K_1, K_2 $\mu_2^{(p)}$ az előző megjegyzésben szereplő mennyiségek, és a p paramétert a $\bar\mu_2^{(p)}$ mérték definiciójában alkalmasan választjuk. Definiáljuk a $G=G^{(p)}$ mértéket, mint az $G=\alpha\bar\mu_1+(1-\alpha)\bar\mu_2^{(p)}$ lineáris kombinációt, ahol az α számot az $\int u\,dG(u)=x$ feltétel határozza meg. Azt állítjuk, hogy elég kis $\varepsilon'>0$, $\delta=$

$\delta(\varepsilon') > 0$ és elég nagy $p = p(\delta) > 0$ választás esetén G valószínűségi mérték, és $I(G\|F_T) < \varepsilon$. Ennek bizonyításához vegyük észre, hogy

$$I(G\|F_T) = \alpha I(\bar\mu_1\|F_T) + (1-\alpha)I(\bar\mu_2\|F_T) = -\alpha\log\mu_1 - (1-\alpha)\log\mu_2^{(p)},$$

és mivel $\int u\,d\bar\mu_1(u) < x$, és $2^p \le \int u\,d\bar\mu_2^{(p)}(u) \le 2^{(p+1)}$, elég nagy p-re $0 < 1-\alpha <$ const. 2^{-p}. Innen $-\alpha\log\mu_1 < -\log\mu_1 < -\log(1-\varepsilon')$, és $-(1-\alpha)\log\mu_2^{(p)} < \delta 2^p\text{const.}\,2^{-p} = \text{const.}\,\delta$. Innen következik, hogy az ε', δ és p alkalmas választásával $I(G\|F_T) < \varepsilon$ elérhető. Az $I(G\|F) < \varepsilon$ egyenlőtlenség alkalmas konstrukcióval a c.) esetben hasonlóan bizonyítható.

A b.) eset könnyen következik a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor 13. feladatának eredményéből. Elég azt az esetet nézni, ha $x = \lim\limits_{t\to\infty}[\log R(t)]'$, mert $x > \lim\limits_{t\to\infty}[\log R(t)]'$ esetén $\sup\limits_{t\ge 0}(tx - \log R(t)) = \infty$. A vizsgálandó esetben válasszuk a G mértéket, mint az x pontba koncentrált mértéket. Ekkor $I(G\|F) = -\log\mu_F(x)$, ahol μ_F jelöli az F mérték által indukált mértéket, azaz a jobboldalon az x pont μ_F mértéke van. Az említett feladatsor 13. feladatának eredménye alapján ez megegyezik az (a) formula baloldalán szereplő szuprémummal.

A feladat utolsó állítását, azt hogy $I(G\|F) < \sup\limits_{t\ge 0}(tx - \log R(t)) + \varepsilon$ elérhető szigorú egyenlőtlenséggel, ha az x számot kissé növeljük, kiolvasható a konstrukciók ellenőrzésével. De ez következik a 4. feladat eredményéből és az utána következő megjegyzésből is. Ezek szerint ugyanis $I(x) < \infty$ a tekintett esetben, és az $I(x)$ függvény folytonossági tulajdonságai alapján megkapjuk a kívánt eredményt, ha az ebben a feladatban bebizonyított állításokat x helyett $x+\varepsilon$-ra alkalmazzuk elég kis $\varepsilon > 0$ számmal.

11.) Tegyük fel először, hogy $q_l > 0$ minden $1 \le l \le k$-ra. Rögzítsünk egy $x_{j_1}, \dots, x_{j_n}$ sorozatot, $x_{j_s} \in X$, $1 \le s \le n$, és legyen $m(l) = m(l, x_{j_1}, \dots, x_{j_n}) = \#\{s\colon 1 \le s \le n,\ x_{j_s} = x_l\}$, $1 \le l \le k$. Ekkor $\chi_n(l) = \chi_n(l, x_{j_1}, \dots, x_{j_n}) = \frac{m(l)}{n}$, és

$$\mu^{(n)}(x_{j_1}, \dots, x_{j_n}) = \prod_{s=1}^n p_{j_s} = \prod_{l=1}^k p_l^{m(l)} = \prod_{l=1}^k \frac{p_l^{m(l)}}{q^{m(l)}} \cdot \prod_{l=1}^k q_l^{m(l)}.$$

Jelölje $\nu^{(n)}$ a ν mérték n-szeres direkt szorzatát önmagával az $\{x_1, \dots, x_k\}^n$ szorzattéren. Ekkor összegezve az utolsó formulát az összes olyan $x_{j_1}, \dots, x_{j_n}$ sorozatra, amelyre $n_n\chi(l) = m(l)$, $l = 1, \dots, k$, valamilyen előírt $m(l)$, $l = 1, \dots, k$, számokkal, azt kapjuk, hogy

$$\frac{\mu^{(n)}(n\chi_n(1) = m(1), \dots, n\chi_n(k) = m(k))}{\nu^{(n)}(n\chi_n(1) = m(1), \dots, n\chi_n(k) = m(k))} = \prod_{l=1}^k \left(\frac{p_l}{q_l}\right)^{m(l)}.$$

Ezért,

$$\prod_{l=1}^k \frac{p_l^{n(q_l+\varepsilon)}}{q_l^{n(q_l-\varepsilon)}}$$

$$\le \frac{\mu^{(n)}(((q_1-\varepsilon)<\chi_n(1)<(q_1+\varepsilon),\dots,(q_k-\varepsilon)<\chi_n(k)<(q_k+\varepsilon))}{\nu^{(n)}((q_1-\varepsilon)<\chi_n(1)<(q_1+\varepsilon),\dots,(q_k-\varepsilon)<\chi_n(k)<(q_k+\varepsilon))}$$
$$\le \prod_{l=1}^{k} \frac{p_l^{n(q_l-\varepsilon)}}{q_l^{n(q-l+\varepsilon)}}.$$

Mivel a nagy számok törvénye alapján

$$\frac{1}{2} \le \nu^{(n)}\left((q_1-\varepsilon)<\chi_n(1)<(q_1+\varepsilon),\dots,(q_k-\varepsilon)<\chi_n(k)<(q_k+\varepsilon)\right) \le 1$$

elég nagy n-re, az előző becslésből következik, hogy

$$\begin{aligned}
&\frac{1}{2}\prod_{l=1}^{k}\frac{p_l^{\varepsilon n}}{q_l^{-\varepsilon n}}\prod_{l=1}^{k}\left(\frac{p_l}{q_l}\right)^{nq_l}\\
&\quad= \mu^{(n)}\left(|\chi_n(l)-q_l|<\varepsilon \text{ minden } 1\le l\le k \text{ indexre}\right)\\
&\quad= \mu^{(n)}\left((q_1-\varepsilon)<\chi_n(1)<(q_1+\varepsilon),\dots,(q_k-\varepsilon)<\chi_n(k)<(q_k+\varepsilon)\right)\\
&\qquad\qquad \le \prod_{l=1}^{k}\frac{p_l^{-\varepsilon n}}{q_l^{\varepsilon n}}\prod_{l=1}^{k}\left(\frac{p_l}{q_l}\right)^{nq_l}
\end{aligned}$$

elég nagy n-re. Mivel $\prod\limits_{l=1}^{k}\left(\frac{p_l}{q_l}\right)^{nq_l} = \exp\left\{-n\sum\limits_{l=1}^{k} q_l \log\frac{q_l}{p_l}\right\}$, a fenti egyenlőtlenségből következik a feladatban megfogalmazott (3) formula, ha $q_l>0$ minden l indexre.

Ha $q_l=0$ valamilyen l-re, akkor a fenti becslés végrehajtható azzal a módosítással, hogy csak az $l\in K$, $K\subset\{1,\dots k\}$, indexeket vesszünk figyelembe a fenti számolásban, ahol $l\in K$ akkor és csak akkor, ha $q_l>0$. A (3) egyenlőtlenségsorozat bal és jobboldala nem változik, ha az $l\in\{1,\dots,k\}$ helyett $l\in K$-ra összegezünk. A középen álló kifejezés értéke nő, ha az $1\le l\le k$ indexek helyett csak az $l\in K$ indexeket tekintjük. Másrészt $\mu^{(n)}(|\chi_n(y,l)-q_l|<\frac{\varepsilon}{k}$ minden $1\le l\le k$ indexre$)\le$ $\mu^{(n)}(|\chi_n(y,l)-q_l|<\varepsilon$ minden $l\in K$ indexre), mert a $\sum\limits_{l\in\{1,\dots,k\}\setminus K} q_l=1$ és $q_l=0$, ha $l\in K$ relációkból következik, hogy a baloldalon olyan halmaz mértékét vettük, amely része a jobboldalon tekintett halmaznak. Innen következik a (3) reláció.

Rátérek az (1) és (2) relációk igazolására. Ha F és G két véges sok pontba koncentrált eloszlás eloszlásfüggvénye, akkor legyen $X=\{x_1,\dots,x_k\}\subset R^1$ az a legkisebb halmaz, amelyre mind az F által meghatározott μ_F mind a G által meghatározott μ_G mértékre $\mu_F(X)=1$, $\mu_G(X)=1$. Legyen $\mu_F(\{x_l\})=p_l$, $\nu_l(\{x_l\})=q_l$, $1\le l\le k$. Ha létezik olyan l index, amelyre $p_l=0$ (és $q_l>0$), akkor $I(G\|F)=\infty$ és elég kis $\varepsilon>0$ számra $P(\sup\limits_x|F_n(x)-G(x)|<\varepsilon)=0$, mert $P(\sup\limits_x|F_n(x)-G(x)|<\varepsilon)\le P(F_n(x+0)-F_n(x-0)-q>-2\varepsilon)$. Ha $p_l>0$ minden $1\le l\le k$-ra, akkor vegyük észre, hogy a feladatban bizonyított egyenlőtlenség bal és jobboldalán $e^{-n(I(G\|F)\pm C(\varepsilon))}$ szerepel. Ezért

$$\begin{aligned}
e^{-n(I(G\|F)+C(\varepsilon/k))} &\leq P\left(|\chi_n(l)-q_l|<\frac{\varepsilon}{k},\ \text{minden } 1\leq l\leq k\text{-ra}\right)\\
&\leq P\left(\sup_{-\infty<x<\infty}|F_n(x)-G(x)|<\varepsilon\right)\\
&\leq P\left(|\chi_n(l)-q_l|<2\varepsilon,\ \text{minden } 1\leq l\leq k\text{-ra}\right)\\
&\leq e^{-n(I(G\|F)-C(2\varepsilon))}
\end{aligned}$$

elég nagy n-re. Innen következik a (+1) és (+2) állítás.

12.) A kívánt egyenlőtlenség bizonyításához elég belátni, hogy minden $r=1,\dots,k$-ra

$$\int_{\mathbf{C}_r}\log\frac{d\nu}{d\mu}(u)\,d\nu(u)\geq\log\frac{\nu(\mathbf{C}_r)}{\mu(\mathbf{C}_r)}\nu(\mathbf{C}_r),$$

Összegezve ugyanis ezeket az egyenlőtlenségeket $k=1,\dots,r$-re megkapjuk a kívánt egyenlőtlenséget. Viszont az előbb felírt egyenlőtlenség a valószínűségi mértékek alkalmas választásával következik a 9. feladatban bebizonyított $I(\nu\|\mu)\geq 0$ egyenlőtlenségből. Bevezetve ugyanis a $\mu_r(\mathbf{B})=\frac{\mu(\mathbf{C}_r\cap\mathbf{B})}{\mu(\mathbf{C}_r)}$ és $\nu_r(\mathbf{B})=\frac{\nu(\mathbf{C}_r\cap\mathbf{B})}{\nu(\mathbf{C}_r)}$ mértékeket, a bizonyítandó egyenlőtlenséget megkapjuk a 9. feladatban bebizonyított $I(\nu\|\mu)\geq 0$ állításból, ha ν-t ν_r-rel és μ-t μ_r-rel helyettesítjük. Ugyanis $\log\frac{d\nu_r}{d\mu_r}=\log\frac{d\nu}{d\mu}-\log\frac{\nu(\mathbf{C}_r)}{\mu(\mathbf{C}_r)}$ a $\mathbf{C}_r$ halmazon, és integrálva ezt az azonosságot a ν_r mérték szerint azt kapjuk, hogy $0\leq\int\log\frac{d\nu_r}{d\mu_r}\,d\nu_r=\frac{1}{\nu(\mathbf{C}_r)}\int\log\frac{d\nu}{d\mu}\,d\nu-\log\frac{\nu(\mathbf{C}_r)}{\mu(\mathbf{C}_r)}$, és ez ekvivalens a bizonyítandó állítással.

Ha ν nem abszolút folytonos a μ mértékre nézve, tekintsük az $\mathcal{C}=(\mathbf{C}_1,\mathbf{C}_2)$, $\mathbf{C}_1=\left\{u\colon\ \frac{d\mu}{d\nu}(u)=0\right\}$, $\mathbf{C}_2=Y\setminus\mathbf{C}_1$ particiót. Ekkor $\nu_{\mathcal{C}}$ sem abszolút folytonos a $\mu_{\mathcal{C}}$ mértékre, és $I(\nu_{\mathcal{C}}\|\mu_{\mathcal{C}})=\infty$. Tekintsük ezután azt az esetet, amikor $I(\nu\|\mu)<\infty$. Ekkor valamilyen $K_2>1>K_1>0$ és N paraméterek segítségével definiáljuk a következő $\mathcal{C}=\{\mathbf{C}_1,\dots,\mathbf{C}_{N+2}\}$ particiót:

$$\mathbf{C}_l=\left\{u\colon\ K_1+\frac{(K_2-K_1)(l-1)}{N}\leq\frac{d\nu}{d\mu}(u)<K_1+\frac{(K_2-K_1)l}{N}\right\},\quad l=1,\dots,N$$

$\mathbf{C}_{N+1}=\left\{u\colon\ \frac{d\nu}{d\mu}(u)\geq K_2\right\}$, $\mathbf{C}_{N+2}=\left\{u\colon\ \frac{d\nu}{d\mu}(u)<K_1\right\}$. (Pontosabban fogalmazva, azokat a $\mathbf{C}_l$ halmazokat vesszük, amelyekre $\mu(\mathbf{C}_l)>0$.) Megmutatjuk, hogy alkalmas K_1, K_2 és N választással $|I(\nu\|\mu)-I(\nu_{\mathcal{C}}\|\mu_{\mathcal{C}})|\leq\varepsilon$.

Ha $K_2>1$-t elég nagynak, $1>K_1>0$-t elég kicsinek választjuk az $I(\nu\|\mu)$ végessége miatt elérhetjük, hogy $\left|\int_{\mathbf{C}_{N+i}}\log\frac{d\nu}{d\mu}(u)\nu(\,du)\right|\leq\frac{\varepsilon}{4}$ legyen, $i=1,2$. Vegyük észre, hogy ha a $\mathbf{C}\subset Y$ halmaz olyan, hogy $a\leq\frac{d\nu}{d\mu}(u)\leq b$ minden $u\in\mathbf{C}$ pontban, $-\infty\leq a<b\leq\infty$, akkor $a\leq\frac{\nu(\mathbf{C})}{\mu(\mathbf{C})}\leq b$. Valóban, mivel $\nu(\mathbf{C})=\int_{\mathbf{C}}\frac{d\nu}{d\mu}(u)\,d\mu(u)$, $\int_{\mathbf{C}}a\,d\mu(u)\leq\nu(\mathbf{C})\leq\int_{\mathbf{C}}b\,d\mu(u)$, ahonnan következik ez az állítás.

A feladat elején szereplő egyenlőtlenség alapján

$$\frac{\varepsilon}{4} \geq \int_{\mathbf{C}_{N+1}} \log \frac{d\nu}{d\mu}(u)\nu(\,du) \;\geq \log \frac{\nu(\mathbf{C}_{N+1})}{\mu(\mathbf{C}_{N+1})}\nu(\mathbf{C}_{N+1}) \geq 0,$$

továbbá

$$\frac{\varepsilon}{4} \geq -\int_{\mathbf{C}_{N+2}} \log \frac{d\nu}{d\mu}(u)\nu(\,du) \geq 0,$$

és mivel $\lim\limits_{u\to 0} u \log u = 0$, és $\log \frac{\nu(\mathbf{C}_{N+2})}{\mu(\mathbf{C}_{N+2})} < 0$, ezért

$$\begin{aligned} 0 \geq \nu(\mathbf{C}_{N+2}) \log \frac{\nu(\mathbf{C}_{N+2})}{\mu(\mathbf{C}_{N+2})} &= \mu(\mathbf{C}_{N+2}) \frac{\nu(\mathbf{C}_{N+2})}{\mu(\mathbf{C}_{N+2})} \log \frac{\nu(\mathbf{C}_{N+2})}{\mu(\mathbf{C}_{N+2})} \\ &\geq \frac{\nu(\mathbf{C}_{N+2})}{\mu(\mathbf{C}_{N+2})} \log \frac{\nu(\mathbf{C}_{N+2})}{\mu(\mathbf{C}_{N+2})} \geq -\frac{\varepsilon}{4}, \end{aligned}$$

ha $K_1 > 0$-t elég kicsinek, és $K_2 > 0$-t elég nagynak választjuk. Ezért a bizonyítás befejezéséhez ebben az esetben elég azt megmutatni, hogy elég nagy $N = N(K_1, K_2)$ számra

$$\sum_{l=1}^{N} \int_{\mathbf{C}_l} \left| \log \frac{d\mu}{d\nu}(u) - \log \frac{\mu(\mathbf{C}_l)}{\nu(\mathbf{C}_l)} \right| d\nu(u) \leq \frac{\varepsilon}{2}.$$

Ez viszont következik abból a tényből, hogy elég nagy N számra

$$\left| \log \frac{d\mu}{d\nu}(u) - \log \frac{\mu(\mathbf{C}_l)}{\nu(\mathbf{C}_l)} \right| \leq \frac{\varepsilon}{2}, \quad \text{ha } u \in \mathbf{C}_l.$$

Végül tekintsük azt az esetet, amikor $I(\nu\|\mu) = \infty$, de a ν mérték abszolút folytonos a μ mértékre. Ez az eset hasonlóan tárgyalható az előző esethez. Mivel $I(\nu\|\mu) = \infty$ választhatunk olyan $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ halmazokat, $\mathbf{C}_1 = \left\{u\colon \frac{d\nu}{d\mu}(u) < K_1\right\}$, $\mathbf{C}_2 = \left\{u\colon K_1 \leq \frac{d\nu}{d\mu}(u) \leq K_2\right\}$, $\mathbf{C}_3 = \left\{u\colon \frac{d\nu}{d\mu}(u) > K_2\right\}$ alkalmas $K_i = K_i(\varepsilon)$-nal, $i = 1, 2$, amelyekre $\nu(\mathbf{C}_1) \geq \frac{1}{2}$, $\int_{\mathbf{C}_2} \frac{d\nu}{d\mu}(u)\, d\nu(u) \geq \frac{2}{\varepsilon}$, és $\log \frac{d\nu}{d\mu}(u) > 1$, ha $u \in \mathbf{C}_3$. Ekkor $\int_{\mathbf{C}_1 \cup \mathbf{C}_3} \log \frac{d\nu}{d\mu}(u)\, d\nu(u) > -\log 2$. Az $I(\nu\|\mu) < \infty$ esetben tárgyalt konstrukció érveléséhez hasonlóan kapjuk, hogy a $\mathbf{C}_2$ halmaznak létezik olyan $\mathbf{C}_2 = \bigcup\limits_{l=1}^{N} \mathbf{B}_l$ particiója, amelyre

$$\sum_{l=1}^{N} \log \frac{\nu(\mathbf{B}_l)}{\mu(\mathbf{B}_l)} \nu(\mathbf{B}_l) \geq \frac{1}{\varepsilon} + \log 2.$$

Ezért az $\mathcal{C} = \{\mathbf{C}_1, \mathbf{C}_3, \mathbf{B}_1, \dots, \mathbf{B}_N\}$ rendszer az $(Y, \mathcal{B})$ tér olyan particiója, amelyre $I\left(\nu_{\mathcal{C}} \| \mu_{\mathcal{C}}\right) \geq \frac{1}{\varepsilon}$. Innen következik a feladat állítása ebben az esetben is.

13.) Tekintsük az összes olyan $\mathbf{A}$ halmazt, amelyre teljesül a fenti approximációs állítás. Nyílván az $\mathcal{A}^0 = \bigcup\limits_{l=1}^{\infty} \mathcal{A}_l$ unió elemei ilyen halmazok. Továbbá $\mathcal{A}^0$ halmazalgebra. Ezért elég azt belátni, hogy az adott tulajdonságú halmazok σ-algebrát alkotnak. Ezt az állítást közvetlenül is lehet ellenőrízni. Például, ha az $\mathbf{A}_l$, $l = 1, 2, \dots$, halmazok mindegyike teljesíti ezt a tulajdonságot, akkor az $\mathbf{A} = \bigcup\limits_{l=1}^{\infty} \mathbf{A}_l$ unió is teljesíti azt. Ugyanis, minden $\varepsilon > 0$-hoz létezik olyan $N = N(\varepsilon)$, amelyre $\mu_j\left(\bigcup\limits_{l=1}^{\infty} \mathbf{A}_l \setminus \bigcup\limits_{l=1}^{N} \mathbf{A}_l\right) \le \frac{\varepsilon}{2}$, $j = 1, \dots, k$, és minden $\mathbf{A}_l$, $l = 1, \dots, N$ halmazhoz létezik olyan $\mathbf{B}_l \in \mathcal{A}^0$ halmaz, amelyre $\mu_j(\mathbf{A}_l \Delta \mathbf{B}_l) \le \frac{\varepsilon}{2N}$, $j = 1, \dots, k$. Ekkor a $\mathbf{B} = \bigcup\limits_{l=1}^{N} \mathbf{B}_l \in \mathcal{A}^0$ halmazra $\mu_j(\mathbf{A} \Delta \mathbf{B}) \le \varepsilon$ minden $j = 1, \dots, k$-ra. Továbbá az is nyilvánvaló, hogy ha egy $\mathbf{A}$ halmazra teljesül az approximációs tulajdonság, akkor az $X \setminus \mathbf{A}$ halmazra is teljesül.

14.) Tekintsük először azt az esetet, amikor a ν mérték abszolút folytonos a μ mértékre nézve. A 12. feladat eredményei alapján tetszőleges $\varepsilon > 0$-ra létezik olyan $\mathcal{B}$ véges partició, amelyre

$$I(\nu\|\mu) \ge I(\nu_{\mathcal{B}}\|\mu_{\mathcal{B}}) > I(\nu\|\mu) - \varepsilon, \qquad \text{ha } I(\nu\|\mu) < \infty,$$

és

$$I(\nu_{\mathcal{B}}\|\mu_{\mathcal{B}}) \ge \frac{1}{\varepsilon}, \qquad \text{ha } I(\nu\|\mu) = \infty.$$

Továbbá az $I(\nu_{\mathcal{A}_n}\|\mu_{\mathcal{A}_n})$ sorozat monoton nő, és a limesze kisebb vagy egyenlő mint $I(\nu\|\mu)$. Ezért elég belátni, hogy elég nagy n-re és egy az Y halmaz egy véges mérhető $\mathcal{B} = \{\mathbf{B}_1, \dots, \mathbf{B}_r\}$ particiójára

$$I\left(\nu_{\mathcal{A}_n}\|\mu_{\mathcal{A}_n}\right) \ge I(\nu_{\mathcal{B}}\|\mu_{\mathcal{B}}) - \frac{\varepsilon}{2}.$$

Legyen $\mathcal{B} = \{\mathbf{B}_1, \dots, \mathbf{B}_r\}$ az Y halmaz olyan véges mérhető particiója, amelyre $\mu(\mathbf{B}_l) > 0$ minden $1 \le l \le r$ indexre. A 13. feladat eredménye segítségével be fogjuk látni, hogy tetszőleges $\delta > 0$-ra és $n > n(\delta)$-ra az $\mathcal{A}_n = \{\mathbf{A}_1, \dots, \mathbf{A}_{p(n)}\}$ partició elemeit r osztályba lehet sorolni úgy, hogy a j-ik, $j = 1, \dots, r$, osztályba tartozó $\mathbf{A}^{(j)}_{j(1)}, \dots, \mathbf{A}^{(j)}_{j(s))}$, $s = s(n, j)$, halmazok $\mathbf{A}_{(j)} = \mathbf{A}_{n,(j)} = \bigcup\limits_{l=1}^{s} \mathbf{A}^{(j)}_{j(l)}$ uniója teljesíti a $\mu(\mathbf{A}_{(j)} \Delta \mathbf{B}_j) \le \delta$ és $\nu(\mathbf{A}_{(j)} \Delta \mathbf{B}_j) \le \delta$, $j = 1, \dots, r$, egyenlőtlenségeket. Elég kis δ szám választásával e tény segítségével fogjuk bizonyítani a kívánt egyenlőtlenséget.

A 13. feladat eredménye alapján minden elég nagy n-re és $\mathbf{B}_j \in \mathcal{B}$-re létezik olyan $\bar{\mathbf{A}}_{(j)} \in \mathcal{A}_n$ halmaz, ($\mathcal{A}_n$-val jelöljük az $\mathcal{A}_n$ partició elemeinek uniójaiból álló σ-algebrát is,) amelyre $\mu(\bar{\mathbf{A}}_j \Delta \mathbf{B}_j) \le \frac{\delta}{4r}$, $\nu(\bar{\mathbf{A}}_j \Delta \mathbf{B}_j) \le \frac{\delta}{4r}$, $j = 1, \dots, r$. Ezek az $\bar{\mathbf{A}}_{(j)}$ halmazok nem feltétlenül diszjunktak, és az uniójuk nem mindig adja ki az egész teret, de mivel a $\mathbf{B}_j$, $j = 1, \dots, r$, halmazok diszjunktak, és az uniójuk az egész

Y tér, ezért $\mu\left(\bigcup\limits_{j=1}^{r}\bar{\mathbf{A}}_{(j)}\right) \ge 1-\frac{\delta}{4}$, $\nu\left(\bigcup\limits_{j=1}^{r}\bar{\mathbf{A}}_{(j)}\right) \ge 1-\frac{\delta}{4}$, $\mu(\bar{\mathbf{A}}_{(j)}\Delta\bar{\mathbf{A}}_{(j')}) \le \frac{\delta}{2r}$ és $\nu(\bar{\mathbf{A}}_{(j)}\Delta\bar{\mathbf{A}}_{(j')}) \le \frac{\delta}{2r}$, ha $j \ne j'$. Definiáljuk az $\mathbf{A}_{(j)} = \bar{\mathbf{A}}_{(j)} \setminus \bigcup\limits_{l=1}^{j-1}\bar{\mathbf{A}}_{(l)}$, $j = 1,\dots,r-1$, és $\mathbf{A}_r = Y \setminus \bigcup\limits_{l=1}^{r-1}\mathbf{A}_l$ halmazokat. Ezek az $\mathbf{A}_{(j)}$ halmazok teljesítik a kívánt feltételeket.

Rögzítsünk egy elég kis $\delta > 0$ számot, és elég nagy n-re vezessük be a $\mathcal{A}_n$ partició $\bar{\mathcal{A}}_n = \{\mathbf{A}_{(1)},\dots,\mathbf{A}_{(s)}\} = \{\mathbf{A}_{n,(1)},\dots,\mathbf{A}_{n,(s)}\}$ durvítását, ahol az $\mathbf{A}_{(j)}$ halmazok a $\mathbf{B}_j$ halmazok előzőekben definiált approximációi. Ekkor $I(\nu_{\bar{\mathcal{A}}_n}\|\mu_{\bar{\mathcal{A}}_n}) \le I(\nu_{\mathcal{A}}\|\mu_{\mathcal{A}})$, és mivel az $\mathbf{A}_{(j)}$ és $\mathbf{B}_{(j)}$, $j = 1,\dots,r$ halmazok μ és ν mértéke csak δ-val térhet el, ezért $\left|I(\nu_{\bar{\mathcal{A}}_n}\|\mu_{\bar{\mathcal{A}}_n}) - I(\nu_{\mathcal{B}}\|\mu_{\mathcal{B}})\right| \le \frac{\varepsilon}{2}$, ha δ-t elég kicsinek választjuk. Innen következik a feladat állítása abban az esetben, ha a ν mérték abszolút folytonos a μ mértékre.

Az az eset, amikor a ν mérték nem abszolút folytonos a μ mértékre hasonlóan tárgyalható, de egyszerűbb. Definiáljuk a $\mathcal{B} = (\mathbf{B}_1, \mathbf{B}_2)$ particiót, ahol

$$\mathbf{B}_1 = \left\{u\colon \frac{d\mu}{d\nu}(u) = 0\right\}, \quad \mathbf{B}_2 = \left\{u\colon \frac{d\mu}{d\nu}(u) > 0\right\}.$$

Ekkor tetszőleges $\delta > 0$-ra és $n \ge n(\delta)$-ra az $\mathcal{A}_n$ particiónak van olyan $\bar{\mathcal{A}} = \{\bar{\mathbf{A}}_1, \bar{\mathbf{A}}_2\}$ durvítása, amelyre $\nu(\bar{\mathbf{A}}_1) \ge \nu(\mathbf{B}_1) - \delta$ és $\mu(\bar{\mathbf{A}}_1) \le \delta$. Innen következik, hogy $I(\nu_{\mathcal{A}}\|\mu_{\mathcal{A}}) \ge I(\nu_{\bar{\mathcal{A}}_n}\|\mu_{\bar{\mathcal{A}}_n})$, és

$$I(\nu_{\bar{\mathcal{A}}}\|\mu_{\bar{\mathcal{A}}}) = \log\frac{\nu(\bar{\mathbf{A}}_1)}{\mu(\bar{\mathbf{A}}_1)}\nu(\bar{\mathbf{A}}_1) + \log\frac{1-\nu(\bar{\mathbf{A}}_1)}{1-\mu(\bar{\mathbf{A}}_1)}(1-\nu(\bar{\mathbf{A}}_1)) \ge \frac{1}{\varepsilon},$$

ha $\delta > 0$ elég kicsi. Innen adódik a feladat állítása ebben az esetben is.

15.) Rögzítsünk egy $\delta_n \to 0$ sorozatot. Az R^1 számegyenes olyan egymásba skatulyázott (nyílt) intervallumokból, egyelemű halmazokból és két félegyenesből álló egymásba skatulyázott véges $\mathcal{A}_n$ particióit tekintjük, amelyekre az $\mathcal{A}_n$ particióban tartalmazott intervallumoknak és félegyeneseknek a G eloszlás által meghatároztott mértéke kisebb mint δ_n.

Lássuk be először, hogy ilyen particiók valóban léteznek. Az $\mathcal{A}_n$ partició megkonstruálása érdekében tekintsünk egy $(-\infty, a)$, és (b, ∞) félegyenest, amelyeknek a G eloszlás által indukált μ_G mértéke kisebb mint δ_n, és legyenek $y_1,\dots,y_k$ azok a pontok amelyekre $\mu_G(y_j) \ge \delta_n$. Definiáljuk a a $\bar{\mu}_G$ mértéket a $\bar{\mu}_G(\mathbf{A}) = \mu_G((\mathbf{A} \cap [a,b]) \setminus \{y_1,\dots,y_k\})$ képlettel. Ekkor minden $x \in R^1$ pontnak van olyan $(x-\eta(x), x+\eta(x))$ környezete, amelyre $\bar{\mu}_G\left((x-\eta(x), x+\eta(x))\right) < \delta$. Ezek lefedik a (kompakt) $[a,b]$ intervallumot, ezért kiválasztható ezen intervallumok közül véges sok, amelyek szintén lefedik az $[a,b]$ intervallumot. Tekintsük a számegyenesnek azt a particióját véges szakaszokra, azok végpontjaira (mint egyelemű halmazra) és két félegyenesre, amelyet ezen intervallumok végpontjai és az $y_1,\dots,y_k$ pontok határoznak meg. Az így definiált $\bar{\mathcal{A}}_n$ partició nem egy pontból álló elemeinek a μ_G mértékes kisebb mint δ. Legyen $\mathcal{A}_n$ a $\bar{\mathcal{A}}_k$, $k = 1, \dots n$, particiók közös finomítása. Ez teljesíti a kívánt feltételeket.

A 14. feladat eredménye alapján tehát tetszőleges $\varepsilon > 0$ számra létezik a számegyenesnek olyan egyelemű halmazokból, nyílt intervallumokból és két félegyenesből álló $\mathcal{A}$ particiója, amelyre

$$\begin{aligned} &|I(G_{\mathcal{A}} \| F_{\mathcal{A}}) - I(G\|F)| \le \frac{\varepsilon}{2} \quad \text{ha } I(G\|F) < \infty \\ &I(G_{\mathcal{A}} \| F_{\mathcal{A}}) > \frac{2}{\varepsilon} \quad \text{ha } I(G\|F) = \infty, \end{aligned}$$

és $\mu_G(\mathbf{A}) \le \frac{\varepsilon}{2}$, ha $\mathbf{A} \in \mathcal{A}$, és $\mathbf{A}$ nem egy pontból áll. (Emlékeztetőül, ha $\mathcal{A} = \{\mathbf{A}_1, \dots, \mathbf{A}_k\}$, akkor minden minden $\mathbf{A}_l \in \mathcal{A}$, $l = 1, \dots, k$, elemnek kijelöljük egy x_l pontját, és az $F_{\mathcal{A}}$ illetve $G_{\mathcal{A}}$ eloszlásfüggvényt azok az eloszlások által meghatározott eloszlásfüggvények, amelyek az x_l értéket $p_l = \mu_F(\mathbf{A}_l)$ illetve $q_l = \mu_G(\mathbf{A}_l)$ valószínűséggel veszik fel.) Legyenek $\xi_1, \xi_2, \dots$, független F eloszlású valószínűségi változók. Ezek meghatározzák az $F_n(x) = \frac{1}{n} \sum\limits_{j=1}^{n} I(\xi_j \le x)$ empirikus eloszlásfüggvényt. Definiáljuk a $\bar{\xi}_j$, $j = 1, 2, \dots$, valószínűségi változókat a következő módon: Ha $\xi_j \in \mathbf{A}_l \in \mathcal{A}$, akkor $\bar{\xi}_j = x_l$. Ekkor $\bar{\xi}_1, \bar{\xi}_2, \dots$ független $F_{\mathcal{A}}$ eloszlású valószínűségi változók, és a $\bar{\xi}_j$ valószínűségi változók konstrukciójából következik, hogy a belőlük elkészített $\bar{F}_n = \bar{F}_{n,\mathcal{A}}$ empirikus eloszlásfüggvény teljesíti az $F_n(x) = \bar{F}_n(x)$, $F_n(x-0) = \bar{F}_n(x-0)$ azonosságot minden olyan x pontban, amely az $\mathcal{A}$ particióban szereplő intervallumok valamelyikének végpontja. Innen, és a $\mathcal{A}$ elemeinek μ_G mértékére adott becslésből következik, hogy

$$\sup_x |F_n(x) - G(x)| - \varepsilon \le \sup_x \left|\bar{F}_n(x) - G_{\mathcal{A}}(x)\right| \le \sup_x |F_n(x) - G(x)| + \varepsilon.$$

Viszont a 12. feladat eredménye alapján

$$-I\left(G_{\mathcal{A}} \| F_{\mathcal{A}}\right) - C\varepsilon \le \frac{1}{n} \log P\left(\sup_x \left|\bar{F}_n(x) - G_{\mathcal{A}}(x)\right| > \varepsilon\right) \le -I\left(G_{\mathcal{A}} \| F_{\mathcal{A}}\right) + C\varepsilon,$$

alkalmas C számmal, ha $n > n(\varepsilon, F, G)$. Az utolsó két egyenlőtlenségből és abból a tényből, hogy $I(G_{\mathcal{A}} \| F_{\mathcal{A}})$ jól approximálja az $I(G\|F)$ I-divergenciát, következik a feladat állítása.

16.) Az $L = L(\varepsilon)$ számot elég nagyra választva elérhetjük, hogy az

$$\bar{F}(x) = P(\xi_1 \le x|\ |\xi_1| \le L)$$

eloszlásfüggvény teljesítse a $\sup\limits_x |\bar{F}(x) - F(x)| \le \varepsilon$ és az $|I(G\|\bar{F}) - I(G\|F)| \le \varepsilon$ egyenlőtlenségeket. (Az utóbbi egyenlőtlenség teljesüléséhez kell az a feltétel, hogy a G eloszlás egy véges intervallumba van koncentrálva.) Ezért a 15. feladat eredményéből ezzel a $\bar{F}$ eloszlásfüggvény választásával következik az ebben a feladatban felírt becslés az $\sup\limits_x |F_n(x) - G(x)| \le \eta$ esemény feltételes valószínűségére a $\{|\xi_k| \le L,\ 1 \le k \le n\}$ feltétel mellett.

A 10. feladat eredménye alapján a nagy eltérés tétel felső becslésének a bizonyításához elég megmutatni, hogy minden $\varepsilon > 0$-ra és olyan G eloszlásra, amelyre $a < \int u\,dG(u) < a + \varepsilon$ és $I(G\|F) < \inf\limits_{t\ge 0}(ta - \log R(t)) + \varepsilon$ teljesül a

$$-\frac{1}{n}\log P\left(\frac{S_n}{n} \in [a,b]\right) \le I(G\|F) + 2\varepsilon$$

egyenlőtlenség. Létezik ugyanis olyan G eloszlás, amelyre $a < \int u\,dG(u) < a + \varepsilon$, és $I(G\|F) < \inf\limits_{t\ge 0}(ta - \log R(t) + \varepsilon$. Sőt, feltehetjük, hogy ez a G eloszlás véges intervallumba van koncentrálva, mert egy elég nagy $K > 0$ számot választva és a G függvényt a $\bar{G}(x) = \frac{G(x)-G(-K)}{G(K)-G(-K)}$, $-K \le x \le K$, helyettesítve olyan függvényt kapunk, amelyik szintén teljesíti a fenti egyenlőtlenségeket. (Pontosabban, a 10. feladatban felhasznált eredmény nem alkalmazható abban a speciális (elfajult) esetben, amikor az $R(t) = Ee^{t\xi_1}$ minden $t \ge 0$-ra értelmezve van, és $\lim\limits_{t\to\infty}[\log R(t)]' = a$. De ezt az esetet, amikor $P(\xi_1 \le a) = 1$ reláció is teljesül, a *Nagy eltérések elmélete; Független, valós értékű valószínűségi változók* feladatsor 13. feladatának eredménye leírja, és ezzel itt nem foglalkozunk. Feltehetjük továbbá, hogy $I(G\|F) < \infty$, mert különben az állítás nyílvánvaló.) Alkalmazzuk a feladat első állítását ezt az állítást kielégítő G és $\bar{F}$ eloszlással, ε helyett $\varepsilon' = \frac{1}{2}\left(\int u\,dG(u) - a\right)$ paraméter választással. Ekkor $\varepsilon' \le \varepsilon$, és ha $\varepsilon > 0$ elég kicsi, akkor az $\varepsilon' \le \frac{1}{2}\left(b - \int u\,dG(u)\right)$ egyenlőtlenség is teljesül. Azt kapjuk, hogy

$$P\left(\omega\colon\ |F_n(x,\omega) - G(x)| \le \eta |\ |\xi_k(\omega)| \le L,\ 1 \le k \le n\right) \ge e^{-nI(G\|\bar{F}) - nC(\eta,\varepsilon')},$$

$|IG(\|F) - I(G\|\bar{F})| \le \varepsilon'$. Ezért a kívánt állítás bizonyításához elég megmutatni azt, hogy elég kis $\eta = \eta(L) > 0$ számra

$$\left\{\omega\colon\ \sup_x |F_n(x,\omega) - G(x)| \le \eta\right\} \cap \mathbf{A}_n \subset \left\{\omega\colon\ \frac{S_n(\omega)}{n} \in [a,b]\right\},$$

ahol $\mathbf{A}_n = \mathbf{A}_n(L) = \{\omega\colon\ |\xi_k(\omega)| \le L,\ 1 \le k \le n\}$, és $L > 0$ ugyanaz a szám, amely az előbb tekintett feltételes valószínűség feltételében szerepel. Innen ugyanis következik, hogy

$$\begin{aligned} P\left(\frac{S_n(\omega)}{n} \in [a,b]\right) &\ge P\left(|F_n(x,\omega) - G(x)| \le \eta \,|\, |\xi_k(\omega)| \le L,\ 1 \le k \le n\right) P(\mathbf{A}_n) \\ &\ge e^{-nI(G\|\bar{F}) - nC(\eta,\varepsilon')} P(|\xi_1| \le L)^n \ge e^{-nI(G\|\bar{F}) - n\bar{C}(\eta,\varepsilon)}, \end{aligned}$$

ahol $\bar{C}(\varepsilon) \to 0$, ha $\varepsilon \to \infty$. Parciális integrálással kapjuk, hogy $\omega \in A_n$ és $\left\{\omega\colon \sup\limits_x |F_n(x,\omega) - G(x)| \le \eta\right\}$ esetén

$$\left|\frac{S_n(\omega)}{n} - \int u\, dG(u)\right| = \left|\int u\, dF_n(u,\omega) - dG(u)\right|$$
$$\le \int_{-L}^{L} |F_n(u,\omega) - G(u)|\, du \le L\eta.$$

Viszont elég kis $\eta > 0$-ra $L\eta \le \int u\, dG(u) - a$ és $L\eta \le b - \int u\, dG(u)$. Innen következik a feladat állítása.

17.) Mivel $\mu \notin \mathbf{F}$, és $\mathbf{F}$ zárt, azaz $X \setminus \mathbf{F}$ nyílt halmaz, ezért létezik az $\mu \in X$ pontnak az $\mathbf{F}$ halmaztól diszjunkt $\mathbf{U} = \{\nu\colon\ |\nu(\mathbf{B}_l) - \mu(\mathbf{B}_l)| < \varepsilon,\ 1 \le l \le k\}$ nyílt környezete, ahol $\mathbf{B}_l$ az $(Y, \mathcal{B})$ tér mérhető részhalmazai. Legyen Definiáljuk a $\mathbf{G}_1 = \{\nu\colon\ |\nu(\mathbf{B}_l) - \mu(\mathbf{B}_l)| < \varepsilon/4,\ 1 \le l \le k\}$ és $\mathbf{G}_2 = X \setminus \{\nu\colon\ |\nu(\mathbf{B}_l) - \mu(\mathbf{B}_l)| \le \varepsilon/2,\ 1 \le l \le k\} = \{\nu\colon\ |\nu(\mathbf{B}_l) - \mu(\mathbf{B}_l)| > \varepsilon/2$ valamely $1 \le l \le k$ indexre.$\}$ halmazokat. Ezek a $\mathbf{G}_1$, $\mathbf{G}_2$ nyílt halmazok teljesítik a T_3 tulajdonságban szereplő feltételeket.

A $\mathbf{K} = \mathbf{K}(\mathbf{B}, \varepsilon) = \{\nu\colon\ |\nu(\mathbf{B}) - \mu(\mathbf{B})| \le \varepsilon\}$ halmaz, ahol $\mathbf{B}$ az $(Y, \mathcal{B})$ tér mérhető részhalmaza, a μ pontot tartalmazó zárt halmaz. Az összes ilyen halmaz metszete szintén zárt, és egyedül a μ pontot tartalmazza. Ezért a $\{\mu\}$ halmaz zárt.

18.) Tekintsük az Y tér egy véges $\mathcal{D} = \{\mathbf{D}_1, \dots, \mathbf{D}_k\}$ véges particióját, és adva két μ és ν mérték az $(X, \mathcal{X})$ téren, legyen $\mu_{\mathcal{D}}$ és $\nu_{\mathcal{D}}$ e mérték vetülete a $\mathcal{D}$ particióra, és $I(\nu_{\mathcal{D}} \| \mu_{\mathcal{D}}) = \sum\limits_{l=1}^{k} \log \frac{\nu(\mathbf{D}_l)}{\mu(\mathbf{D}_l)} \nu(\mathbf{D}_l)$ ezen vetületmértékek I-divergenciája. A $I(\nu_{\mathcal{D}} \| \mu_{\mathcal{D}})$ függvény rögzített μ mértékkel mint a ν mérték függvénye alulról félig folytonos a τ topológia szerint. Ugyanis azon ν mértékek halmaza, amelyekre $I(\nu_{\mathcal{D}} \| \mu_{\mathcal{D}}) = \infty$ nyílt halmaz, (megegyezik azon ν mértékek halmazával, amelyekre létezik a $\mathcal{D}$ particiónak olyan $\mathbf{D}_j$ eleme, amelyre $\mu(\mathbf{D}_j) = 0$, és $\nu(\mathbf{D}_j) > 0$),), a komplementer zárt halmazon pedig folytonos. Viszont a 14. feladat eredménye alapján választhatunk $\mathcal{D}_n$ particióknak olyan sorozatát, amelyekre $\lim\limits_{n\to\infty} I(\nu_{\mathcal{D}_n} \| \mu_{\mathcal{D}_n}) = \sup\limits_n I(\nu_{\mathcal{D}_n} \| \mu_{\mathcal{D}_n}) = I(\nu_{\mathcal{D}} \| \mu_{\mathcal{D}})$. Mivel alulról félig folytonos függvények szuprémuma alulról félig folytonos, innen következik a feladat első állítása.

A feladat második állításának bizonyítása érdekében vegyük a számegyenes olyan intervallumokból és félegyenesekből álló $\mathcal{D}_n = \{\mathbf{D}_{1,n}, \dots, \mathbf{D}_{k(n),n}\}$ növekvő véges partícióból álló sorozatát, $n = 1, 2, \dots$, amelyre a $\bigcup \mathcal{D}_n$ halmazrendszer generálja a $\mathcal{B}$ Borel σ-algebrát. Ekkor alkalmazva az első rész bizonyítását az $I(\nu_{\mathcal{D}_n} \| \mu_{\mathcal{D}_n})$, $n = 1, 2, \dots$, I-divergenciák sorozatára a 14. feladat eredményének segítségével $n \to \infty$ határátmenettel megkapjuk a kívánt állítást.

19.) Tekintsünk egy olyan $\nu \in \mathbf{G}$ mértéket, amelyre $I(\nu\|\mu) \le \inf\limits_{\chi \in \mathbf{G}} I(\chi\|\mu) + \varepsilon$, és tekintsük ennek olyan $\mathbf{U}$ környezetét, amelyre $\mathbf{U} \subset \mathbf{G}$, és $\mathbf{U}$ a következő alakú: $\mathbf{U} = \{\chi\colon\ \chi \in X, |\chi(\mathbf{B}_1) - \nu(\mathbf{B}_l)| < \varepsilon,\ l = 1, 2, \dots, k\}$ alkalmas $\varepsilon > 0$ számmal és $\mathcal{B} = \{\mathbf{B}_1, \dots, \mathbf{B}_l\}$ az Y tér alkalmas particiója. Tekintsük a μ és ν mértékek vetületét ezekre a particiókra. Jelöljük ezeket $\mu_{\mathcal{B}}$ és $\nu_{\mathcal{B}}$-vel, és vizsgáljuk ezek I-divergenciáját. A 12. feladat eredménye alapján $I(\nu_{\mathcal{B}}\|\mu_{\mathcal{B}})=\sum\limits_{l=1}^{p}\log\frac{\nu(\mathbf{B}_l)}{\mu(\mathbf{B}_l)}\nu(\mathbf{B}_l)\le I(\nu\|\mu)$. Ezért a 11. feladat eredményét felhasználva kapjuk, hogy $I(\nu_{\mathcal{B}}\|\mu_{\mathcal{B}})\le \inf\limits_{\chi\in\mathbf{G}} I(\chi\|\mu)+\varepsilon$, és $\mu_n(\mathbf{G})\ge\mu_n(\mathbf{U})\ge e^{-n(I(\nu_{\mathcal{B}}\|\mu_{\mathcal{B}})-C\varepsilon)}$, alkalmas >0 számmal, ha $n\ge n(\varepsilon)$. Ezért

$$\mu_n(\mathbf{G})\ge\inf_{\chi\in\mathbf{G}} e^{-n(I(\chi\|\mu)-\varepsilon-C\varepsilon)},\quad \text{ha } n\ge n(\varepsilon),$$

Innen következik a feladat állítása.

20.) A feladat állítása kapcsolódik a funkcionálanalízis egyik klasszikus eredményéhez, amely szerint egy Banach tér duális terének zárt egységgömbje a gyenge topológiában kompakt. Természetesebbnek tűnik viszont e tételre való hivatkozás helyett a bizonyítás módszerét adaptálni. Ez a Tyihonov tétel alkalmazása, a topologikus tér természetes beágyazása a $[0,1]$ intervallum (a szokásos topológiával) nagy számosságú példányának direkt szorzatába. Ez utóbbi tér kompakt a Tyihonov tétel szerint.

Tekintsünk az $(Y,\mathcal{B})$ minden mérhető $\mathbf{B}\subset\mathcal{B}$ részhalmazára egy $(Z_{\mathbf{B}},\mathcal{C}_{\mathbf{B}})$ teret, amely a $[0,1]$ intervallum a szokásos topológiával. Legyen ezek direkt szorzata a szorzattopológiával $(\mathbf{Z},\mathcal{C})=\prod\limits_{\mathbf{B}\in\mathcal{B}}(Z_{\mathbf{B}},\mathcal{C}_{\mathbf{B}})$. A $\mathbf{Z}$ tér felfogható mint azokból a függvényekből álló halmaz, amelyek az Y tér $\mathcal{B}$ mérhető halmazait képezik le a $[0,1]$ intervallumra. A $\mathcal{C}$ topológia nyílt halmazait pedig ezen a téren a $\{z_{\mathbf{B}_1}\in\mathbf{G}_1,\dots,z_{\mathbf{B}_k}\in\mathbf{G}_k\}$ halmazok generálják, ahol $k=1,2,\dots,$, $\mathbf{B}_1,\dots,\mathbf{B}_k$ az Y halmaz mérhető halmazai, $\mathbf{G}_1,\dots,\mathbf{G}_k$ pedig a $[0,1]$ intervallum nyílt részhalmazai. Vegyük észre, hogy $X\subset\mathbf{Z}$, és az X téren bevezetett $\mathcal{X}$ τ topológia a $\mathbf{Z}$ téren definiált $\mathcal{C}$ topológia megszorítása az X térre. Így az $(X,\mathcal{X})$ teret beágyaztuk egy kompakt topologikus térbe. Tudjuk, hogy egy kompakt halmaz zárt részhalmazai szintén kompaktak. Ezenkívül azt is tudjuk (a 18. feladat eredménye), hogy minden $\mu\in X$-re az $I(\cdot\|\mu)$ I-divergencia alulról félig folytonos függvény az $(X,\mathcal{X})$ téren. Egy alulról félig folytonos $I(\cdot)$ függvény $\{x\colon I(x)\le R\}$ tipusú nívóhalmazai zártak. Ezen tények alapján próbáljuk meg bebizonyítani a feladat állítását. A problémát az okozza, hogy az I-divergencia a $(\mathbf{Z},\mathcal{C})$ kompakt térnek csak az X részhalmazán van értelmezve, és ennek topológiai tulajdonságairól semmit sem tudunk. E nehézségen úgy segítünk, hogy az X teret beágyazzuk a $\mathbf{Z}$ tér egy $\mathbf{Z}_0$ zárt részhalmazába (ez a $\mathbf{Z}_0$ halmaz valójában az X tér lezártja, de erre a tényre nincs szükségünk), és az I-divergencia függvényt kiterjesztjük az X halmazról a kompakt $\mathbf{Z}_0$ halmazra úgy, hogy a kiterjesztett függvény is alulról félig folytonos. Ráadásul a kiterjesztett függvény értéke a $\mathbf{Z}_0\setminus X$ halmazon ∞. E tények segítségével nem lesz nehéz bizonyítani a feladat állítását.

Definiáljuk a $\mathbf{Z}_0\subset\mathbf{Z}$ halmazt mint a következő metszetet: $\mathbf{Z}_0=\bigcap\limits_{\mathcal{U}}\mathbf{Z}_{\mathcal{U}}$, ahol az $\mathcal{U}$ indexhalmaz jelöli az Y halmaz összes véges particióját mérhető halmazokra, és az

Y tér egy $\mathcal{U} = \{\mathbf{B}_1, \dots, \mathbf{B}_k\}$, $\mathbf{B}_l \in \mathcal{B}$, $l = 1, \dots, k$ véges particiójára legyen

$$\mathbf{Z}_{\mathcal{U}} = \left\{ z\colon\ z \in \mathbf{Z},\ \sum_{l=1}^{k} z_{\mathbf{B}_l} = 1 \right\}.$$

Ekkor $\mathbf{Z}_0$ a $\mathbf{Z}$ halmaz zárt részhalmaza, mivel a zárt $\mathbf{Z}_{\mathcal{U}}$ halmazok metszete. A $\mathbf{Z}_0$ halmaz az $(Y, \mathcal{B})$ tér mérhető halmazain értelmezett végesen additiv nem negatív értékű 1-re normált halmazfüggvényeiből áll, ezért $X \subset \mathbf{Z}_0$.

Az I-divergencia kiterjesztésének és tulajdonságainak bizonyítása érdekében jegyezzük meg, hogy ha ν és μ két valószínűségi mérték, akkor

$$I(\nu\|\mu) = \sup_{\mathcal{U}} I(\nu_{\mathcal{U}}\|\mu_{\mathcal{U}}) = \sup_{\{\mathbf{B}_1,\dots,\mathbf{B}_k\}} \sum_{l=1}^{k} \log \frac{\nu(\mathbf{B}_l)}{\mu(\mathbf{B}_l)} \nu(\mathbf{B}_l), \quad \mathcal{U} = \{\mathbf{B}_1, \dots, \mathbf{B}_k\}$$

ahol a szuprémum az Y tér $\mathcal{U} = \{\mathbf{B}_1, \dots, \mathbf{B}_l\}$ véges (mérhető) particiójára vétetik. (Ez következik a 12. feladat eredményéből.) Definiáljuk az $I(\nu\|\mu)$ függvényt (rögzített μ mértékkel) minden $\nu \in \mathbf{Z}_0$-re, azaz akkor is, ha ν additív, de nem σ-additív halmazfüggvény az előbb felírt szuprémum segítségével. Ekkor $I(\cdot\|\mu)$ alulról félig folytonos, mert alulról félig folytonos függvények szuprémuma. Belátjuk, hogy az így kiterjesztett I-divergenciára $I(\nu\|\mu) = \infty$, ha $\nu \in \mathbf{Z}_0 \setminus X$ (és $\mu \in X$). Ugyanis az, hogy a pozitív értékű, additív μ halmazfüggvény σ-additív azzal ekvivalens, hogy tetszőleges $\mathbf{B}_n \in \mathcal{B}$ halmazsorozatra, amelyre $\mathbf{B}_n \supset \mathbf{B}_{n+1}$, $n = 1, 2, \dots$, és $\bigcap\limits_{n=1}^{\infty} \mathbf{B}_n = \emptyset$, $\lim\limits_{n\to\infty} \mu(\mathbf{B}_n) = 0$. Mivel $\mu \in X$, és $\nu \in \mathbf{Z}_0 \setminus X$, létezik olyan leszálló $\mathbf{B}_n \in \mathcal{B}$ az üres halmazhoz tartó $\mathbf{B}_n$ halmazsorozat, amelyre $\lim\limits_{n\to\infty} \mu(\mathbf{B}_n) = 0$ és $\lim\limits_{n\to\infty} \nu(\mathbf{B}_n) = \alpha > 0$. Ezért tekintve az $\mathcal{U}_n = \{\mathbf{B}_n, Y \setminus \mathbf{B}_n\}$ particiókat $I(\nu\|\mu) \ge \lim\limits_{n\to\infty} I(\nu_{\mathcal{U}_n}\|\mu_{\mathcal{U}_n}) = \infty$.

A fentiek alapján könnyen bizonyítható a feladat állítása. Tudjuk, hogy $\mathbf{Z}_0$ a $(\mathbf{Z}, \mathcal{C})$ tér zárt részhalmaza, és a $\mathbf{Z}_0$-ra kiterjesztett $I(\cdot\|\mu)$ halmaz alulról félig folytonos a $(\mathbf{Z}_0, \mathcal{C}_0)$ téren, ahol $\mathcal{C}_0$ a $\mathcal{C}$ topológia megszorítása a $\mathcal{Z}_0$ halmazra. Ezért az $\mathbf{A}(R) = \{\nu\colon\ \nu \in \mathbf{Z}_0,\ I(\nu\|\mu) \le R\}$ halmaz a $(\mathbf{Z}_0, \mathcal{C}_0)$ tér zárt, következésképp kompakt részhalmaza minden $0 \le R < \infty$ számra. Viszont $\mathbf{A}(R) \subset X$. Ezért $\mathbf{A}(R) = A(R) \cap X$ nemcsak a $(\mathbf{Z}_0, \mathcal{C}_0)$ téren, hanem annak megszorításán, az $(X, \mathcal{X})$ téren is kompakt.

21.) Mint az előző feladatban láttuk, az $(X, \mathcal{X})$ tér beágyazható egy kompakt T_2 térbe. (A $[0, 1]$ tér önmagával vett nagy példányszámú direkt szorzatába.) A topológia standard eredményei alapján egy kompakt T_2 tér T_4 tér. (Tetszőleges két diszjunkt zárt halmaznak létezik diszjunkt nyílt környezete.) Egy T_4 tér egyben T_ρ tér, és annak tetszőleges altere, tehát például az $(X, \mathcal{X})$ tér is T_ρ tér.

22.) Rögzítsünk egy $L = L(\varepsilon) > \inf\limits_{\nu\in\mathbf{K}} I(\nu) - \varepsilon$ számot, ha ez az infimum véges, és $L = L(\varepsilon) = \frac{1}{\varepsilon}$, ha ez az infimum végtelen. Válasszunk minden $\nu \in \mathbf{K}$ mértéknek egy olyan $\mathbf{G}(\nu, \varepsilon)$ nyílt környezetét, amelyre $\liminf\limits_{n\to\infty} -\frac{1}{n} \log \mathbf{Q}_n(\mathbf{G}(\nu, \varepsilon)) > L - \varepsilon$.

Ezen halmazok lefedik a $\mathbf{K}$ halmazt. Vegyük a $\mathbf{K}$ halmaznak egy ilyen halmazokból álló véges $\mathbf{G}_1 = \mathbf{G}(\nu_1, \varepsilon)$, ... , $\mathbf{G}_m = \mathbf{G}(\nu_m, \varepsilon)$ fedését. Ekkor elég nagy n indexre

$$\mathbf{Q}_n(\mathbf{K}) \le \sum_{l=1}^{m} \mathbf{Q}_n(\mathbf{G}_l) \le \text{const.} \sup_{1\le k\le m} e^{-n(L-\varepsilon)} \le \text{const.} \sup_{\nu\in\mathbf{K}} e^{-n(I(\nu)-2\varepsilon)}.$$

Innen következik a (iii′) tulajdonság.

Lássuk be, hogy az $I(\cdot\|\mu)$ függvény és a $Q_n = \mu^{(n)}$ mértékek, ahol $\mu^{(n)}$ a μ_n empirikus mérték eloszlása, teljesítik a feladat feltételeit. Tekintsük az Y tér egy alkalmas véges $\mathcal{C} = \{\mathbf{C}_1, \dots, \mathbf{C}_k\}$ partícióját, vegyünk egy kis $\delta > 0$ számot, és definiáljuk egy $\nu \in X$ pont alkalmas nyílt környezetét a következő módon: $\mathbf{G}(\nu, \varepsilon) = \{\chi\colon\ \chi \in X, |\chi(\mathbf{C}_j) - \nu(\mathbf{C}_j)| < \delta,\ j = 1, \dots k\}$, $\delta = \delta(\varepsilon)$. A 12. feladat eredménye alapján alkalmas $\mathcal{C}$ partició választásával elérhető, hogy az $I(\nu_\mathcal{C}\|\mu_\mathcal{C}) \ge I(\nu\|\mu) - \frac{\varepsilon}{2}$ ha $I(\nu\|\mu) < \infty$, és $I(\nu_\mathcal{C}\|\mu_\mathcal{C}) > \frac{1}{\varepsilon}$, ha $I\nu\|\mu) = \infty$ egyenlőtlenségek teljesüljenek. Másrészt a 11. feladatban bizonyított (3) egyenlőtlenség jobboldala alapján $\mathbf{Q}_n(\mathbf{G}(\nu, \varepsilon)) < e^{-n(I(\nu_\mathcal{C}\|\mu_\mathcal{C})-\varepsilon/2)}$, ha a $\delta = \delta(\varepsilon) > 0$ számot elég kicsinek választjuk, és $n \ge n(\delta)$. (Jelen esetben a (3) formulától eltérő jelölést használunk. A $\mathbf{Q}_n(\mathbf{G}(\nu, \varepsilon))$ kifejezés játssza a (3) formulában szereplő $\mu^{(n)}(\cdot)$ valószínűség szerepét, és ott $\delta = \delta(\varepsilon)$-t írunk ε helyett. Továbbá a $\mu^{(n)}(\cdot)$ valószínűségre a (3) formulában adott felső becslés a mostani jelölésben a $e^{-nI(\nu\|\mu)-nC(\delta)}$ kifejezéssel egyenlő.) Ezért $\liminf\limits_{n\to\infty} -\frac{1}{n} \log \mathbf{Q}_n(\mathbf{G}(\nu, \varepsilon)) \ge I(\nu\|\mu) - \varepsilon$, ha $I(\nu\|\mu) < \infty$, és nagyobb mint $\frac{1}{\varepsilon}$, ha $I(\nu\|\mu) = \infty$.

23.) Tetszőleges $\mathcal{C} \in \mathcal{C}^*$-ra, $\varepsilon \ge 0$-ra

$$\mu^{(n)}(\mu_n \in \mathbf{F}) \le \mu^{(n)}\left(\mu_{n\mathcal{C}} \in \mathbf{F}_\mathcal{C}\right) \le \mu^{(n)}\left(\mu_{n\mathcal{C}} \in [\mathbf{F}_\mathcal{C}]\right) \le \sup_{\nu\in[\mathbf{F}_\mathcal{C}]} e^{-n(I(\nu_\mathcal{C}\|\mu_\mathcal{C})-\varepsilon)},$$

ha $n > n(\varepsilon, \mathcal{C})$. Ez például következik a 22. feladat eredményéből, és abból a tényből, hogy $[\mathbf{F}_\mathcal{C}]$ kompakt halmaz az $(X_\mathcal{C}, \mathcal{A}_\mathcal{C})$ téren. Itt felhasználtuk, hogy egy véges halmazon definiált valószínűségi mértékeknek egy a τ topológia szerint zárt részhalmaza egyben kompakt is ezen topológia szerint. (Egyébként ez az egyenlőtlenség elemien is levezethető.) Innen logaritmust véve kapjuk, hogy

$$\liminf_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mu_n \in \mathbf{F}) \ge \inf_{\nu\in[\mathbf{F}_\mathcal{C}]} I\left(\nu_\mathcal{C}\|\mu_\mathcal{C}\right).$$

Szuprémumot véve $\mathcal{C} \in \mathcal{C}^*$-re kapjuk a kívánt állítást.

24.) Tekintsünk egy $L > \sup\limits_{\mathcal{C}\in\mathcal{C}^*} \inf\limits_{\nu_\mathcal{C}\in[\mathbf{F}_\mathcal{C}]} I\left(\nu_\mathcal{C}\|\mu_\mathcal{C}\right)$ számot. A (4) formulában szereplő egyenlőtlenség bizonyításához azt kell belátni, hogy ez az L szám felső becslés az egyenlőtlenség jobboldalán szereplő kifejezésre is. Ennek érdekében először lássuk be azt, hogy létezik olyan $\nu \in (X, \mathcal{A})$ mérték, amelyre $\nu_\mathcal{C} \in [\mathbf{F}_\mathcal{C}]$ minden $\mathcal{C} \in \mathcal{C}^*$ véges particióra, és $I(\nu\|\mu) \le L$. Viszont nem állítjuk, — legalábbis egyelőre, — hogy a $\nu \in \mathbf{F}$ reláció is teljesül.

A $\mathbf{D}_\mathcal{C} = \{\nu\colon\ \nu \in X,\ \nu_\mathcal{C} \in [\mathbf{F}_\mathcal{C}]\,,\ I(\nu\|\mu) \le L\}$ halmaz kompakt minden $\mathcal{C} \in \mathcal{C}^*$ véges particióra, mivel a 20. feladat alapján egy kompakt és egy zárt halmaz metszete. (A $\{\nu\colon\ \nu \in X, \nu_\mathcal{C} \in [\mathbf{F}_\mathcal{C}]\}$ halmaz zárt. Ez következik abból, hogy a $\nu \to \nu_\mathcal{C}$ leképezés folytonos.) Továbbá a $\mathbf{D}_\mathcal{C}$ halmaz nem üres, mivel létezik olyan $\nu_\mathcal{C} \in [\mathbf{F}_\mathcal{C}]$ mérték, amelyre $I(\nu_\mathcal{C}\|\mu_\mathcal{C}) \le L$, és arra a $\nu \in X$ mértékre, amelyet a $\nu(\mathbf{D}) = \frac{\mu(\mathbf{D})}{\mu(\mathbf{C}_l)}\nu_\mathcal{C}(x_l)$, $l = 1,\dots,k$, képlet definiál, ha $\mathbf{D} \subset \mathbf{C}_l$, ahol $\mathcal{C} = \{\mathbf{C}_1,\dots,\mathbf{C}_k\}$, és x_l a $\mathbf{C}_l$-hez rendelt pont, $I(\nu\|\mu) = I(\nu_\mathcal{C}\|\mu_\mathcal{C}) \le L$. A $I(\nu\|\mu) = I(\nu_\mathcal{C}\|\mu_\mathcal{C})$ azonosság azért igaz, mert $\frac{d\nu}{d\mu}(y) = \frac{\nu(\mathbf{C}_j)}{\mu(\mathbf{C}_j)}$ minden $y \in \mathbf{C}_j$ pontban. Továbbá az így definiált ν mérték „vetülete” a $\mathcal{C}$ particióra a fenti $\nu_\mathcal{C}$ mérték, amely eleme az $[\mathbf{F}_\mathcal{C}]$ halmaznak. Azt akarjuk belátni, hogy $\bigcap\limits_{\mathcal{C}\in\mathcal{C}^*} \mathbf{D}_\mathcal{C} \ne \emptyset$. A $\mathbf{D}_\mathcal{C}$ halmazok kompaktsága miatt elég azt belátni, hogy véges sok $\mathbf{D}_{\mathcal{C}_1}, \dots, \mathbf{D}_{\mathcal{C}_p}$ halmaz metszete nem üres. Legyen $\bar{\mathcal{C}}$ olyan véges partició, amely finomabb mint a $\mathcal{C}_1,\dots,\mathcal{C}_p$ particiók mindegyike, és legyen $\nu \in \mathbf{D}_{\bar{\mathcal{C}}}$. Azt állítjuk, hogy ekkor $\nu \in \mathbf{D}_{\mathcal{C}_l}$ minden $l = 1,\dots,p$-re. Ehhez azt kell belátni, hogy mivel $\mathcal{C}_l$ durvítása a $\bar{\mathcal{C}}$ particiónak, ha $\nu_{\bar{\mathcal{C}}} \in [\mathbf{F}_{\bar{\mathcal{C}}}]$, akkor $\nu_{\mathcal{C}_l} \in [\mathbf{F}_{\mathcal{C}_l}]$. Ez utóbbi állítás bizonyítása érdekében vegyük észre, hogy ha $\nu_{\bar{\mathcal{C}}} \in [\mathbf{F}_{\bar{\mathcal{C}}}]$, akkor létezik olyan $\nu^{(n)} \in X$, $n = 1,2,\dots$, mértéksorozat, amelyre $\nu^{(n)}_{\bar{\mathcal{C}}} \in \mathbf{F}_{\bar{\mathcal{C}}}$, azaz $\nu^{(n)}_{\bar{\mathcal{C}}}$ eleme a $\mathbf{F}_{\bar{\mathcal{C}}}$ halmaznak és nemcsak annak lezártjának, és $\lim\limits_{n\to\infty} \nu^{(n)}_{\bar{\mathcal{C}}}(x_m) = \nu_{\bar{\mathcal{C}}}(x_m)$ minden $m = 1,\dots,p$-re, ahol $\bar{\mathcal{C}} = \{\bar{\mathbf{C}}_1,\dots,\bar{\mathbf{C}}_p\}$, és x_m a $\bar{\mathbf{C}}_m$ halmazhoz rendelt pont. Mivel a $\mathcal{C}_l$ partició elemei ilyen (diszjunkt) $\bar{\mathbf{C}}_m$ halmazok uniói, innen következik, hogy $\nu^{(n)}_{\mathcal{C}_l} \in \mathbf{F}_{\mathcal{C}_l}$, és $\lim\limits_{n\to\infty} \nu^{(n)}_{\mathcal{C}_l} = \nu_{\mathcal{C}_l}$. Ezért $\nu_{\mathcal{C}_l} \in [\mathbf{F}_{\mathcal{C}_l}]$.

Legyen $\nu \in \mathbf{D}_\mathcal{C}$ minden $\mathcal{C} \in \mathcal{C}^*$ particióra. Ekkor, mivel $\nu_\mathcal{C} \in [\mathbf{F}_\mathcal{C}]$ minden $\mathcal{C} \in \mathcal{C}^*$ particióra, és $\mathbf{F}$ zárt halmaz, ezért $\nu \in \mathbf{F}$. Valóban, ha $\nu \notin \mathbf{F}$ lenne, akkor mint a 17. feladat megoldásában, az $(X,\mathcal{X})$ tér T_3 tulajdonságának bizonyításában láttuk, létezne olyan $\mathcal{C} = \{\mathbf{C}_1,\dots,\mathbf{C}_k\} \in \mathcal{C}^*$ partició, és $\varepsilon > 0$ szám, amelyekre az $\mathbf{U} = \{\chi\colon\ \chi \in X,\ |\chi(\mathbf{C}_l) - \nu(\mathbf{C}_l)| < \varepsilon$ minden $1 \le l \le k$-ra$\}$ halmaz a $\nu \in X$ pontnak, a $\mathbf{V} = \{\chi\colon\ \chi \in X,\ |\chi(\mathbf{C}_l) - \nu(\mathbf{C}_l)| > 2\varepsilon$ valamely $1 \le l \le k$-ra$\}$ halmaz az $\mathbf{F}$ zárt halmaznak a nyílt környezete, és $\mathbf{U} \cap \mathbf{V} = \emptyset$. Ez viszont ellentmond annak, hogy $\nu_\mathcal{C} \in [\mathbf{F}_\mathcal{C}]$ a $\mathcal{C} = \{\mathbf{C}_1,\dots,\mathbf{C}_k\}$ particióval.

A bizonyítandó egyenlőtlenség jobboldalán szereplő kifejezés felülről becsülhető a $\sup\limits_{\mathcal{C}\in\mathcal{C}^*} I\left(\nu_\mathcal{C}\|\mu_\mathcal{C}\right)$ mennyiséggel, ahol egy olyan alkalmasan választott $\nu \in \mathbf{F}$ mértéket veszünk, amelyre $\nu \in \mathbf{D}_\mathcal{C}$ minden $\mathcal{C} \in \mathcal{C}^*$ particióra. Mivel erre a ν mértékre $\sup\limits_{\mathcal{C}\in\mathcal{C}^*} I\left(\nu_\mathcal{C}\|\mu_\mathcal{C}\right) \le I(\nu\|\mu) \le L$ minden $\mathcal{C} \in \mathcal{C}^*$ particióra, innen következik a (4) formulában felírt egyenlőtlenség. Végül a (4) formulában felírt azonosság a 12. feladat következménye.

A 23. feladat eredményéből és a 24. feladatban már bebizonyított (4) formulájából következik a nagy eltérés tétel iii.) tulajdonsága a Szanov tétel feltételeinek teljesülése esetén.

25.) Mutassuk meg először azt, hogy a feladatban definiált $\mathbf{F}$ halmaz zárt, azaz, ha $\nu \notin \mathbf{F}$, akkor létezik a ν pontnak olyan $\mathbf{U}(\nu)$ környezete, amelyre $\mathbf{U}(\nu) \cap \mathbf{F} = \emptyset$. Minden $\nu \in X$ mértékre létezik olyan $\varepsilon > 0$ szám, amelyre $\nu(\{(0,\varepsilon)\}) \le \frac{1}{3}$.

Ezért az $\mathbf{F}$ halmaz és a $\nu \in X$ pont $\mathbf{U}_1(\nu) = \left\{\chi\colon |\chi(\{(0,\varepsilon)\}) - \nu(\{(0,\varepsilon)\})| < \frac{1}{3}\right\}$ környezetének a metszete véges. Legyen $N = [\varepsilon^{-1}]$, ahol $[\cdot]$ egész részt jelent. Ekkor $k > N$ indexre $\nu_k \notin \mathbf{U}_1(\nu)$. Ha $\nu \notin F$ akkor válasszunk minden $k \le N$ indexre olyan $\varepsilon_k > 0$ számot, amelyre vagy $\left|\nu\left(\left\{\frac{1}{k}\right\}\right) - \nu_k\left(\left\{\frac{1}{k}\right\}\right)\right| > \varepsilon_k$ vagy $\left|\nu\left(\left\{\frac{3}{4}\right\}\right) - \nu_k\left(\left\{\frac{3}{4}\right\}\right)\right| > \varepsilon_k$. Mivel $\nu \ne \nu_k$, ez lehetséges. Legyen $\varepsilon = \min\limits_{k \le N} \varepsilon_k$, és definiáljuk az

$$U_2(\nu) = \left\{\chi\colon \left|\chi\left(\left\{k^{-1}\right\}\right) - \nu\left(\left\{k^{-1}\right\}\right)\right| > \varepsilon, \text{ ha } k \le N\right\}$$

és

$$U_3(\nu) = \left\{\chi\colon \left|\chi\left(\left\{\frac{3}{4}\right\}\right) - \nu\left(\left\{\frac{3}{4}\right\}\right)\right| > \varepsilon\right\}$$

halmazokat. Ekkor az $U(\nu) = U_1(\nu) \cap U_2(\nu) \cap U_3(\nu)$ halmaz a ν mértéknek az $\mathbf{F}$ halmaztól diszjunkt környezete. Tehát az $\mathbf{F}$ halmaz zárt.

Az $\mathbf{F}_{\mathcal{C}}$ halmaz a következő $\bar\nu_k$ mértékekből áll: $\bar\nu_k(\{x_1\}) = 1 - \frac{1}{k}$, $\bar\nu_k(\{x_2\}) = \frac{1}{k}$. Ezek a mértékek konvergálnak ahhoz a $\bar\nu$ mértékhez, amelyre $\bar\nu(\{x_1\}) = 1$, és ez nem szerepel a $\bar\nu_k$ mértékek között. Innen következik, hogy az $\mathbf{F}_{\mathcal{C}}$ halmaz nem zárt.

26.) Tetszőleges $y_1, \dots, y_k$, $y_j \in Y$, $j = 1, \dots, k$ pontokra, $1 \le k \le n$, és $j_1, \dots, j_k$, $\sum\limits_{p=1}^{k} j_p = n$, pozitív egész számokra definiáljuk a

$$\begin{aligned}\mathbf{G}(y_1, \dots, y_k, j_1, \dots, j_k) &= \mathbf{G}(y_1, \dots, y_k, j_1, \dots, j_k, n)\\ &= \left\{\mu\colon\ \mu \in X,\ \left|\mu(\{y_p\}) - \frac{j_p}{n}\right| < \frac{1}{n}\right\}\end{aligned}$$

halmazokat. Adva egy $\mathbf{B} \subset Y$ halmaz, definiáljuk a

$$\mathbf{G} = \bigcup_{k=1}^{n} \bigcup_{((y_1,\dots,y_k),(j_1,\dots,j_k))\colon\ y_l \in \mathbf{B},\ 1 \le l \le k} \mathbf{G}(y_1, \dots, y_k, j_1, \dots, j_k)$$

halmazt. Ez a halmaz nyílt, mert minden $\mathbf{G}(y_1, \dots, y_k, j_1, \dots, j_k)$ halmaz nyílt. Továbbá a $\xi_1, \dots, \xi_n$ minta által meghatározott μ_n empirikus mértékre, $\mu_n \in \mathbf{G}$ akkor és csak akkor, ha az összes ξ_j valószínűségi változó értéke a $\mathbf{B}$ halmazban van. Ugyanis $\mu_n(\{y\}) = \frac{j}{n}$, $0 \le j \le n$ minden $y \in Y$-ra, és az összes ξ_j akkor és csak akkor esik az $\mathbf{A}$ halmazba, ha az általa meghatározott μ_n mérték eleme valamelyik a $\mathbf{G}$ definiciójában szereplő $\mathbf{G}(y_1, \dots, y_k, j_1, \dots, j_k)$ halmazban.

Legyenek $\xi_1, \cdots, \xi_n$ független, a $[0,1]$ intervallumon egyenletes eloszlású valószínűségi változók. Jelölje μ_n azt az empirikus mértéket, amelyet ezek a valószínűségi változók határoznak meg. Legyen $\mathbf{B} \subset [0,1]$ nem mérhető, 0 belső és 1 külső mértékű halmaz. Ekkor az az esemény, hogy mindegyik ξ_j, $1 \le j \le n$, valószínűségi változó ebbe a halmazba esik nem mérhető, és ennek az eseménynek nincs valószínűsége. Viszont az előbbi érvelésben megmutattuk, hogy létezik olyan nyílt $\mathbf{G}$ halmaz a valószínűségi mértékek terén, amelyre ez az esemény megegyezik a $\{\mu_n \in \mathbf{G}\}$ eseménnyel.

27.) Az egyszerűség miatt tekintsük először csak az $I(\mu_G \| \mu_F) < \infty$ esetet. Mivel $I(\cdot \| \mu_F)$ alulról félig folytonos függvény a 18. feladat eredménye szerint, ezért

$$\mathbf{H} = \mathbf{H}(\eta) = \{\mu_H \colon I(\mu_H \| \mu_F) > I(\mu_G \| \mu_F) - \eta\}$$

nyílt halmaz a valószínűségi mértékek terében minden $\eta > 0$ számra mind a τ mind a τ_0 topológia szerint, és $\mu_G \in \mathbf{H}$. (Itt azt a jelölést használtuk, hogy μ_H a H eloszlásfüggvény által indukált Stieltjes mérték a számegyenesen.)

Belátjuk, hogy alkalmas $\varepsilon_0 = \varepsilon_0(\eta) > 0$ küszöbindexre igaz, hogy minden $0 < \varepsilon < \varepsilon_0$ számra létezik olyan $\mathbf{G}_\varepsilon$ nyílt és $\mathbf{F}_\varepsilon$ zárt halmaz, amelyekre $\mu_G \in \mathbf{G}_\varepsilon$, $\mathbf{G}_\varepsilon \subset \mathbf{K}_\varepsilon \subset \mathbf{H}(\eta)$, és $\mathbf{K}_\varepsilon \subset \mathbf{F}_\varepsilon \subset \mathbf{H}(\eta)$, ahol $\mathbf{K}_\varepsilon = \{H \colon \sup\limits_x |H(x) - G(x)| < \varepsilon\}$. (A $\mathbf{K}_\varepsilon$ halmaz definiciójában azonosítottam a H és G eloszlásfüggvényt az általuk meghatározott μ_H és μ_G Stieltjes mértékkel.) Innen következik a feladat állítása.

Ennek megmutatása érdekében vegyük észre, hogy $\mathbf{G}_\varepsilon \subset \mathbf{K}_\varepsilon$ miatt

$$\limsup_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{G}_\varepsilon) \ge \limsup_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{K}_\varepsilon),$$

és $I(\mu_G \| \mu_F) \ge \limsup\limits_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{G}_\varepsilon)$ a nagy eltérés tétel ii.) tulajdonsága és a $\mu_G \in \mathbf{G}_\varepsilon$ reláció miatt. Innen

$$I(\mu_G \| \mu_F) \ge \limsup_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{K}_\varepsilon).$$

Hasonlóan, $\mathbf{K}_\varepsilon \subset \mathbf{F}_\varepsilon$ miatt

$$\liminf_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{K}_\varepsilon) \ge \liminf_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{F}_\varepsilon),$$

és $I(\mu_G \| \mu_F) - \eta \le \liminf\limits_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{F}_\varepsilon)$ a nagy eltérés tétel iii.) tulajdonsága és az $\mathbf{F}_\varepsilon \subset \mathbf{H}(\eta)$ reláció miatt. Innen

$$I(\mu_G \| \mu_F) - \eta \le \liminf_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{K}_\varepsilon).$$

Tehát

$$\begin{aligned} I(\mu_G \| \mu_F) &\ge \limsup_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{K}_\varepsilon) \\ &\ge \liminf_{n\to\infty} -\frac{1}{n} \log \mu^{(n)}(\mathbf{K}_\varepsilon) \ge I(\mu_G \| \mu_F) - \eta. \end{aligned}$$

Mivel $F_n \in \mathbf{K}_\varepsilon$ egy F_n n-elemű empirikus eloszlásfüggvényre akkor és csak akkor, ha $\sup_x |F_n(x) - G(x)| < \varepsilon$, valójában azt láttuk be, hogy

$$\begin{aligned} I(\mu_G \| \mu_F) &\geq \limsup_{n\to\infty} -\frac{1}{n} \log P\left(\sup_x |F_n(x) - G(x)| < \varepsilon\right) \\ &\geq \liminf_{n\to\infty} -\frac{1}{n} \log P\left(\sup_x |F_n(x) - G(x)| < \varepsilon\right) \geq I(\mu_G \| \mu_F) - \eta. \end{aligned}$$

Innen $\varepsilon \to 0$ majd $\eta \to 0$ határátmenetet alkalmazva megkapjuk a kívánt állítást.

Be kell még látnunk a felhasznált tartalmazási tulajdonságokat. Először azt mutatjuk meg, hogy $\mathbf{K}_\varepsilon \subset \mathbf{H}(\eta)$, ha $\varepsilon < \varepsilon_0(\eta)$. A bizonyításban azt használjuk fel, hogy $\mathbf{H}(\eta)$ nyílt halmaz a τ_0 topológiában, és $\mu_G \in \mathbf{H}(\eta)$. Innen következik, hogy létezik olyan $\mathbf{S} \subset \mathbf{H}(\eta)$, $\mu_G \in \mathbf{S}$ halmaz, amely $\mathbf{S} = \bigcap_j \{\mu_H\colon\ |\mu_H(B_j) - \mu_G(B_j)| < \bar\eta\}$, alakban írható, ahol egy véges metszetet vettünk, a metszetben szereplő minden B_j halmaz alkalmas intervallum vagy félegyenes, és $\bar\eta > 0$ elég kis pozitív szám. Itt azt használtuk ki, hogy milyen bázis segítségével adtuk meg a τ_0 topológiát.

Viszont könnyen látható, hogy $\{\mu_H\colon \sup_x |H(x) - G(x)| < \varepsilon\} \subset \mathbf{S}$, azaz, $\mathbf{K}_\varepsilon \subset \mathbf{S} \subset \mathbf{H}(\eta)$, ha $\varepsilon \leq \frac{\bar\eta}{2}$, mert $|\mu_H(B_j) - \mu_G(B_j)| < 2\varepsilon$, ha $\sup_x |H(x) - G(x)| < \varepsilon$. Sőt innen az is következik, hogy $\mathbf{K}_\varepsilon \subset \mathbf{F}_\varepsilon \subset \mathbf{H}(\eta)$, ahol $\mathbf{K}_\varepsilon = \{\mu_H\colon \sup_x |H(x) - G(x)| < \varepsilon\}$. Továbbá $\mathbf{K}_\varepsilon$ zárt halmaz a τ_0 vagy τ topológiában, mert a $\mathbf{K}_\varepsilon$ halmaz a $\{\mu_H\colon\ |H(x) - G(x)| \leq \varepsilon\}$ zárt halmazok metszete minden $-\infty < x < \infty$ számra. Be kell még látni, hogy minden $\varepsilon > 0$ számra létezik olyan $\mathbf{G}_\varepsilon$ nyílt halmaz, amelyre $\mu_G \in \mathbf{G}_\varepsilon \subset \mathbf{K}_\varepsilon$.

Itt hivatkozunk a 15. feladat megoldásának elején elvégzett konstrukcióra, amely szerint minden $\delta > 0$ számra létezik a számegyenesnesnek olyan véges particiója B_j halmazokra, amelyre minden B_j halmaz vagy egy pont vagy egy olyan nyílt intervallum, esetleg félegyenes, amelyre $\mu_G(B_j) < \delta$. Alkalmazzuk ezt az eredményt a $\delta = \frac{\varepsilon}{2}$ számra. Ekkor nem nehéz belátni, hogy elég kis $\bar\eta = \bar\eta(\varepsilon) > 0$ számra a $\mathbf{G}_\varepsilon = \bigcap_J \{\mu_H\colon\ |\mu_H(B_j) - \mu_G(B_j)| < \eta\}$ halmaz teljesíti a kívánt feltételeket.

Az $I(\mu_G \| \mu_F) = \infty$ eset hasonlóan tárgyalható, bár az kissé egyszerűbb.

Felhasznált topológiai ismeretek.

Egy $(X,\mathcal{U})$ topológikus tér egy X halmazból és annak $\mathcal{U}$ nyílt részhalmazaiból álló rendszer. A nyílt halmazok $\mathcal{U}$ rendszere teljesíti a következő tulajdonságokat:

1.) Az $\emptyset$ üres, és az X teljes halmaz nyílt halmaz.
2.) Véges sok nyílt halmaz metszete nyílt.
3.) Nyílt halmazok uniója mindig nyílt halmaz.

Egy nyílt halmaz komplementerét zárt halmaznak nevezik.

Egy topológikus tér definiciójában a nyílt halmazok $\mathcal{U}$ rendszerét gyakran nem közvetlenül definiálják, hanem megadják annak egy bázisát. Ismertetem ennek definicióját.

Topológikus tér bázisának a definiciója. *Egy $(X,\mathcal{U})$ topologikus térnek akkor és csak akkor bázisa egy az X halmaz bizonyos részhalmazaiból álló $\mathcal{B}$ halmazrendszer, ha $\mathcal{B} \subset \mathcal{U}$, és minden $U \in \mathcal{U}$ nyílt halmazhoz és $x \in U$ ponthoz, létezik egy olyan $B \in \mathcal{B}$ halmaz, amelyre $x \in B$, és $B \subset U$.*

Be lehet látni, hogy egy $(X,\mathcal{U})$ topológikus tér $\mathcal{B}$ bázisa egyértelműen meghatározza a $\mathcal{U}$ topológiát, (a $\mathcal{B}$-beli halmazok uniói a nyílt halmazok). Egy X tér részhalmazaiból álló $\mathcal{B}$ rendszer akkor és csak akkor lehet egy $(X,\mathcal{U})$ topológikus tér bázisa valamely alkalmas $\mathcal{U}$ topológiával, ha minden $B_1, B_2 \in \mathcal{B}$ halmazpárhoz, és $x \in B_1 \cap B_2$ ponthoz létezik olyan $B \in \mathcal{B}$ halmaz, amelyre $x \in B \subset B_1 \cap B_2$.

Be szokták vezetni az úgynevezett $\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$, $\mathbf{T}_\rho$ és $\mathbf{T}_4$ tulajdonságú topológikus tereket. Ezek a tulajdonságok azt fejezik ki, hogy egy $\mathcal{U}$ topológia mennyire tud szétválasztani bizonyos tulajdonságú diszjunkt halmazokat.

$\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$, $\mathbf{T}_\rho$, és $\mathbf{T}_4$ tulajdonságú topológikus terek definiciója. *Legyen adva egy $(X,\mathcal{U})$ topológikus tér. Az ebben szereplő $\mathcal{U}$ topológia $\mathbf{T}_1$ tulajdonságú, ha minden $x \in X$ pontra az $\{x\}$ halmaz zárt. Az $\mathcal{U}$ topológia $\mathbf{T}_2$ tulajdonságú, ha minden $x, y \in X$, $x \neq y$ pontpárra létezik olyan G és H nyílt halmaz az $\mathcal{U}$ topológiában, amelyre $x \in G$, $y \in H$, és $G \cap H = \emptyset$. (A $\mathbf{T}_2$ tulajdonságú tereket Hausdorff tereknek is nevezik az irodalomban.)*

A további definiciókban feltesszük, hogy $\mathbf{T}_1$ terekkel foglalkozunk, azaz olyan topológikus terekkel, amelyekben minden egy pontból álló halmaz zárt. Az $\mathcal{U}$ topológia $\mathbf{T}_3$ tulajdonságú, ha minden F zárt halmazra és $x \notin F$ pontra létezik olyan G és H nyílt halmaz, amelyekre $F \subset G$, $x \in H$, és $G \cap H = \emptyset$. Az $\mathcal{U}$ topológia $\mathbf{T}_\rho$ tulajdonságú, ha minden F zárt halmazra, és $x \notin F$ pontra létezik olyan folytonos $f(\cdot)$ függvény az $(X,\mathcal{U})$ téren, amelyre $f(u) = 0$, ha $u \in F$, és $f(x) = 1$. Az $\mathcal{U}$ topológia $\mathbf{T}_4$ tulajdonságú, ha minden olyan F_1 és F_2 zárt halmazpárra, amelyre $F_1 \cap F_2 = \emptyset$ létezik olyan G és H nyílt halmaz, amelyekre $F_1 \subset G$, $F_2 \subset H$ és $G \cap H = \emptyset$.

A fenti definicióban az egymást követő fogalmak a megelőző fogalmaknál (szigorúan) erősebb megkötést tartalmaznak.

A topológia egyik fontos fogalma a kompakt halmaz. Ennek definicióját, illetve a kompakt halmazok legfontosabb tulajdonságait ismertetem az alábbiakban. Különösen fontos lesz számunkra az alább megfogalmazandó Tyihonov tétel.

Egy $(X,\mathcal{U})$ topológikus tér $A \subset X$ halmaza kompakt, ha az A halmaz minden nyílt halmazokkal történő fedéséből kiválasztható egy véges fedés, azaz ha G_t, $t \in T$, valamely T indexhalmazzal nyílt halmazoknak egy olyan rendszere, amelyre $A \subset \bigcup\limits_{t \in T} G_t$, akkor létezik a T indexhalmaznak olyan $\{t_1, \dots, t_n\} \subset T$ véges részhalmaza, amelyre $A \subset \bigcup\limits_{j=1}^{n} G_{t_j}$. Egy $(X,\mathcal{U})$ topológikus tér kompakt, ha az X halmaz kompakt a benne szereplő $\mathcal{U}$ topológia szerint.

Egy $\mathbf{T}_2$ topológikus tér minden kompakt halmaza zárt. Kompakt halmaz zárt részhalmazai kompaktak. Szemléletesen mondva, a kompakt halmazok a kis zárt halmazok.

A topológia egyik fontos eredménye a Tyihonov tétel, amely szerint kompakt topológikus terek (Descartes) szorzata szintén kompakt topológikus tér. E tétel megfogalmazása érdekében definiáljuk topológikus terek (Descartes) szorzatát.

Először X_t, $t \in T$, halmazok $\prod\limits_{t \in T} X_t$ Descartes szorzatát definiáljuk az általános esetben, amikor a T indexhalmaz nem feltétlenül véges vagy megszámlálható. $\prod\limits_{t \in T} X_t$ az összes olyan $x(t)$, $t \in T$, függvényből áll, amelyekre $x(t) \in X_t$ minden $t \in T$ pontban.

Legyen adva topológikus tereknek egy$(X_t,\mathcal{U}_t)$, $t \in T$, halmaza. Ezek Descartes szorzata az az $(X,\mathcal{U})$ topológikus tér, amelyre $X = \prod\limits_{t \in T} X_t$, és a nyílt halmazok $\mathcal{U}$ halmazának egy $\mathcal{B}$ bázisát a következő módon definiáljuk. $V \in \mathcal{B}$ akkor és csak akkor, ha létezik olyan véges $S = \{t_1, \dots, t_n\} \subset T$ halmaz, és $U_{t_j} j \in \mathcal{U}_{t_j}$, az $(X_{t_j},\mathcal{U}_{t_j})$ térben nyílt halmazok rendszere, $1 \le j \le n$, amelyekre $V = \prod\limits_{t \in T} Y_t$, ahol $Y_t = U_t$, ha $t \in S$, és $Y_t = X_t$, ha $t \in T \setminus S$. Ezután megfogalmazom a Tyihonov tételt.

Tyihonov tétel. *Legyen adva kompakt topológikus tereknek egy $(X_t,\mathcal{U}_t)$, $t \in T$, halmaza. Ezek Descartes szorzata szintén kompakt topológikus tér.*

Még egy eredményre lesz szükségünk. Ez a következő állítás. Ha $(X,\mathcal{U})$ kompakt topológikus tér, és azonkívül még $\mathbf{T}_2$ tér is, akkor egyben $\mathbf{T}_4$ tér is.

Bizonyításvázlat. Először azt látjuk be, hogy $(X,\mathcal{U})$ $\mathbf{T}_3$ tér. Legyen $F \subset X$ zárt, ezért kompakt halmaz, és $x \in X \setminus F$. Akkor minden $y \in F$ ponthoz lehet találni olyan G_y és H_y nyílt halmazokból álló párt, amelyre $G_y \cap H_y = \emptyset$, $y \in G_y$, és $x \in H_y$. A G_y halmazok uniója lefedi az F kompakt halmazt, ezért létezik véges sok $G_{y_1}, \dots, G_{y_k}$ halmaz, amelyre $G = \bigcup\limits_{j=1}^{k} G_{y_j}$ nyílt halmaz, és $F \subset G$. Másrészt $H = \bigcap\limits_{j=1}^{k} H_{y_j}$ is nyílt halmaz, $x \in H$, és $G \cap H = \emptyset$. Innen következik a $\mathbf{T}_3$ tulajdonság.

Az $(X,\mathcal{U})$ tér $\mathbf{T}_4$ tulajdonsága hasonlóan látható be annak $\mathbf{T}_3$ tulajdonsága alapján. Két diszjunkt zárt F_1 és F_2 halmazt kell venni, és az $F_1 = F$ választással és az F_2 halmazzal úgy érvelve, mint korábban az x ponttal kapjuk a kívánt állítást.