A Valószínűségszámítás II. előadássorozat hetedik témája.

FELTÉTELES VALÓSZÍNŰSÉG ÉS FELTÉTELES VÁRHATÓ ÉRTÉK

A feltételes valószínűség és feltételes várható érték fogalmát nulla valószínűséggel bekövetkező feltételek esetén is definiálják, tehát olyan esetekben is, amikor a hagyományos, a bevezető valószínűségszámítás előadásban tanult definiciónak nincs értelme. De e fogalmak általánosításai az ilyen esetekre a valószínűségszámítás legnehezebb fogalmai közé tartoznak. Ezek jobb megértése érdekében megpróbálom először megmutatni egy példa segítségével, hogy milyen szemléletes képet akarunk ennek a definiciónak a segítségével megfogalmazni. Tekintsük a következő példát.

Van tíz darab lámpánk, ezek élettartama egymástól független, (exponenciális eloszlással és) 25 óra várható értékkel. Egy foglalatba betesszük az első lámpát, hogy bevilágítson egy termet. Ha a lámpa kiégett azonnal kicseréljük a következő lámpára. Első kérdés: Mit várunk várunk, mennyi ideig elegendő a tíz lámpa együttesen a terem bevilágításához? A természetes válasz az, hogy ez a tíz lámpa együttes élettartamának a várható értéke, azaz $10\times25=250$ óra. A második kérdés a következő: Megfigyeljük, hogy melyik időpontban cseréljük ki a második lámpát. Mit várunk ennek az információnak az ismeretében a tíz lámpa együttes élettartamára? Ha ez a csere 48 óra 22 perc múlva történik meg, akkor a természetes becslés a 10 lámpa együttes élettartamára 48 óra 22 perc plusz 8×25 óra, azaz 248 óra 22 perc. Ha ez a csere 51 óra 19 perc múlva történik, akkor hasonlóan azt várjuk, hogy ez az időtartam 251 óra 19 perc.

A fenti példa nem önmaga miatt érdekes számunkra, hanem azért, mert rámutat arra, hogy természetes vizsgálni a következő kérdést. Ha adva van egy esemény vagy egy valószínűségi változó, akkor érdekelhet minket ennek az eseménynek a valószínűsége vagy ennek a valószínűségi változónak a várható értéke. Előfordulhat, hogy más eseményeknek a bekövetkezését vagy be nem következését, más valószínűségi változók felvett értékét meg tudjuk figyelni, és ezen plusz információk ismeretében akarjuk megbecsülni a minket érdeklő esemény valószínűségét vagy valószínűségi változó várható értékét. Ekkor természetes olyan becslést adni, amely ezeket a plusz információkat is figyelembe veszi. A kérdés az, hogy hogyan tegyük ezt. Az előző példában is ilyen kérdést fogalmaztam meg. Ott meg akartuk becsülni tíz valószínűségi változó összegének a várható értékét azon feltétel mellett, hogy az első két változó összegének az értéke ismert. Olyan példát tekintettünk, ahol a megfigyelt valószínűségi változó, az első két lámpa összélettartama, folytonos eloszlású valószínűségi változó, ezért nulla annak a valószínűsége, hogy az egy előírt értéket vesz fel. Tehát a keresett valószínűségi változó feltételes várható értékére olyan feltétel mellett vagyunk kiváncsiak, hogy egy nulla valószínűségi esemény következett be. Viszont a bevezető valószínűség előadásban definiált feltételes valószínűség (és feltételes várható érték) csak abban az esetben volt értelmes, ha a feltétel valószínűsége nem nulla.

Szeretnénk a feltételes valószínűség és feltételes várható érték fogalmának olyan általánosítását adni, amely lehetővé teszi, hogy a feltételes valószínűségről és feltételes várható értékről beszélhessünk nulla valószínűségi feltételek mellett is. Emellett azt is elvárjuk, hogy a bevezetett fogalmak megfeleljenek szemléletes képünknek, amelyet, legalábbis egyelőre, csak homályosan tudunk megfogalmazni. Lehetséges egy ilyen definiciót adni, de ehhez szükségünk van a mértékelmélet néhány mély eredményére. Ennek az előadásnak a célja a feltételes valószínűség és várható érték definiciójának az ismertetése az általános esetben, illetve e fogalmak legfontosabb tulajdonságainak a tárgyalása.

Az előző példa azt sugallja, hogy egy A halmaz feltételes valószínűségét feltéve bizonyos $\xi_1,\dots,\xi_k$ valószínűségi változók értékét úgy érdemes definiálni, mint a $\xi_1,\dots,\xi_k$ valószínűségi változók alkalmas (Borel) mérhető függvényét, azaz

$$P(A|\xi_1=x_1,\dots,\xi_k=x_k)=f_A(x_1,\dots,x_k),$$

ahol $f_A(x_1,\dots,x_k)$ Borel mérhető függvény az R^k k-dimenziós euklideszi térben, és azt méri mennyire valószínű az A esemény feltéve, hogy $\xi_1=x_1,\dots,\xi_k=x_k$. Ezt a valószínűséget implicit módon tudjuk definiálni. Azt várjuk a $P(A\cap B)=P(B)P(A|B)$ azonosság analógiájára, hogy bármely A halmazra és

$$B=\{\omega\colon\ \xi_1(\omega),\dots,\xi_k(\omega))\in[x_1,x_1+dx_1]\times\cdots\times[x_k,x_k+dx_k]\}$$

alakú B halmazra teljesül a

$$\begin{aligned}&P\left((\xi_1,\dots,\xi_k)\in[x_1,x_1+dx_1]\times\cdots\times[x_k,x_k+dx_k]\cap A\right)\\&\qquad=P\left((\xi_1,\dots,\xi_k)\in[x_1,x_1+dx_1]\times\cdots\times[x_k,x_k+dx_k]\right)f_A(x_1,\dots,x_k)\end{aligned}$$

azonosság. Ez az azonosság azonban semmitmondó, mert annak mind a két oldala nullával egyenlő. Viszont tartalmas állítássá válhat, ha megfelelő módon kiintegráljuk. Alkalmas integrálás azt sugallja, hogy teljesülnie kell a

$$\begin{aligned}&P\left(A\cap\{(\xi_1,\dots,\xi_k)\in B\}\right)\\&\qquad=\int_{(x_1,\dots,x_k)\in B}f_A(x_1,\dots,x_k)P\left(\xi_1\in[x_1,x_1+dx_1],\dots,\xi_k\in[x_k,x_k+dx_k]\right)\\&\qquad=\int_{(x_1,\dots,x_k)\in B}f_A(x_1,\dots,x_k)F(\,dx_1,\dots,\,dx_k)\end{aligned}\tag{$*$}$$

azonosságnak, ahol B az R^k k-dimenziós tér tetszőleges (Borel) mérhető halmaza, $F(x_1,\dots,x_k)$ pedig a k-dimenziós $(\xi_1,\dots,\xi_k)$ valószínűségi változó eloszlásfüggvénye. Az analízis egyik fontos eredményéből, a később ismertetendő Radon–Nikodym tételből következik, hogy létezik olyan $f_A(\cdot,\cdot,\cdot)$ függvény amely teljesíti a $(*)$ azonosságot minden mérhető B halmazra, és ezek az azonosságok lényegében egyértelműen meghatározzák ezt az f_A feltételes valószínűséget. Az előző mondatban használt 'lényegében egyértelműen' kifejezés értelmének a pontos magyarázatára később még visszatérek.

Hasonló gondolatmenet segítségével definiálhatjuk egy η, $E|\eta|<\infty$, valószínűségi változó $g_\eta(x_1,\dots,x_k)=E(\eta|\xi_1=x_1,\dots,\xi_k=x_k)$ feltételes várható értékét feltéve a $\xi_1=x_1,\dots,\xi_k=x_k$ feltételeket. Ez olyan (Borel-mérhető) $g_\eta(x_1,\dots,x_k)$ függvény, amelyre az

$$\begin{aligned}E\eta(\omega)I\left(\{\omega\colon\ (\xi_1(\omega),\dots,\xi_k(\omega))\in B\}\right)&=\int_{\{\omega\colon\ (\xi_1(\omega),\dots,\xi_k(\omega))\in B\}}\eta(\omega)\,dP(\omega)\\&=\int_B g_\eta(x_1,\dots,x_k)F(dx_1,\dots,dx_k)\end{aligned}\tag{$**$}$$

relációk teljesülnek a k-dimenziós R^k euklideszi tér tetszőleges B Borel mérhető részhalmazára, ahol $I(B)$ jelöli egy $B \subset \Omega$ halmaz indikátor függvényét, és $F(x_1, \dots, x_k)$ a $(\xi_1, \dots, \xi_k)$ véletlen vektor eloszlásfüggvénye. Azt, hogy ilyen $g_\eta(x_1, \dots, x_k)$ függvény valóban létezik, és ez a g_η függvény az $F(x_1, \dots, x_k)$ eloszlásfüggvény által meghatározott Lebesgue–Stieltjes mérték szerint egy valószínűséggel meg van határozva szintén a Radon–Nikodym tétel segítségével láthatjuk be.

Meg fogom mutatni, hogy a fent vázolt logika és a Radon–Nikodym tétel segítségével bevezethetjük a feltételes valószínűség és feltételes várható érték egy általános, és szemléletes elvárásainknak megfelelő definicióját. Sőt, ezeket a fogalmakat némileg általánosabb esetben fogom definiálni. Előfordulhat ugyanis, hogy azon előzetes információink, amelyek alapján egy halmaz valószínűségére vagy egy valószínűségi változó értékére becslést akarunk adni nem tömöríthetőek véges sok valószínűségi változó megfigyelt értékébe. Annak érdekében, hogy a feltételes valószínűség és feltételes várható érték fogalmát természetes módon tudjuk tárgyalni ilyen általánosabb esetben is, először azt kell tisztáznunk, hogy mit jelent az általános esetben az, hogy bizonyos megfigyelt események függvényeként akarunk valamit megbecsülni.

Ha meg tudunk figyelni megszámlálható sok eseményt, akkor meg tudjuk figyelni ezek unióját, metszetét, illetve mindegyik esemény komplementerét. Ez azt jelenti, hogy a megfigyelhető események σ-algebrát alkotnak. Ezért az általános kérdés úgy fogalmazható meg, hogy adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn egy $\mathcal{F} \subset \mathcal{A}$ σ-algebra, (amely tartalmazza az általunk megfigyelt eseményeket, amelyek mindegyikéről tudjuk, hogy bekövetkezett-e vagy sem) és egy A esemény vagy egy η valószínűségi változó, amelyre $E|\eta| < \infty$, akkor hogyan definiáljuk a $P(A|\mathcal{F})(\omega)$ feltételes valószínűséget illetve $E(\eta|\mathcal{F})(\omega)$ feltételes várható értéket feltéve az $\mathcal{F}$ σ-algebrát. Az, hogy az $\mathcal{F}$ σ-algebra ismeretében akarjuk megbecsülni az A halmaz valószínűségét illetve η valószínűségi változó értékét a $P(A|\mathcal{F})(\omega)$ és $E(\eta|\mathcal{F})(\omega)$ feltételes valószínűség definiciójában azt jelenti, hogy

i.) A $P(A|\mathcal{F})(\omega)$ feltételes valószínűség $\mathcal{F}$ mérhető valószínűségi változó.

i.′) Az $E(\eta|\mathcal{F})(\omega)$ feltételes várható érték $\mathcal{F}$ mérhető valószínűségi változó.

Az $\mathcal{F}$ σ-algebra szerinti feltételes valószínűség és feltételes várható érték definicióját a $(*)$ képlethez vezető érveléshez hasonlóan a következő módon próbáljuk definiálni a $P(A|\mathcal{F})(\omega)$ feltételes valószínűséget és $E(\eta|\mathcal{F})(\omega)$ feltételes várható értéket.

ii.) $P(A \cap B) = \int_B P(A|\mathcal{F})(\omega)\, dP(\omega)$ minden $\mathcal{F}$ mérhető B halmazra.

ii.′) $\int_B \eta(\omega)\, dP(\omega) = \int_B E(\eta|\mathcal{F})(\omega)\, dP(\omega)$ minden $\mathcal{F}$ mérhető B halmazra.

Ezután bevezetem a feltételes valószínűség és feltételes várható érték definicióját az általános esetben.

A feltételes valószínűség és feltételes várható érték definiciója. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőnek egy $\mathcal{F} \subset \mathcal{A}$ rész-σ-algebrája. Legyen továbbá adva egy $A \in \mathcal{A}$ esemény vagy egy $\eta(\omega)$, $E|\eta(\omega)| < \infty$, valószínűségi változó ezen a valószínűségi mezőn. Az A esemény $P(A|\mathcal{F})(\omega)$ feltételes valószínűsége feltéve a $\mathcal{F}$ σ-algebrát olyan* *valószínűségi változó, amely teljesíti az i.) és ii.) tulajdonságokat. Az $\eta(\omega)$ valószínűségi változó $E(\eta|\mathcal{F})(\omega)$ feltételes várható értéke feltéve a $\mathcal{F}$ σ-algebrát olyan valószínűségi változó, amely teljesíti az i.′) és ii.′) tulajdonságokat.*

Az előző definició mintájára bevezetem az előadás elején tárgyalt feltételes valószínűség és feltételes várható érték fogalmát feltéve bizonyos valószínűségi változók értékeinek az ismeretét.

A feltételes valószínűség és feltételes várható érték definiciója feltéve bizonyos valószínűségi változók értékét. *Legyen adva egy A esemény vagy egy η valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn valamint véges sok $\xi_1, \dots, \xi_k$ valószínűségi változó ugyanezen a valószínűségi mezőn. Az A esemény feltételes valószínűsége feltéve, hogy $\xi_1 = x_1, \dots, \xi_k = x_k$ egy olyan $f_A(x_1, \dots, x_k)$ k-változós Borel mérhető függvény, amely teljesíti a $(*)$ relációt. Egy η, $E|\eta| < \infty$, szintén az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált valószínűségi változó feltételes várható értéke feltéve, hogy $\xi_1 = x_1, \dots, \xi_k = x_k$ egy olyan $g_\eta(x_1, \dots, x_k)$ k-változós Borel mérhető függvény, amely teljesíti a $(**)$ relációt.*

Meg kell mutatni, hogy a fenti tulajdonságokkal rendelkező $P(A|\mathcal{F})(\omega)$ feltételes valószínűség és $E(\xi|\mathcal{F})(\omega)$ feltételes várható érték valóban létezik, illetve, hogy ezek a tulajdonságok meghatározzák őket. Meg fogom mutatni, hogy ez következik az alábbi megfogalmazandó Radon–Nikodym tételből. Ezenkívül meg kívánjuk érteni az általános esetben defininiált $P(A|\mathcal{F})$ illetve $E(\eta|\mathcal{F})$ feltételes valószínűség és feltételes várható érték kapcsolatát az előzőleg speciális esetben definiált $P(A|\xi_1 = x_1, \dots, \xi_k = x_k)$ és $E(\eta|\xi_1 = x_1, \dots, \xi_k = x_k)$ kifejezésekkel, amelyek létezését szintén meg kell indokolni.

Először megfogalmazok egy kérdést, amelyre a Radon–Nikodym tétel ad választ. Erre az eredményre szükségünk van a feltételes valószínűség és várható érték létezésének az indoklásában is.

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

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

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

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

Radon–Nikodym tétel. *Legyen adva egy Ω tér, rajta egy $\mathcal{F}$ σ-algebra, továbbá ezen az $\mathcal{F}$ σ-algebrán egy μ (σ-véges) mérték és ν (véges) előjeles mérték. (Az, hogy egy ν előjeles mérték véges azt jelenti, hogy $|\nu(A)| < \infty$ minden $A \in \mathcal{F}$ halmazra.) Tegyük fel, hogy a ν előjeles mérték abszolut folytonos a μ mértékre nézve. Akkor létezik olyan az Ω téren definiált valós értékű $\mathcal{F}$ mérhető $f(\omega)$ függvény, amelyre $\int |f(\omega)|\,d\mu(\omega) < \infty$, és $\nu(C) = \int_C f(\omega)\,d\mu(\omega)$ minden $C \in \mathcal{F}$ halmazra. Továbbá ez az $f(\omega)$ függvény egyértelműen meg van határozva a következő értelemben. Ha két $f_1(\omega)$ és $f_2(\omega)$ $\mathcal{F}$ mérhető függvény teljesíti a fenti relációt minden $C \in \mathcal{F}$ halmazra, akkor $f_1(\omega) = f_2(\omega)$ a μ mérték szerint majdnem minden $\omega \in \Omega$ pontban.*

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

Következmény. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, azon egy $\mathcal{F} \subset \mathcal{A}$ σ-algebra, és egy $\eta(\omega)$, $E|\eta(\omega)| < \infty$ valószínűségi változó vagy $A \in \mathcal{A}$ esemény. Létezik a definció követelményeit teljesítő $E(\eta|\mathcal{F})(\omega)$ feltételes várható érték, illetve $P(A|\mathcal{F})(\omega)$ feltételes valószínűség, és az $E(\eta|\mathcal{F})(\omega)$ illetve $P(A|\mathcal{F})(\omega)$ valószínűségi változók egy valószínűséggel meg van határozva.*

A következmény indoklása. Tekintsük a ν, $\nu(B) = \int_B \eta(\omega)\,dP(\omega)$, $B \in \mathcal{F}$, illetve $\bar{\nu}$, $\bar{\nu}(B) = P(A \cap B)$, $B \in \mathcal{F}$, előjeles mértékeket valamint a P valószínűségi mértéket az $(\Omega, \mathcal{F})$ mérhető téren. (Fontos, hogy az $\mathcal{F} \subset \mathcal{A}$ és nem az $\mathcal{A}$ σ-algebrát tekintjük.) Mind a ν, mind a $\bar{\nu}$ mérték abszolut folytonos a P mértékre nézve. Ezért a Radon–Nikodym tétel alapján definiálhatjuk a ν, illetve $\bar{\nu}$ előjeles mértékek $\frac{d\nu}{dP}(\omega)$, illetve $\frac{d\bar{\nu}}{dP}(\omega)$ Radon–Nikodym deriváltját az $(\Omega, \mathcal{F})$ téren. Ekkor az $E(\eta|\mathcal{F})(\omega) = \frac{d\nu}{dP}(\omega)$ és $P(A|\mathcal{F})(\omega) = \frac{d\bar{\nu}}{dP}(\omega)$ valószínűségi változók teljesítik az $E(\eta|\mathcal{F})(\omega)$ feltételes várható értékre illetve $P(B|\mathcal{F})(\omega)$ feltételes valószínűségre előírt feltételeket. Az, hogy a feltételes várható érték és feltételes valószínűség egy valószínűséggel meg van határozva, könnyen látható.

Azt kell kihasználni, hogy ha egy az $(\Omega, \mathcal{F}, P)$ téren definiált $\mathcal{F}$ mérhető ζ valószínűségi változóra $\int_B \zeta(\omega)\, dP(\omega) = 0$ minden $B \in \mathcal{F}$ halmazra, akkor a ζ valószínűségi változó egy valószínűséggel nulla. Innen következik, hogy ha ζ_1 és ζ_2 teljesíti az $E(\eta|\mathcal{F})$ feltételes várható értékre előírt feltételeket akkor a $\zeta_1(\omega) - \zeta_2(\omega)$ különbség egy valószínűséggel nulla.

1. megjegyzés: Legyen adva egy $(\Omega, \mathcal{A})$ mérhető tér, azon egy (σ-additív) μ mérték. A Radon–Nikodym tételben egy hasznos reprezentációját adtuk a μ mérték szerint abszolut folytonos ν előjeles mértéknek. Léteznek a μ mérték szerint nem abszolut folytonos előjeles mértékek is. Vannak például úgynevezett a μ mértékre nézve szinguláris (előjeles) mértékek, amelyek egy a μ mérték szerint nulla mértékű halmazba vannak koncentrálva. Ez azt jelenti, hogy létezik az Ω térnek olyan $X \subset \Omega$ részhalmaza, amelyre $\mu(X) = 0$, és minden $B \subset \Omega \backslash X$ halmazra $\nu(B) = 0$. Tetszőleges ν véges előjeles mérték az $(\Omega, \mathcal{A})$ téren egyértelműen felbontható $\nu = \nu_1 + \nu_2$ alakban, ahol ν_1 a μ mérték szerint abszolut folytonos, ν_2 pedig a μ mértékre szinguláris előjeles mérték.

Tekintsük azt a speciális esetet, amikor μ a Lebesgue mérték a számegyenesen. Ekkor a μ mértékre szinguláris mértékre példa egy véges vagy megszámlálható sok pontba koncentrált mérték. Vannak bonyolultabb a Lebesgue mértékre szinguláris mértékek is, olyanok, amelyek szerint minden pont mértéke nulla. A szinguláris mértékek minél pontosabb leírása fontos kérdése a mértékelméletnek. De mivel ez nem tartozik a valószínűségszámítás témakörébe, ezért ezzel a problémával itt nem foglalkozunk.

2. megjegyzés: A Radon–Nikodym tétel tipikus egzisztencia tétel. Nem ismert explicit módszer a Radon–Nikodym derivált kiszámolására az általános esetben. Ez a fő oka annak, hogy a feltételes valószínűséggel és feltételes várható értékkel csak nehezen tudunk számolni. Van azonban egy fontos speciális eset, amikor a feltételes valószínűségekkel és feltételes várható értékekkel jól tudunk számolni. Ha olyan valószínűségi változók függvényeivel dolgozunk, amelyeknek létezik együttes sűrűségfüggvénye, és azt ismerjük, akkor explicit módon ki tudjuk számolni a minket érdeklő feltételes valószínűségeket és feltételes várható értékeket. (A sűrűségfüggvény ismerete úgy értendő ebben az esetben, hogy mindazon valószínűségi változók együttes sűrűségfüggvényét ismerjük, amelyek vagy a feltételben szerepelnek vagy a vizsgálandó esemény vagy valószínűségi változó tőlük függ.) A matematikai statisztikában ilyen jellegű problémákkal találkozunk. Erre a kérdésre később még visszatérek.

A $g_\eta(x_1, \dots, x_k) = E(\eta|\xi_1 = x_1, \dots, \xi_k = x_k)$ feltételes várható érték létezését az $E(\eta|\mathcal{F})$ feltételes várható érték létezéséhez hasonlóan igazolhatjuk, csak ekkor más μ mértékkel és ν előjeles mértékkel kell dolgozunk. Ekkor mind a μ mértéket mind a ν előjeles mértéket az R^k k-dimenziós euklideszi tér Borel mérhető halmazain definiáljuk. A μ mérték a $(\xi_1, \dots, \xi_k)$ véletlen vektor eloszlása az R^k tér $B \in \mathcal{B}$ Borel mérhető részhalmazain, azaz $\mu(B) = P((\xi_1, \dots, \xi_k) \in B)$, $B \in \mathcal{B}$, és

$$\nu(B) = \int_{\{\omega\colon\, (\xi_1(\omega), \dots, \xi_k(\omega)) \in B\}} \eta(\omega)\, dP(\omega)$$

minden $B \in \mathcal{B}$ halmazra. A Radon–Nikodym tétel alapján létezik olyan $g_\eta(x_1, \dots, x_k)$ Borel mérhető függvény a k-dimenziós euklideszi téren, amelyre

$$\nu(B) = \int_B g_\eta(x_1, \dots, x_k)\mu(\,dx_1, \dots, \,dx_k), \quad B \in \mathcal{B}.$$

Be lehet látni, hogy ez a g_η függvény a $g_\eta(x_1, \dots, x_k) = E(\eta|\xi_1 = x_1, \dots, \xi_k = x_k)$ feltételes várható érték, azaz $g_\eta(x_1, \dots, x_k)$ egy olyan Borel mérhető függvény, amely teljesíti a $(**)$ relációt minden $B \in \mathcal{B}$ halmazra. Valóban,

$$\int_B g_\eta(x_1, \dots, x_k)\mu(\,dx_1, \dots, \,dx_k) = \int_B g_\eta(x_1, \dots, x_k)\,dF(x_1, \dots, x_k),$$

ahol $F(x_1, \dots, x_k)$ a $(\xi_1, \dots, \xi_k)$ eloszlásfüggvénye, és $\nu(B)$ a $(**)$ baloldalán szereplő mennyiség. Továbbá azt sem nehéz belátni, hogy a $g_\eta(x_1, \dots, x_k)$ függvény a $\xi_1, \dots, \xi_k$ valószínűségi változók μ eloszlása szerint 1 valószínűséggel meg van határozva.

A feltételes valószínűség fogalmával nem kell külön foglalkoznunk, mert a feltételes valószínűség és feltételes várható érték definiciójából következik, hogy tetszőleges mérhető A halmazra és annak $I_A(\omega)$ indikátor függvényére $P(A|\mathcal{F})(\omega) = E(I_A(\omega)|\mathcal{F})(\omega)$ és $P(A|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k) = E(I_A(\omega)|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k)$.

Meg kívánom tárgyalni az $E(\eta|\mathcal{F})(\omega)$ és

$$g_\eta(x_1, \dots, x_k) = E(\eta|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k)$$

feltételes várható értékek közötti kapcsolatot abban az esetben, ha $\mathcal{F} = \sigma(\xi_1, \dots, \xi_k)$, a $\xi_1, \dots, \xi_k$ valószínűségi változók által generált σ-algebra, azaz a legszűkebb olyan σ-algebra, amelyre nézve az összes $\xi_1(\omega), \dots, \xi_k(\omega)$ valószínűségi változó mérhető függvény. A két fogalom közötti kapcsolatot a következő tételben fogalmazom meg.

Tétel a feltételes várható érték különböző definiciói közötti kapcsolatról. *Legyen adva véges sok $\xi_1, \dots, \xi_k$ valamint rajtuk kívül egy η, $E|\eta| < \infty$, valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, és jelölje $\mathcal{F}$ a $\xi_1, \dots, \xi_k$ valószínűségi változók által generált σ-algebrát. Tekintsük a korábban definiált $E(\eta|\mathcal{F})$ és $g_\eta(x_1, \dots, x_k) = E(\eta|\xi_1 = x_1, \dots, \xi_k = x_k)$ feltételes várható értékeket. Ezek teljesítik az alábbi azonosságot.*

$$E(\eta|\mathcal{F})(\omega) = g_\eta(\xi_1(\omega), \dots, \xi_k(\omega)) \quad \textit{1 valószínűséggel.}$$

Megjegyzés. Az egyszerűség és áttekinthetőbb fogalmazás érdekében csak véges sok $\xi_1, \dots, \xi_k$ valószínűségi változó esetében definiáltam a feltételes valószínűséget és feltételes várható értéket a $\xi_1 = x_1, \dots, \xi_k = x_k$ alakú feltételek mellett. Definiálhattam volna e fogalmakat akkor is, ha a feltételek végtelen sok valószínűségi változó előírt értékeit tartalmazzák, és megfogalmazhattam volna a megfelelő eredményeket ebben az esetben is. A megfogalmazás és jelölés ekkor bonyolultabb lett volna, de nem lépett volna fel új matematikai nehézség.

A tétel bizonyítása felhasznál bizonyos mértékelméleti eredményeket, amelyeket az alábbi tételben fogalmazok meg. E tétel bizonyítását a kiegészítésben adom meg.

Tétel valószínűségi változók mérhető függvényeinek a jellemzéséről. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és azon $\xi_1(\omega), \dots, \xi_k(\omega)$ valószínűségi változók. Jelölje $\mathcal{F}$ a $\xi_1(\omega), \dots, \xi_k(\omega)$ valószínűségi változók által generált σ-algebrát. Egy η valószínűségi változó akkor és csak akkor mérhető erre az $\mathcal{F}$ σ-algebrára, ha létezik a k-dimenziós R^k euklideszi térben olyan Borel mérhető $f(x_1, \dots, x_k)$ függvény, amelyre $\eta(\omega) = f(\xi_1(\omega), \dots, \xi_k(\omega))$. Abban a speciális esetben, amikor $\eta(\omega)$ egy $B \in \mathcal{F}$ halmaz indikátorfüggvénye az $\eta(\omega) = f(\xi_1(\omega), \dots, \xi_k(\omega))$ előállításban szereplő f választható úgy, mint az R^k Euklideszi tér egy Borel–mérhető B_1 halmzának indikátor függvénye. Továbbá, ha két k-változós $f_1(x_1, \dots, x_k)$ és $f_2(x_1, \dots, x_k)$ Borel mérhető függvényre $f_1(\xi_1, \dots, \xi_k) = f_2(\xi_1, \dots, \xi_k)$ egy valószínűséggel, akkor $f_1(x_1, \dots, x_k) = f_2(x_1, \dots, x_k)$ a $(\xi_1, \dots, \xi_k)$ véletlen vektor μ eloszlása szerint majdnem minden $(x_1, \dots, x_k)$ pontban.*

A feltételes várható érték különböző definiciói közötti kapcsolatról szóló tétel bizonyítása. Mind az $E(\eta|\mathcal{F})(\omega)$ mind a $g_\eta(\xi_1(\omega), \dots, \xi_k(\omega))$ valószínűségi változó $\mathcal{F}$ mérhető. Ezért elég azt igazolni, hogy $g_\eta(\xi_1(\omega), \dots, \xi_k(\omega))$ teljesíti a ii.′) relációt. Ez ugyanis azt jelenti, hogy ezt a valószínűségi változót is tekinthetjük az $E(\eta|\mathcal{F})$ feltételes várható érték egyik verziójának.

A kivánt állítás igazolásának érdekében tekintsük az Ω halmaz következő T transzformációját az R^k euklideszi térbe: $T(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))$. Vegyük észre, hogy T mértéktartó transzformációja az $(\Omega, \mathcal{A}, P)$ térnek az $(R^k, \mathcal{B}^k, \mu_F)$ térbe, ahol μ_F a $(\xi_1, \dots, \xi_k)$ vektor eloszlása. Továbbá a *valószínűségi változók mérhető függvényeinek a jellemzéséről* szóló tétel alapján minden $B \in \mathcal{F}$ halmaz $B = \{\omega\colon (\xi_1(\omega), \dots, \xi_k(\omega)) \in B_1\}$ alakba írható alkalmas B_1 Borel mérhető halmazzal. Ezért a mértéktartó transzformációk által indukált integráltranszformációk tulajdonságaiból és a $(**)$ relációból következik, hogy tetszőleges $B \in \mathcal{F}$ és neki megfelelő B_1 halmazra

$$\begin{aligned}
\int_B g_\eta(\xi_1(\omega), \dots, \xi_k(\omega))\, dP(\omega) &= \int_{\{\omega\colon (\xi_1(\omega), \dots, \xi_k(\omega) \in B_1\}} g_\eta(\xi_1(\omega), \dots, \xi_k(\omega))\, dP(\omega) \\
&= \int_{B_1} g_\eta(x_1, \dots, x_k)\, d\mu_F(x_1, \dots, x_k) \\
&= \int_{\{\omega\colon (\xi_1(\omega), \dots, \xi_k(\omega) \in B_1\}} \eta(\omega)\, dP(\omega) = \int_B \eta(\omega)\, dP(\omega).
\end{aligned}$$

Ez azt jelenti, hogy

$$E(\eta|F)(\omega) = g_\eta(\xi_1(\omega), \dots, \xi_k(\omega))$$

választással teljesül a ii.′) reláció.

Láttuk, hogy ha $\mathcal{F}$ valamely $\xi_1, \dots, \xi_k$ valószínűségi változók által generált σ-algebra, akkor egy η valószínűségi változó $E(\eta|\mathcal{F})$ feltételes várható értéke kifejezhető a $g_\eta(x_1, \dots, x_k) = E(\eta|\xi_1 = x_1, \dots, \xi_k = x_k)$ k-változós függvény segítségével. Megfordítva, a $g_\eta(x_1, \dots, x_k)$ függvény is előállítható az $E(\eta|\mathcal{F})$ feltételes várható érték segítségével. Ugyanis a *valószínűségi változók mérhető függvényeinek a jellemzéséről* szóló tétel alapján az $E(\eta|\mathcal{F})$ feltételes várható érték felírható $E(\eta|\mathcal{F}) = h(\xi_1, \dots, \xi_k)$ alakban alkalmas Borel mérhető $h(x_1, \dots, x_k)$ függvény segítségével, és nem nehéz belátni, hogy ez a $h(x_1, \dots, x_k)$ függvény választható a $g_\eta(x_1, \dots, x_k)$ függvénynek.

Példa: A feltételes várható érték definiciójának jobb megértése érdekében tekintsük a következő példát: Legyen az $(\Omega, \mathcal{A}, P)$ valószínűségi mező az $\Omega = [0,1] \times [0,1]$ egységnégyzet, rajta a szokásos $\mathcal{A}$ Borel σ-algebrával, és legyen $P = \lambda$, a Lebesgue mérték az egységnégyzet Borel-mérhető részhalmazain. Legyen $\mathcal{F}$ az $A \times [0,1]$, $A \in \mathcal{B}_1$ alakú halmazokból álló σ-algebra, ahol $\mathcal{B}_1$ a $[0,1]$ intervallumon generált σ-algebrát jelöli. Tekintsünk egy tetszőleges (mérhető és integrálható) $f(x,y)$ függvényt az egységnégyzeten (az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn), és számoljuk ki az $E(f(x,y)|\mathcal{F})$ feltételes várható értéket az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn.

Ha az $f(x,y)$ függvény valóban függ mind a két változójától, akkor nem $\mathcal{F}$ mérhető függvény. Viszont definiáljuk a $g_0(x) = \int_0^1 f(x,y)\,dy$ és $g(x,y) = g_0(x)$ függvényeket. (A $g(x,y)$ függvény valójában nem függ az y koordinától, viszont tekinthető egy az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált valószínűségi változónak, és mivel nem függ az y koordinától (és Borel mérhető), ezért $\mathcal{F}$ mérhető. Azt állítom, hogy $E(f(x,y)|\mathcal{F}) = g(x,y)$. Ehhez azt kell ellenőrizni, hogy $\int_{A\times[0,1]} g(x,y)\,dx\,dy = \int_{A\times[0,1]} f(x,y)\,dx\,dy$. Viszont

$$\int_{A\times[0,1]} g(x,y)\,dx\,dy = \int_A g_0(x)\,dx = \int_A \left(\int_0^1 f(x,y)\,dy\right)\,dx = \int_{A\times[0,1]} f(x,y)\,dx\,dy,$$

amint állítottuk.

Az előző példa módosított változata: Legyen az $(\Omega, \mathcal{A}, P)$ valószínűségi mező az $\Omega = [0,1] \times [0,1]$ egységnégyzet, rajta a szokásos $\mathcal{A}$ Borel σ-algebrával. Rögzítsünk egy olyan $h(x,y)$ függvényt az egységnégyzeten, amelyre $h(x,y) \ge 0$ minden (x,y) pontban, $\int_0^1 \int_0^1 h(x,y)\,dx\,dy = 1$, és legyen a P mérték az egységnégyzeten a Lebesgue mérték szerint $h(x,y)$ sűrűségfüggvénnyel rendelkező mérték, azaz legyen $P(B) = \int_B h(x,y)\,dx\,dy$ az egységnégyzet minden Borel-mérhető B részhalmazán. Legyen $\mathcal{F}$ az $A \times [0,1]$ alakú halmazokból álló σ-algebra, ahol $A \in \mathcal{B}_1$, és $\mathcal{B}_1$ a $[0,1]$ intervallumon generált σ-algebrát jelöli. Tekintsünk egy tetszőleges (mérhető és a P mérték szerint integrálható) $f(x,y)$ függvényt az egységnégyzeten, és számoljuk ki az $E(f(x,y)|\mathcal{F})$ feltételes várható értéket az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn.

A módosított példában felvetett probléma megoldása. A keresett feltételes várható érték a következő: Definiáljuk a

$$g_0(x) = \int_0^1 f(x,y) \frac{h(x,y)}{\int_0^1 h(x,y)\,dy}\,dy = \frac{\int_0^1 f(x,y)h(x,y)\,dy}{\int_0^1 h(x,y)\,dy}$$

és $g(x,y) = g_0(x)$ függvényeket. Azt állítom, hogy a $g(x,y)$ függvény a keresett feltételes várható érték. Ez a függvény nem függ az y koordinától, így $\mathcal{F}$ mérhető. Azt állítom, hogy $E(f(x,y)|\mathcal{F}) = g(x,y)$. Azt kell ellenőrizni, hogy $\int_{A\times[0,1]} g(x,y)h(x,y)\,dx\,dy = \int_{A\times[0,1]} f(x,y)h(x,y)\,dx\,dy$ minden mérhető $A \subset [0,1]$ halmazra. Viszont

$$\begin{aligned}\int_{A\times[0,1]} g(x,y)h(x,y)\,dx\,dy &= \int_A \left(\int_0^1 h(x,y)\,dy\right) g_0(x)\,dx \\ &= \int_A \left(\int_0^1 f(x,y)h(x,y)\,dy\right) dx = \int_{A\times[0,1]} f(x,y)h(x,y)\,dx\,dy,\end{aligned}$$

amint állítottam. A kapott eredmény megfelel szemléletes képünknek, amely szerint rögzített x_0 számra az $E(f(x,y)|x = x_0)$ feltételes várható értéket úgy számolhatjuk ki, hogy az $f(x_0,y)$ függvényt kiintegráljuk az y változó szerint, de nem a $h(x_0,y)$ sűrűségfüggvény, hanem ennek normalizáltja a $\frac{h(x_0,y)}{\int_0^1 h(x_0,y)\,dy}$ sűrűségfüggvény szerint. Később tárgyalni fogok olyan eredményeket, amelyek az itt kapott képletet speciális esetként tartalmazzák.

Következő témánk a feltételes várható érték legfontosabb tulajdonságainak ismertetése, és az, hogy hogyan lehet a feltételes várható értékekkel számolni.

Az alábbi tételben felsorolom a feltételes várható érték legfontosabb tulajdonságait. Érdemes észrevenni, hogy ezek a tulajdonságok megfelelnek szemléletes képünknek. Olyan tényeket fejeznek ki például, hogy amennyiben a feltételben szereplő $\mathcal{F}$ σ-algebra független a ξ valószínűségi változótól, amelynek a feltételes várható értékét számoljuk, akkor ez a σ-algebra semmilyen információt nem ad a ξ valószínűségi változó viselkedéséről, ezért az $E(\xi|\mathcal{F})$ feltételes várható érték megegyezik az $E\xi$ várható értékkel (5. tulajdonság). Ha viszont a ξ valószínűségi változó $\mathcal{F}$ mérhető, akkor $\mathcal{F}$ ismeretében őt is ismerjük, ezért feltételes várható értéke megegyezik a valódi értékével, sőt ha egy $\xi\eta$ szorzat feltételes várható értékét számoljuk, akkor ξ kiemelhető a feltételes várható érték elé. (6. tulajdonság.)

Rögzítsünk egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőt, és legyen $\mathcal{F} \subset A$ az ebben a valószínűségi mezőben szereplő $\mathcal{A}$ σ-algebra egy rész-σ-algebrája. Az alábbi tulajdonságokban szereplő valószínűségi változók az előbb rögzített valószínűségi mezőn vannak definiálva.

Tétel a feltételes várható érték tulajdonságairól.

1. *Ha* $\sum\limits_{k=1}^{\infty} |a_k| E|\xi_k| < \infty$, $\mathcal{F} \subset A$ *tetszőleges* σ*-algebra, akkor*

$$E\left(\sum_{k=1}^{\infty} a_k\xi_k \middle| \mathcal{F}\right)(\omega) = \sum_{k=1}^{\infty} a_k E(\xi_k|\mathcal{F})(\omega) \quad \textit{majdnem minden } \omega\in\Omega \textit{ pontban.}$$

2. *Ha* $E|\xi| < \infty$, $\mathcal{G} \subset \mathcal{F} \subset \mathcal{A}$ *tetszőleges* σ*-algebrák, akkor* $E\left(E(\xi|\mathcal{F})|\mathcal{G}\right)(\omega) = E(\xi|\mathcal{G})(\omega)$ *majdnem minden* $\omega \in \Omega$ *pontban. Speciálisan* $E\left(E\xi|\mathcal{F}\right)) = E\xi$.
3. *Ha* ξ *olyan valószínűségi változó, amelyre* $P(a \le \xi \le b) = 1$ *alkalmas* $-\infty < a < b < \infty$ *számokkal,* $\mathcal{F} \subset \mathcal{A}$ σ*-algebra, akkor* $a \le E(\xi|\mathcal{F})(\omega) \le b$ *majdnem minden*

$\omega \in \Omega$ pontban. Általánosabban, ha $\xi(\omega) \geq \eta(\omega)$ majdnem minden $\omega \in \Omega$ pontban, akkor $E(\xi|\mathcal{F})(\omega) \geq E(\eta|\mathcal{F})(\omega)$ majdnem minden $\omega \in \Omega$ pontban.

4. *$E(\xi|\mathcal{A})(\omega) = \xi(\omega)$, ahol $\mathcal{A}$ a valószínűségi mező definiciójában szereplő ,,legnagyobb" σ-algebra. Ha $\mathcal{A}_0$ jelöli a triviális σ-algebrát, amelyik csak az Ω és $\emptyset$ üres halmazból áll, akkor $E(\xi|\mathcal{A}_0)(\omega) = E\xi$.*
5. *Ha $E|\xi| < \infty$, $\mathcal{F} \subset \mathcal{A}$, a ξ valószínűségi változó független az $\mathcal{F}$ σ-algebrától, azaz ha tetszőleges $F \in \mathcal{F}$ és a számegyenesen lévő Borel mérhető $B \subset R^1$ halmazokra, $P(\{\omega\colon\ \xi(\omega) \in B\} \cap F) = P(\{\omega\colon\ \xi(\omega) \in B\})P(F)$, akkor $E(\xi|\mathcal{F})(\omega) = E\xi$.*
6. *Ha $E\xi^2 < \infty$, $E\eta^2 < \infty$, $\mathcal{F} \subset \mathcal{A}$, a ξ valószínűségi változó $\mathcal{F}$ σ-algebrára mérhető valószínűségi változó, azaz tetszőleges a számegyenesen lévő Borel mérhető $B \subset R^1$ halmazra, $\{\omega\colon\ \xi(\omega) \in B\} \in \mathcal{F}$, akkor $E(\xi\eta|\mathcal{F})(\omega) = \xi(\omega)E(\eta|\mathcal{F})(\omega)$ majdnem minden $\omega \in \Omega$ pontban. Speciálisan, ebben az esetben $E(\xi|\mathcal{F})(\omega) = \xi(\omega)$ majdnem minden $\omega \in \Omega$ pontban.*

A tétel bizonyítása. Az 1. tulajdonság bizonyításához azt kell belátni, hogy minden $F \in \mathcal{F}$ halmazra

$$\int_F \left(\sum_{k=1}^\infty a_k E(\xi_k|\mathcal{F})(\omega)\right) P(\,d\omega) = \int_F \left(\sum_{k=1}^\infty a_k \xi_k)(\omega)\right) P(\,d\omega).$$

Ez az állítás viszont igaz, mert a feltételes várható érték definiciója és az integrál alapvető tulajdonságai alapján

$$\begin{aligned}\int_F \left(\sum_{k=1}^\infty a_k E(\xi_k|\mathcal{F})(\omega)\right) P(\,d\omega) &= \sum_{k=1}^\infty a_k \int_F E(\xi_k|\mathcal{F})(\omega) P(\,d\omega) \\ &= \sum_{k=1}^\infty a_k \int_F \xi_k(\omega) P(\,d\omega) = \int_F \left(\sum_{k=1}^\infty a_k \xi_k(\omega)\right) P(\,d\omega).\end{aligned}$$

A fenti számolásban a Lebesgue tétel biztosította az integrálok végtelen összegeivel végrehajtott számolások jogosságát, illetve azt a tény, hogy a $\sum\limits_{k=1}^\infty |a_k| E|\xi_k| < \infty$ feltételből a $\sum\limits_{k=1}^\infty |a_k| E|E(\xi_k|\mathcal{F})| < \infty$ reláció is következik. Ez azért igaz, mert $|E(\xi_k|\mathcal{F})| \leq E(|\xi_k||\mathcal{F})$, amit nem nehéz közvetlenül belátni, de következik a 3.) tulajdonságból (aminek bizonyításában nem használtuk az 1.) tulajdonságot), és $E(E(|\xi_k||\mathcal{F})) = E|\xi_k|$.

A 2. tulajdonság első állításának belátásához azt kell ellenőrizni, hogy az adott feltételek mellett minden $G \in \mathcal{G}$ halmazra

$$\int_G \xi(\omega) P(\,d\omega) = \int_G E(\xi|\mathcal{F})(\omega) P(\,d\omega),$$

mert a feltételes várható érték definiciója alapján

$$\int_G E(\xi|\mathcal{F})(\omega) P(\,d\omega) = \int_G E(E(\xi|\mathcal{F})|\mathcal{G})(\omega) P(\,d\omega)).$$

A kívánt azonosság viszont igaz, mert az adott feltételek mellett $G \in \mathcal{F}$. A második állítás belátása érdekében tekintsük a triviális csak a teljes és üres halmazból álló $\mathcal{A}_0 = \{\emptyset, \Omega\}$ σ-algebrát, és vegyük észre, hogy $\mathcal{A}_0 \subset \mathcal{F}$ minden a valószínűségi mezőn definiált $\mathcal{F}$ σ-algebrára, másrészt $E(\eta|\mathcal{A}_0) = E\eta$ tetszőleges η valószínűségi változóra. Innen $EE(\xi|\mathcal{F}) = E(E(\xi|\mathcal{F})|\mathcal{A}_0) = E(\xi|\mathcal{A}_0) = E\xi$.

A 3. tulajdonság azért érvényes, mert, ha $\xi(\omega) \ge \eta(\omega)$ majdnem minden $\omega \in \Omega$ pontban, akkor

$$\int_F E(\xi|\mathcal{F})(\omega)P(\,d\omega) = \int_F \xi(\omega)P(\,d\omega) \ge \int_F \eta(\omega)P(\,d\omega) = \int_F E(\eta|\mathcal{F})(\omega)P(\,d\omega)$$

minden $F \in \mathcal{F}$ halmazra. Innen következik, hogy az

$$F_0 = \{\omega\colon\ E(\xi|\mathcal{F})(\omega) < E(\xi|\mathcal{F})(\omega)\} \in \mathcal{F}$$

halmazra $P(F_0) = 0$. Ellenkező esetben ugyanis nem teljesülne a fenti egyenlőtlenség az $F_0 = F$ halmazra. Mivel az azonosan konstans függvények feltételes várható értékei önmagukkal egyenlő, ezért, ha $a \le \xi \le b$ egy valószínűséggel, akkor

$$a = E(a|\mathcal{F}) \le E(\xi|\mathcal{F}) \le E(b|\mathcal{F}) = b.$$

A 4. tulajdonság bizonyítása triviális, ezért azt elhagyom.

Az 5. tulajdonság bizonyításához azt kell belátni, hogy ha $F \in \mathcal{F}$, és az $\mathcal{F}$ σ-algebra független a ξ valószínűségi változótól, akkor

$$\int_F \xi(\omega)P(\,d\omega) = \int_F E\xi(\omega)P(\,d\omega) = P(F)E\xi.$$

Viszont ebben az esetben az F halmaz $I_F(\omega)$ indikátor függvénye független a $\xi(\omega)$ valószínűségi változótól, ezért $\int_F \xi(\omega)P(\,d\omega) = EI_F(\omega)\xi(\omega) = EI_F(\omega)E\xi = P(F)E\xi$.

A 6. tulajdonság bizonyításához azt kell belátni, hogy ha $E\xi^2 < \infty$, $E\eta^2 < \infty$ és ξ $\mathcal{F}$ mérhető valószínűségi változó, akkor minden $F \in \mathcal{F}$ halmazra

$$\int_F \xi(\omega)E(\eta|\mathcal{F})(\omega)P(\,d\omega) = \int_F \xi(\omega)\eta(\omega)P(\,d\omega).$$

Abban az esetben, ha a ξ valószínűségi változó egy B $\mathcal{F}$ mérhető halmaz indikátor függvénye, akkor ez az állítás nyilvánvaló, mert ekkor $B \cap C \in \mathcal{F}$, ezért

$$\begin{aligned}\int_F \xi(\omega)E(\eta|\mathcal{F})(\omega)P(\,d\omega) &= \int_{F\cap B} E(\eta|\mathcal{F})(\omega)P(\,d\omega)\\ &= \int_{F\cap B} \eta(\omega)P(\,d\omega) = \int_F \xi(\omega)\eta(\omega)P(\,d\omega).\end{aligned}$$

Ezután az 1. tulajdonságból következik, hogy a bizonyítandó azonosság érvényes olyan $\xi(\omega) = \sum c_j I_{B_j}(\omega)$ lépcsős függvényekre is, melyeket véges sok $B_j \in \mathcal{F}$ halmaz indikátorfüggvényének lineáris kombinációjaként kapunk. Mivel tetszőleges $\mathcal{F}$ mérhető ξ valószínűségi változó jól approximálható ilyen lépcsős függvényekkel, innen standard határátmenettel be lehet látni az állítást. Ennek részleteit elhagyom.

Megfogalmazom és bebizonyítom a fenti állítások egy érdekes következményét az alábbi tételben.

Tétel a feltételes várható érték egy optimum tulajdonságáról. *Legyen adva egy ξ valószínűségi változó, amelyre $E\xi^2 < \infty$, és egy $\mathcal{F}$ σ-algebra, $\mathcal{F} \subset \mathcal{A}$, egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ekkor az $E(\xi|\mathcal{F})$ feltételes valószínűségnek megvan a következő optimum tulajdonsága:*

$$E\left[\xi(\omega) - E(\xi|\mathcal{F})(\omega)\right]^2 = \inf_{\substack{\eta \ \mathcal{F} \text{ mérhető} \\ \text{valószínűségi változó, } E\eta^2<\infty}} E(\xi(\omega) - \eta(\omega))^2.$$

1. megjegyzés: A fenti tétel azt fejezi ki, hogy ha a ξ valószínűségi változót az $\mathcal{F}$ σ-algebra eseményeitől függő valószínűségi változóval, azaz $\mathcal{F}$ mérhető valószínűségi változóval akarjuk becsülni, és két becslés közül azt tekintjük jobbnak a másiknál, amelyikre a becslés és a ξ valószínűségi változó közötti különbség négyzetének a várható értéke kisebb, akkor az $E(\xi|\mathcal{F})$ feltételes várható érték az optimális becslés. Megjegyzem, hogy a bevezető valószínűségszámítás előadásban szerepelt egy eredmény, amelyik a fenti állítás speciális esetének tekinthető. Nevezetesen beláttuk, hogy $\operatorname{Var}\xi = E(\xi - E\xi)^2 \le E(\xi - a)^2$ tetszőleges a valós számra. Ha definiáljuk azt a triviális $\mathcal{F}_0 = \{\emptyset, \Omega\}$ σ-algebrát, amelyik csak az üres halmazból és a biztos eseményből áll, akkor erre a σ-algebrára csak a konstans függvények mérhetőek, és erre a $\mathcal{F}_0$ σ-algebrára nézve $E(\xi|\mathcal{F}_0) = E\xi$. Ezért az előbb megfogalmazott tétel az ebben a megjegyzésben megfogalmazott állítást jelenti ebben a speciális esetben.

2. megjegyzés: Az $(\Omega, \mathcal{A}, P)$ téren mérhető és négyzetesen integrálható függvények (valószínűségi változók) egy úgynevezett L_2 teret alkotnak, ami Hilbert tér és az $\mathcal{F}$ mérhető, négyzetesen integrálható függvények e tér egyik alterét alkotják. Ebben az interpretációban a fenti eredmény azt mondja ki, hogy a Hilbert tér ezen alterének a Hilbert tér egy ξ eleméhez legközelebbi eleme az $E(\xi|\mathcal{F})$ valószínűségi változó. A Hilbert terek úgy tekinthetők, mint (esetleg) végtelen dimenziós euklideszi terek. Tudjuk, hogy egy euklideszi térben egy ponthoz egy altérben levő pontok közül a legközelebbi pont e pont merőleges vetülete az altérre, továbbá ez a tulajdonság érvényes Hilbert terekre is. Az alább ismertetett bizonyítás tulajdonképpen ezen geometriai kép által sugallt módszer kidolgozása a most vizsgált esetben.

A tétel bizonyítása. Legyen η tetszőleges $\mathcal{F}$ mérhető négyzetesen integrálható függvény. Belátjuk, hogy

$$E\eta(\omega)(\xi(\omega) - E(\xi|\mathcal{F})(\omega)) = 0 \quad \text{azaz} \quad E\eta(\omega)\xi(\omega) = E\eta(\omega)E(\xi|\mathcal{F})(\omega).$$

(Ez jelenti azt, hogy $\xi(\omega) - E(\xi|\mathcal{F})(\omega)$ a $\xi(\omega)$ valószínűségi változó ortogonális vetülete az $\mathcal{F}$ mérhető és négyzetesen integrálható függvények alterére.) Ez az állítás azért igaz, mert

$$E\eta(\omega)\xi(\omega) = E(E(\eta\xi|\mathcal{F})(\omega)) = E\eta(\omega)E(\xi|\mathcal{F})(\omega)$$

a feltételes várható érték előző tételben megfogalmazott 2. és 6. tulajdonságai alapján. Innen tetszőleges η $\mathcal{F}$ mérhető, négyzetesen integrálható függvényre

$$\begin{aligned}
E(\eta-\xi)^2 &= E\left((\eta - E(\xi|\mathcal{F})) + (E(\xi|\mathcal{F}) - \xi)\right)^2 \\
&= E(\eta - E(\xi|\mathcal{F}))^2 + E\left(E(\xi|\mathcal{F}) - \xi\right)^2 + 2E(\eta - E(\xi|\mathcal{F}))(E(\xi|\mathcal{F}) - \xi)) \\
&= E(\eta - E(\xi|\mathcal{F}))^2 + E\left(E(\xi|\mathcal{F}) - \xi\right)^2 \ge E(\xi - E(\xi|\mathcal{F}))^2,
\end{aligned}$$

mert $E(\eta - E(\xi|\mathcal{F}))(E(\xi|\mathcal{F}) - \xi)) = E\eta(E(\xi|\mathcal{F}) - \xi)) - E(E(\xi|\mathcal{F}))(E(\xi|\mathcal{F}) - \xi)) = 0$. ($E(E(\xi|\mathcal{F}))(E(\xi|\mathcal{F}) - \xi)) = 0$, mert $E(\xi|\mathcal{F})$ $\mathcal{F}$ mérhető valószínűségi változó, így $\eta = E(\xi|\mathcal{F})$ választással is alkalmazhatjuk az $E\eta(\xi - E(\xi|\mathcal{F})) = 0$ azonosságot.) Az $E(\eta-\xi)^2 \ge E(\xi - E(\xi|\mathcal{F}))^2$ egyenlőtlenségből következik a tétel állítása.

Az előző tételben megfogalmazott eredmény fontos mind a valószínűségszámításban, mind a statisztikában. A statisztikában alapvető kérdés az, hogy hogyan lehet egy ismeretlen mennyiséget (jelen esetben egy valószínűségi változót) ismereteink alapján (jelen esetben a $\mathcal{F}$ σ-algebra, illetve annak ismeretében, hogy e σ-algebra mely halmazai következtek be, és melyek nem) minél jobban megbecsülni. A becslés jóságának természetes mérése az, hogy milyen kicsi a becslés és becsült mennyiség közötti különbség négyzetének a várható értéke. A fenti eredményt úgy lehet interpretálni, hogy az $E(\xi|\mathcal{F})(\omega)$ feltételes várható érték a $\xi(\omega)$ valószínűségi változó legjobb közelítése valamely $\mathcal{F}$ mérhető valószínűségi változóval (az $L_2(\Omega, \mathcal{A}, P)$ térben). Egy komoly kellemetlenség, hogy a feltételes várható érték kiszámítása nagyon bonyolult feladat. Egy fontos speciális esetben azonban, amikor normális eloszlású vektor bizonyos koordinátáinak ismeretében a többi koordináta feltételes várható értékét akarjuk kiszámítani ez a probléma viszonylag egyszerű. Erről szól a következő feladat. Csak azt a speciális esetet tekintem, amikor egy kétdimenziós normális eloszlású vektor egyik koordinátájának a feltételes várható értékét akarjuk kiszámolni a másik koordináta ismeretében. De ezt a feladatot viszonylag könnyen általánosíthatjuk.

Feladat:

Legyen (ξ,η) egy két-dimenziós normális eloszlású véletlen vektor. Számítsuk ki az $E(\xi|\eta)$ feltételes várható értéket.

Megoldás: Láttuk, (lásd a többváltozós centális határeloszlástétel előadás eredményeit) hogy $\xi = a\eta + \zeta$ alakban írható, ahol az a konstans alkalmas választásával (nevezetesen az $a = \frac{\mathrm{Cov}\,(\xi,\eta)}{\mathrm{Var}\,\eta}$ választással) elérhető, hogy a $\zeta = \xi - a\eta$ és η valószínűségi változók függetlenek legyenek. Ezzel az a választással

$$\begin{aligned}
E(\xi|\eta) &= E((a\eta+\zeta)|\eta) = aE(\eta|\eta) + E(\zeta|\eta) = a\eta + E\zeta = a(\eta - E\eta) + E\xi \\
&= \frac{\mathrm{Cov}\,(\xi,\eta)}{\mathrm{Var}\,\eta}(\eta - E\eta) + E\xi
\end{aligned}$$

a várható értéknek az előző tételben szereplő 1., 5. és 6. tulajdonságai alapján.

Megjegyzés: Vegyük észre, hogy a fenti feladat eredménye szerint egy normális eloszlású véletlen vektor egyik koordinátájának a feltételes várható értéke a feltételben szereplő koordináta lineáris függvénye. Ez az állítás érvényes az itt nem tárgyalt magasabb dimenziójú normális vektorokra természetes módon megfogalmazható általánosabb probléma esetében is. Láttuk, egy előző eredményben, hogy egy valószínűségi változó feltételes várható értéke feltéve bizonyos más valószínűségi változókat megegyezik a tekintett valószínűségi változónak a feltételben szereplő valószínűségi változók segítségével megadható legjobb közelítésével az L_2 normában. A fenti feladat eredménye (illetve annak itt meg nem fogalmazott magasabb dimenziós általánosítása) azt mondja, hogy abban a speciális esetben, ha egy normális vektor koordinátáinak a feltételes várható értékét akarjuk kiszámolni feltéve más koordináták értékeit, akkor ez a legjobb L_2-normában vett közelítés egyben a legjobb (a feltételben szereplő valószínűségi változókkal kifejezhető) L_2 normában vett *lineáris* közelítés. E tény alapvető szerepet játszik sok elméleti statisztikai vizsgálatban.

Láttuk korábban, hogy egy véletlen vektor függvényeinek a várható értékét ki tudjuk számolni a véletlen vektor eloszlása szerinti alkalmas (Lebesgue) integrál segítségével. Felmerül a kérdés: Lehet-e hasonló kalkulust kidolgozni a feltételes várható értékek kiszámolására? A feltételes várható értékeket szeretnénk a feltételes eloszlások segítségével kiszámolni. Erre a kérdésre lényegében pozitív választ lehet adni. A megfelelő kalkulus kidolgozása érdekében a következő két problémát kell megoldani.

i. Adjunk meg jó feltételes eloszlásokat, azaz definiáljuk a $P(\xi \in A|\mathcal{F})(\omega)$ feltételes valószínűségek rendszerét minden B Borel mérhető halmazra úgy, hogy jól tudjunk vele számolni.

ii. Találjuk meg az integrálformulákat a feltételes várható értékek kiszámolására.

E két probléma közül a második a kevésbé nehéz. Az első probléma kapcsán komoly nehézségek lépnek fel. Ezek azzal függnek össze, hogy Lebesgue integrált csak *(σ-additív)* mértékek szerint tudunk alkalmazni, ezért a feltételes eloszlásokat úgy kell definiálni, hogy azok majdnem minden ω-ra mértékek legyenek. A feltételes valószínűségek felsorolt tulajdonságai közül az első egy ilyen jellegű tulajdonságot fogalmaz meg, ha az ott felírt azonosságot $\xi_k = I(\omega\colon\ \xi(\omega) \in A_k)$ alakú valószínűségi változókra alkalmazzuk. De a feltételes valószínűség csak egy valószínűséggel van definiálva. Elképzelhető, hogy a megkövetelt azonosságok közül az egyik egy null mértékű halmazon nem teljesül. Akkor ezt ki tudjuk javítani a feltételes valószínűségek módosításával egy null mértékű halmazon. De még a legegyszerűbb esetekben is kontinum sok egy valószínűséggel teljesülő azonosságot kell egyszerre teljesíteni. Az, hogy ez lehetséges korántsem magától értetődő.

Van egy fontos speciális eset, amikor a kívánt tulajdonságú feltételes eloszlásoknak nemcsak a létezését tudjuk biztosítani, hanem azokat explicit módon meg is tudjuk adni, így jól tudunk velük számolni. Erről szól a következő, statisztikai alkalmazásokban fontos eredmény. Ebben a következő kérdéssel foglalkozunk.

Egy $k+l$ dimenziós $(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)$ véletlen vektornak ismerjük a (létező) $f(x_1,\dots,x_k,y_1,\dots,y_l)$ sűrűségfüggvényét, és ki akarjuk számítani e véletlen vektor valamely $h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)$ függvényének a feltételes várható értékét az $\eta_1=y_1$, $\dots$, $\eta_l=y_l$ feltételek mellett. Ennek érdekében egy olyan formulát bizonyítunk a $(\xi_1,\dots,\xi_k)$ véletlen vektor feltételes eloszlására, feltéve az $\eta_1=y_1,\dots,\eta_l=y_l$ feltételeket, amely lehetővé teszi, hogy a közönséges várható értékek kiszámolására tanult formulákat adaptáljuk erre az esetre is, és kiszámítsuk a minket érdeklő feltételes várható értékeket.

Tétel feltételes eloszlások kiszámolásáról. *Legyen $(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)$ egy olyan véletlen vektor az R^{k+l} $k+l$ dimenziós euklideszi térben, amelynek létezik sűrűségfüggvénye, amit $f(x_1,\dots,x_k,y_1,\dots,y_l)$-lel fogok jelölni.*

Ebben az esetben létezik a $(\xi_1,\dots,\xi_k)$ véletlen vektornak feltételes sűrűségfüggvénye feltéve az $\eta_1=y_1,\dots,\eta_l=y_l$ feltételeket, és az megadható az

$$f(x_1,\dots,x_k|y_1,\dots,y_l)=\frac{f(x_1,\dots,x_k,y_1,\dots,y_l)}{g(y_1,\dots,y_l)},$$

képlet segítségével, ahol

$$g(y_1,\dots,y_l)=\int f(x_1,\dots,x_k,y_1,\dots,y_l)\,dx_1,\dots dx_k,$$

azaz az $(\eta_1,\dots,\eta_l)$ véletlen vektor sűrűségfüggvénye. Ez azt jelenti, hogy tetszőleges Borel mérhető $A\subset R^k$ halmazra

$$\begin{aligned}P&(\{\omega\colon\ (\xi_1(\omega),\dots,\xi_k(\omega))\in A\}|\eta_1(\omega)=y_1,\dots,\eta_l(\omega)=y_l)\\&=\int_A f(x_1,\dots,x_k|y_1,\dots,y_l)\,dx_1\dots dx_k.\end{aligned}$$

A tétel bizonyítása. Annak igazolásához, hogy a fenti feltételes valószínűségeket az adott módon ki lehet számolni a $(*)$ formulát kell ellenőrizni ebben az esetben, azaz azt kell megmutatni, hogy minden $A\subset R^k$ és $B\in R^l$ Borel mérhető halmazpárra

$$\begin{aligned}&P(\{\omega\colon\ (\xi_1(\omega),\dots,\xi_k(\omega))\in A\}\cap\{\omega\colon\ (\eta_1(\omega),\dots,\eta_l(\omega))\in B\})\\&\quad=\int\limits_B\left[\int\limits_A f(x_1,\dots,x_k|y_1,\dots,y_l)\,dx_1\dots dx_k\right]g(y_1,\dots,y_l)\,dy_1\dots dy_l.\end{aligned}$$

Ez az azonosság viszont érvényes, mert

$$\begin{aligned}\int_B&\left[\int_A f(x_1,\dots,x_k|y_1,\dots,y_l)\,dx_1\dots dx_k\right]g(y_1,\dots,y_l)\,dy_1\dots dy_l\\&=\int_{A\times B} f(x_1,\dots,x_k,y_1,\dots,y_l)\,dx_1\dots dx_k\,dy_1\dots dy_l\\&=P(\{\omega\colon\ (\xi_1(\omega),\dots,\xi_k(\omega))\in A\}\cap\{\omega\colon\ (\eta_1(\omega),\dots,\eta_l(\omega))\in B\}).\end{aligned}$$

Megfogalmazom az előző tétel egy megfelelőjét, amely azt állítja, hogy a számunkra érdekes feltételes eloszlások léteznek az általános esetben. E tétel bizonyítását, amely nem egyszerű, a kiegészítésben adom meg. Az eredmény szépséghibája az, hogy csak egzisztenciatétel, és nem ad lehetőséget a feltételes eloszlások effektiv kiszámolására.

Tétel reguláris feltételes eloszlások létezéséről. *Legyen $(\xi_1(\omega),\dots,\xi_k(\omega))$ egy k-dimenziós valószínűségi vektor egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, $\mathcal{F}\subset\mathcal{A}$ σ-algebra. Jelölje $\mathcal{B}$ a Borel σ-algebrát az R^k k-dimenziós euklideszi téren. A $(\xi_1(\omega),\dots,\xi_k(\omega))$ véletlen vektornak létezik az $\mathcal{F}$ σ-algebra szerinti reguláris feltételes eloszlása, azaz meg lehet adni egy olyan $F(B,\omega)$, $F(B,\omega)$: $\mathcal{B}\times\Omega\to R^1$, függvényt, amelyre*

i.) *$F(B,\cdot)$ minden $B\in\mathcal{B}$ halmazra $\mathcal{F}$ mérhető valószínűségi változó.*

ii.) *$F(\cdot,\omega)$ minden $\omega\in\Omega$ pontra valószínűségi mérték az R^k k-dimenziós euklideszi tér $\mathcal{B}$ σ-algebráján, azaz $F(R^k,\omega)=1$, $0\le F(B,\omega)\le 1$ minden $B\in\mathcal{B}$ halmazra, $F\left(\bigcup\limits_{k=1}^{\infty}B_k,\omega\right)=\sum\limits_{k=1}^{\infty}F(B_k,\omega)$ minden diszjunkt $B_k\in\mathcal{B}$, $k=1,2,\dots,$ halmazokból álló rendszerre.*

iii.) *$F(B,\omega)=P\left((\xi_1(\omega),\dots,\xi_k(\omega))\in B|\mathcal{F}\right)(\omega)$ minden $B\in\mathcal{B}$ halmazra. Ez azt jelenti, hogy az $F(B,\omega)$ függvény úgy tekinthető, mint a $P\left((\xi_1(\omega),\dots,\xi_k(\omega))\in B|\mathcal{F}\right)(\omega)$ feltételes valószínűség egyik verziója.*

A következő tételben megfogalmazom azt az eredményt, amely miatt az előző két eredmény hasznos volt. Ebben egy véletlen vektor függvényének a várható értékét számoljuk ki, mint ennek a függvénynek a véletlen vektor eloszlása szerinti integrált. De e tétel alkalmazásához szükséges a reguláris feltételes eloszlások létezése. Ezt biztosítja általános esetben az előző tétel. Az előtte tárgyalt *tétel feltételes eloszlások kiszámolásáról* eredménye lehetővé teszi e reguláris feltételes eloszlás explicit kiszámolását egy fontos speciális esetben. A tétel bizonyítását a kiegészítésben írom le.

Tétel a feltételes várható érték kiszámolásáról reguláris feltételes eloszlások segítségével. *Legyen adva egy $(\Omega,\mathcal{A},P)$ valószínűségi mező, azon egy $\mathcal{F}\subset\mathcal{A}$ σ-algebra, egy $\xi=(\xi_1,\dots,\xi_k)$ k-dimenziós és $\eta=(\eta_1,\dots,\eta_l)$ l-dimenziós véletlen véletlen vektor. Legyen továbbá az $(\eta_1,\dots,\eta_l)$ véletlen vektor $\mathcal{F}$ mérhető, és jelölje $F(B,\omega)$, $B\in\mathcal{B}^k$, $\omega\in\Omega$, a $(\xi_1,\dots,\xi_k)$ véletlen vektor reguláris feltételes eloszlását feltéve az $\mathcal{F}$ σ-algebrát. Legyen $h(x_1,\dots,x_k,y_1,\dots,y_l)$ egy $k+l$ változós Borel mérhető függvény, amelyre $E|h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|<\infty$. Az $E\left(h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|\mathcal{F}\right)(\omega)$ feltételes várható értéket ki lehet számítani a következő képlet segítségével:*

$$\begin{aligned}&E\left(h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|\mathcal{F}\right)(\omega)\\&\qquad=\int h(x_1,\dots,x_k,\eta_1(\omega),\dots,\eta_l(\omega))F(dx_1,\dots,dx_k,\omega).\end{aligned}$$

Megjegyzés: A fenti eredmény azt mondja ki, hogy a feltételes várható értéket a természetes módon számolhatjuk ki egy reguláris feltételes eloszlás szerint. Azoknak a valószínűségi változóknak, amelyek $\mathcal{F}$ mérhetőek, tudjuk az értékét a feltételben szereplő $\mathcal{F}$ σ-algebra ismeretében, ezért természetes ezek értékeit behelyettesíteni a tekintett valószínűségi változóknak abba a függvényébe, amelynek feltételes várható értékét ki akarjuk számolni. Ezután ezt a feltételes várható értéket úgy számolhatjuk ki, hogy a helyettesítés után kapott függvényt integráljuk a többi valószínűségi változó $\mathcal{F}$ σ-algebra szerinti feltételes eloszlása szerint. Megjegyzem, hogy ez az eredmény, illetve annak albb ismertetett változata a *feltételes eloszlások kiszámolásáról szóló tétel* eredményével együtt egyszerű eljárást ad *az előző példa módosított változata* néven tekintett feladat megoldására.

Megfogalmazom a feltételes várható érték kiszámításáról szóló tétel megfelelőjét arra az esetre, ha az $E(h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|\eta_1=y_1,\dots,\eta_l=y_l)$ feltételes várható értéket akarjuk kiszámolni a $P((\xi_1,\dots,\xi_k)\in A|\eta_1=y_1,\dots,\eta_l=y_l)$ feltételes reguláris eloszlás ismeretében.

A feltételes várható érték kiszámolásáról szóló tétel egy változata. *Legyen adva egy $(\Omega,\mathcal{A},P)$ valószínűségi mező, azon egy egy $\xi=(\xi_1,\dots,\xi_k)$ k-dimenziós és $\eta=(\eta_1,\dots,\eta_l)$ l-dimenziós véletlen vektor. Legyen adva továbbá a $\xi=(\xi_1,\dots,\xi_k)$ véletlen vektor $F(B|y_1,\dots,y_l)$ feltételes eloszlása feltéve az $(\eta_1,\dots,\eta_k)$ véletlen vektort, azaz legyen $F(B|y_1,\dots,y_l)$, $B\in\mathcal{B}^k$, $y_j\in R$, $1\le j\le l$, egy függvény a következő tulajdonságokkal:*

a) Minden $(y_1,\dots,y_l)$ vektorra $F(\cdot|y_1,\dots,y_l)$ valószínűségi mérték az $(R^k,\mathcal{B}^k)$ euklideszi térben.

b) Minden $B\in\mathcal{B}^k$ halmazra $F(B|\cdot,\cdot,\cdot)$ Borel mérhető függvény az $(R^l,\mathcal{B}^l)$ térben.

c) $P((\xi_1,\dots,\xi_k)\in B|\eta_1=y_1,\dots,\eta_l=y_l)=F(B|y_1,\dots,y_l)$.

Legyen $h(x_1,\dots,x_k,y_1,\dots,y_l)$ egy $k+l$ változós Borel mérhető függvény, amelyre $E|h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|<\infty$. Ekkor

$$\begin{aligned}&E\left(h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|\eta_1=y_1,\dots,\eta_l=y_l\right)\\&\qquad=\int h(x_1,\dots,x_k,y_1,\dots,y_l)F(dx_1,\dots,dx_k|y_1,\dots,y_l).\end{aligned}$$

Ha létezik $f(x_1,\dots,x_k|y_1,\dots,y_l)$ feltételes sűrűségfüggvénye az $F(B|y_1,\dots,y_l)$ feltételes eloszlásfüggvénynek, azaz létezik olyan $f(\cdots|\cdots)$ függvény, amelyre teljesül az

$$F(B|y_1,\dots,y_k)=\int_B f(x_1,\dots,x_k|y_1,\dots,y_l)\,dx_1\dots dx_k\quad\text{minden } B\in\mathcal{B}^k \text{ halmazra}$$

azonosság, akkor

$$\begin{aligned}&E\left(h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|\eta_1=y_1,\dots,\eta_l=y_l\right)\\&\qquad=\int h(x_1,\dots,x_k,y_1,\dots,y_l)f(x_1,\dots,x_k|y_1,\dots,y_l)\,dx_1\dots dx_k.\end{aligned}$$

Befejezésül a következő két feladat megoldását tárgyalom.

1. feladat:

Legyen ζ és η két független valószínűségi változó, $G(u,v)$ két-változós mérhető függvény. Mutassuk meg, hogy $EG(\zeta,\eta)|\eta=y)=EG(\zeta,y)$.

2. feladat:

Legyen (ξ,η) két-dimenziós normális eloszlású véletlen vektor. Mutassuk meg, hogy $P(\xi<x|\eta=y)=\Phi_{m,\sigma}(x)$, $m=E\xi+\frac{\mathrm{Cov}\,(\xi,\eta)(y-E\eta)}{\mathrm{Var}\,\eta}$ és $\sigma^2=\mathrm{Var}\,\xi-\dfrac{\mathrm{Cov}\,(\xi,\eta)^2}{\mathrm{Var}\,\xi}$ paraméterekkel, ahol $\Phi_{m,\sigma}(\cdot)$ az m várható értékű és σ szórású normális eloszlású valószínűségi változó eloszlásfüggvénye.

Megjegyzés: Az 1. feladat állítása heurisztikusan természetes. Ugyanis, ha $\eta=y$, akkor $G(\zeta,\eta)=G(\zeta,y)$ és mivel η és ζ független valószínűségi változók, ezért η ismerete semmilyen információt nem ad a ζ valószínűségi változó viselkedéséről. Ez azt sugallja, hogy az $EG(\zeta,\eta)|\eta=y)$ feltételes várható értéket úgy kapjuk meg, hogy a $G(\zeta,y)$ valószínűségi változó várható értékét számoljuk ki a (ζ eredeti eloszlása szerint).

A 2. feladat állítását az 1. feladat állításnak és annak a ténynek a segítségével tudjuk belátni, hogy egy két-dimenziós véletlen normális eloszlású vektor első koordinátája kifejezhető a második koordináta és egy attól független normális eloszlású valószínűségi változó lineáris kombinációjaként. Vegyük észre, hogy a 2. feladat eredménye szerint ξ feltételes eloszlása rögzített $\eta=y$ feltétel esetén normális eloszlás, amelynek szórása nem függ az $\eta=y$ feltételtől.

Az 1. feladat megoldása: Legyen $F(u)$ a ζ, $H(y)$ az η valószínűségi változó eloszlásfüggvénye. Ekkor $EG(\zeta,y)=\int G(u,y)F(\,du)$, és azt kell belátnunk, hogy tetszőleges Borel-mérhető A halmazra $\int_A EG(\zeta,y)H(\,dy)=EG(\zeta,\eta)I_A(\eta)$. Viszont

$$
\begin{aligned}
EG(\zeta,\eta)I_A(\eta) &= \int I_A(y)G(u,y)F(\,du)H(\,dy)\\
&= \int I_A(y)\left[\int G(u,y)F(\,du)\right]H(\,dy)=\int_A EG(\zeta,y)H(\,dy)
\end{aligned}
$$

a Fubini tétel alapján, ahol $I_A(\cdot)$ az A halmaz indikátorfüggvényét jelöli.

A 2. feladat megoldása. A többdimenziós normális eloszlásról szóló előadás eredményei alapján a (ξ,η) véletlen vektor ξ koordinátája előállítható $\xi=a\eta+\zeta$ alakban $a=\frac{\mathrm{Cov}\,(\xi,\eta)}{\mathrm{Var}\,\eta}$ választással úgy, hogy $\zeta=\xi-a\eta$ és η független, normális eloszlású valószínűségi változók. Ezért $P(\xi<x|\eta=y)=P(a\eta+\zeta<x|\eta=y)=P(ay+\zeta<x)$ az 1. feladat eredménye alapján. (Ez az 1. feladat eredményéből következik, ha azt a következő $G(u,v)=G_x(u,v)$ függvényre alkalmazzuk: $G(u,v)=1$, ha $u+av<x$, és $G(u,v)=0$,

ha $u+av\ge 0$, ahol $a=\frac{\mathrm{Cov}\,(\xi,\eta)}{\mathrm{Var}\,\eta}$.) Viszont ζ normális eloszlású valószínűségi változó $E\xi-aE\eta=E\xi-\frac{\mathrm{Cov}\,(\xi,\eta)E\eta}{\mathrm{Var}\,\eta}$ várható értékkel és $\mathrm{Var}\,\xi+a^2\mathrm{Var}\,\eta-2a\mathrm{Cov}\,(\xi,\eta)=\mathrm{Var}\,\xi-\frac{\mathrm{Cov}\,(\xi,\eta)^2}{\mathrm{Var}\,\xi}$ szórásnégyzettel. Másrészt $Eay+\zeta=ay+E\zeta$, és $\mathrm{Var}\,(ay+\zeta)=\mathrm{Var}\,\zeta$, ahonnan következik, hogy a feladat megoldásában megjelenő normális eloszlásnak valóban az ott megadott m várható értéke és σ^2 szórásnégyzete van.

Kiegészítés.

Először a következő az előadás fő részében megfogalmazott eredmény bizonyítását ismertetem.

A valószínűségi változók mérhető függvényeinek a jellemzéséről szóló tétel bizonyítása. Az analízis általános eredményeiből következik, hogy ha $g(x_1,\dots,x_k)$ Borel mérhető függvény, akkor a $g(\xi_1,\dots,\xi_k)$ valószínűségi változó mérhető a $\xi_1,\dots,\xi_k$ valószínűségi változók által generált $\mathcal{F}$ σ-algebrára. Az állítás megfordítását először a $\mathcal{F}$ mérhető indikátor függvényekre látom be, azaz megmutatom, hogy ha $A\in\mathcal{F}$, akkor létezik olyan $g(x_1,\dots,x_k)$ Borel mérhető függvény, amelyre $g(\xi_1(\omega),\dots,\xi_k(\omega))=1$, ha $\omega\in A$, és $g(\xi_1(\omega),\dots,\xi_k(\omega))=0$, ha $\omega\notin A$. Továbbá ez a $g(x_1,\dots,x_k)$ függvény választható úgy, mint az R^k tér egy Borel mérhető halmaz indikátor függvénye.

Abban a speciális esetben, ha $A=\prod\limits_{j=1}^{k} I(\{\omega\colon\ \xi_j(\omega)\in B_j\})$ valamely $B_1,\dots,B_k$ Borel mérhető halmazokkal, és $I(B)$ egy B halmaz indikátor függvényét jelöli, az állítás nyilvánvaló a $g(x_1,\dots,x_k)=1$, ha $x_j\in B_j$, $1\le j\le k$, és $g(x_1,\dots,x_k)=0$ egyébként függvényválasztással. Ezért elég belátni, hogy a kívánt tuladjonsággal rendelkező halmazok σ-algebrát alkotnak. Ez viszont következik a következő észrevételből. Ha $I(\omega\colon\ \omega\in A)=g(\xi_1(\omega),\dots,\xi_k(\omega))$, akkor $I(\omega\colon\ \omega\notin A)=1-g(\xi_1(\omega),\dots,\xi_k(\omega))$, és ha $I(\omega\colon\ \omega\in A_j)=g_j(\xi_1(\omega),\dots,\xi_k(\omega))$, $j=1,2,\dots$, akkor $I(\omega\colon\ \omega\in\bigcup\limits_{j=1}^{\infty}A_j)=g(\xi_1(\omega),\dots,\xi_k(\omega))$ $g(x_1,\dots,x_k)=\max\limits_{1\le j<\infty} g(x_1,\dots,x_k)$ választással, és az így definiált $g(x_1,\dots,x_k)$ függvény szintén egy Borel mérhető halmaz indikátor függvénye.

Ezután könnyen látható, hogy amennyiben az η valószínűségi változó lépcsős függvény az $(\Omega,\mathcal{F},P)$ téren, azaz felírható $\eta=\sum\limits_{j=1}^{l}c_lI(A_j)$ véges összeg alakban valamely $A_j\in\mathcal{F}$ diszjunkt halmazok összegeként, akkor létezik $\eta=g(\xi_1,\dots,\xi_k)$ alakú előállítása alkalmas Borel mérhető függvény segítségével. Továbbá, ha η_j, $j=1,2\dots$, lépcsős függvények, $\sup\eta_j(\omega)<\infty$ minden $\omega\in\Omega$ elemi eseményre, akkor az $\eta=\sup\eta_j$ valószínűségi változó is előállítható a kívánt $\eta=g(x_1,\dots,x_k)$ alakban.

Valóban, legyen $\eta_j=g_j(\xi_1,\dots,\xi_k)$, $j=1,2,\dots$, és definiáljuk a $g(x_1,\dots,x_k)=\sup g_j(x_1,\dots,x_k)$ függvényt, ha $\sup g_j(x_1,\dots,x_k)<\infty$, és legyen $g(x_1,\dots,x_k)=0$, ha $\sup g_j(x_1,\dots,x_k)=\infty$. Ez megadja a kívánt előállítást. Mivel tetszőleges pozitív $\mathcal{F}$ mérhető valószínűségi változó előállítható, mint monoton növekvő elemi függvények limesze, innen következik az ilyen valószínűségi változók kívánt alakú előállítása is. Ezután egy általános η valószínűségi változót felírva $\eta=\eta_+-\eta_-$ alakban, ahol $\eta_+=$

$\max(\eta,0)$, $\eta_-=-\min(\eta,0)$, megkapjuk a kívánt reprezentációt tetszőleges $\mathcal{F}$ mérhető valószínűségi változóra.

A kívánt előállítás (a tétel értelmében vett) egyértelműségének belátása érdekében vegyük egy η valószínűségi változó két

$$\eta(\omega)=g_1(\xi_1(\omega),\dots,\xi_k(\omega)) \quad \text{és} \quad \eta(\omega)=g_2(\xi_1(\omega),\dots,\xi_k(\omega))$$

előállítását és a $h(x_1,\dots,x_k)=g_1(x_1,\dots,x_k)-g_2(x_1,\dots,x_k)$ különbségfüggvényt. Ekkor $h(\xi_1(\omega),\dots,\xi_k(\omega))=0$ majdnem minden $\omega\in\Omega$ elemi eseményre, és ez csak úgy lehetséges, ha $h(x_1,\dots,x_k)=0$ a $(\xi_1,\dots,x_k)$ véletlen vektor μ eloszlása szerint majdnem minden pontban.

A feltételes reguláris eloszlás létezéséről szóló tételnek alábbi általánosabb alakját bizonyítom, amelyben értéküket tetszőleges teljes szeparábilis metrikus térben felvevő valószínűségi változókat tekintek.

Tétel reguláris feltételes eloszlások létezéséről. *Legyen $\xi(\omega)$ egy értékeit valamely (U,ρ) teljes szeparábilis metrikus térben felvevő valószínűségi változó egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, és legyen adva a valószínűségi mező valamely $\mathcal{F}\subset\mathcal{A}$ rész-σ-algebrája. Jelölje $\mathcal{B}$ a Borel σ-algebrát az (U,ρ) metrikus téren. A $\xi(\omega)$ valószínűségi változónak létezik az $\mathcal{F}$ σ-algebra szerinti reguláris feltételes eloszlása, azaz meg lehet adni egy olyan $F(B,\omega)$, $F(B,\omega)$: $\mathcal{B}\times\Omega\to R^1$ függvényt, amelyre*

i.) *$F(B,\cdot)$ minden $B\in\mathcal{B}$ halmazra $\mathcal{F}$ mérhető valószínűségi változó.*

ii.) *$F(\cdot,\omega)$ minden $\omega\in\Omega$ pontra valószínűségi mérték az (U,ρ) metrikus tér $\mathcal{B}$ σ-algebráján, azaz $F(U,\omega)=1$, $0\le F(B,\omega)\le 1$ minden $B\in\mathcal{B}$ halmazra, és*

$$F\left(\bigcup_{k=1}^{\infty}B_k,\omega\right)=\sum_{k=1}^{\infty}F(B_k,\omega)$$

minden diszjunkt $B_k\in\mathcal{B}$, $k=1,2,\dots$, halmazokból álló rendszerre.

iii.) *$F(B,\omega)=P\left(\xi(\omega)\right)\in B|\mathcal{F})(\omega)$ minden $B\in\mathcal{B}$ halmazra. Ez azt jelenti, hogy az $F(B,\omega)$ függvény úgy tekinthető, mint a $P\left(\xi(\omega)\right)\in B|\mathcal{F})(\omega)$ feltételes valószínűség egyik verziója. Ez úgy is megfogalmazható, hogy $F(B,\omega)$ annak a Q_B mértéknek a a P valószínűségi mérték szerinti $\frac{dQ_B}{dP}$ Radon–Nikodym deriváltja az $(\Omega,\mathcal{F},P)$ mértéktéren, amelyet a $Q_B(A)=P(B\cap A)$, $A\in\mathcal{F}$, képlet definiál.*

A tétel bizonyítása előtt megfogalmazok néhány a bizonyításban hasznos állítást lemma formájában. Ezek bizonyítását itt nem tárgyalom.

1. lemma. *Egy $(X,\mathcal{A})$ mérhető tér megszámlálható sok $\mathcal{A}$ mérhető halmazából álló $\mathcal{C}$ halmazrendszert tartalmazó legszűkebb algebra szintén megszámlálható számosságú.*

2. lemma. *Legyen (X,ρ) egy szeparábilis teljes metrikus tér, és μ egy véges mérték e tér $\mathcal{A}$ Borel σ-algebráján. Ekkor minden $B\in\mathcal{A}$ halmazra és $\varepsilon>0$ számra létezik olyan $K\subset B$ kompakt halmaz, amelyre $\mu(K)\ge\mu(B)-\varepsilon$.*

3. lemma. *Legyen* $K_1 \supseteq K_2 \supseteq \cdots$ *egymásba skatulyázott kompakt halmazok családja egy* (X,ρ) *szeparábilis teljes metrikus térben, amelyek metszete üres, azaz* $\bigcap\limits_{n=1}^{\infty} K_n = \emptyset$*. Ekkor létezik olyan* n *index, amelyre* $K_n = \emptyset$*.*

4. lemma. *Legyen* μ *egy (végesen) additív, nem negatív halmazfüggvény egy* X *tér* $\mathcal{C}$ *algebráján. Ez a* μ *halmazfüggvény* σ*-additív a* $\mathcal{C}$ *algebrán, ha teljesíti a következő folytonossági tulajdonságot: Minden olyan egymásba skatulyázott* $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$*,* $B_n \in \mathcal{C}$*,* $n = 1,2,\dots$*, halmazrendszerre, amelyre* $\bigcap\limits_{n=1}^{\infty} B_n = \emptyset$ *teljesül a* $\lim\limits_{n\to\infty} \mu(B_n) = 0$ *reláció.*

Az ötödik lemma megfogalmazása előtt bevezetem a következő definiciót.

Monoton halmazosztályok definiciója. *Ha egy* X *halmaz részhalmazainak* $\mathcal{C}$ *család-ja olyan tulajdonságú, hogy* $C_n \in \mathcal{C}$*,* $C_n \subseteq C_{n+1}$*,* $n = 1,2,\dots$*, esetén* $C_\infty = \bigcup\limits_{n=1}^{\infty} C_n \in \mathcal{C}$*, és* $C_n \in \mathcal{C}$*,* $C_n \supseteq C_{n+1}$*,* $n = 1,2,\dots$*, esetén* $C_\infty = \bigcap\limits_{n=1}^{\infty} C_n \in \mathcal{C}$*, akkor* $\mathcal{C}$*-t monoton halmazosztálynak nevezzük.*

5. lemma. *Egy* $\mathcal{C}$ *algebrát tartalmazó monoton halmazosztály tartalmazza a* $\mathcal{C}$ *algebra által generált* σ*-algebrát is.*

A tétel bizonyítása. Először megmutatom, hogy a tétel bizonyítását a következő 'A tulajdonság' igazolására lehet visszavezetni.

> *A tulajdonság:* Létezik egy (megszámlálható sok halmazból álló) $\mathcal{C} \subset \mathcal{B}$ algebra az (X,ρ) metrikus téren úgy, hogy a $\mathcal{B}$ σ-algebra a $\mathcal{C}$ algebra által generált legszűkebb σ-algebra, valamint a $Q(C,\omega) = P(\xi(\omega) \in C|\mathcal{F})(\omega)$ feltételes valószínűségeknek egy verziója minden $C \in \mathcal{C}$ halmazra és egy olyan $\Omega_0 \subset \Omega$ a $P(\Omega_0) = 1$ és $\Omega_0 \in \mathcal{F}$ tulajdonságokat teljesítő halmaz, amelyekre a $P(\cdot|\mathcal{F})(\omega)$ halmazfüggvény 1-re normált mérték a $\mathcal{C}$ algebrán minden $\omega \in \Omega_0$ elemi eseményre.

Először megmutatom, hogy feltehetjük, hogy az 'A tulajdonság'-ot teljesítő rendszerben $\Omega_0 = \Omega$. Ennek érdekében rögzítsünk egy tetszőleges P_0 valószínűségi mértéket a $\mathcal{C}$ algebrán, és terjesszük ki a $Q(C,\omega) = P(\xi \in C|\mathcal{F})(\omega)$ függvényt az $\omega \in \Omega_0$ halmazról az $\omega \in \Omega$ halmazra a következő képlet segítségével: $Q(C,\omega) = P(C|\mathcal{F})(\omega) = P_0(C)$ minden $C \in \mathcal{C}$ halmazra és $\omega \in \Omega \setminus \Omega_0$ elemi eseményre.

Ezután tekintsük minden $\omega \in \Omega$ elemi eseményre a $Q(\cdot,\omega)$ mérték egyértelmű kiterjesztését a $\mathcal{C}$ algebráról a $\mathcal{B}$ σ-algebrára. Azt állítom, hogy az így kapott $Q(B,\omega)$, $B \in \mathcal{B}$, $\omega \in \Omega$, függvény választható, mint a ξ valószínűségi változó $Q(B,\omega) = F(B,\omega) = P(\xi(\omega) \in B|\mathcal{F})(\omega)$ feltételes eloszlása az $\mathcal{F}$ σ-algebra szerint. Ehhez azt kell belátni, hogy a $Q(B,\omega)$ függvény teljesíti a tételben felsorolt (i), (ii) és (iii) tulajdonságot.

Tekintsük azon B halmazok $\mathcal{B}_0$ osztályát, amelyek teljesítik az (i) tulajdonságot. Ekkor $\mathcal{C} \subset \mathcal{B}_0$, és nem nehéz belátni, hogy $\mathcal{B}_0$ monoton halmazosztály. Ezért az

5. lemma alapján $\mathcal{B} \subseteq \mathcal{B}_0$, tehát az (i) tulajdonság teljesül. A (ii) tulajdonság teljesülése nyilvánvaló. A (iii) tulajdonság igazolásához azt kell ellenőrizni, hogy $P(B \cap F) = \int_F Q(C,\omega)\, dP(\omega)$ minden $F \in \mathcal{F}$ és $B \in \mathcal{B}$ halmazra. Ennek érdekében definiáljuk minden $F \in \mathcal{F}$ halmazra a $\mu_{1,F}(B) = P(B \cap F)$ és $\mu_{2,F}(B) = \int_F Q(B,\omega)\, dP(\omega)$, $B \in \mathcal{B}$, halmazfüggvényeket. Mivel mind $\mu_{1,F}$ mind $\mu_{2,F}$ mérték, és $\mu_{1,F}(C) = \mu_{2,F}(C)$ minden $C \in \mathcal{C}$ halmazra, a mértékek kiterjesztésének egyértelműsége miatt egy algebráról az általa generált legszűkebb σ-algebrára a két mérték megegyezik, és a (iii) tulajdonság is teljesül.

Ezután be kell látni az 'A tulajdonság'-ot. Felhasználva azt, hogy (X,ρ) szeparábilis metrikus tér, tudunk választani megszámlálható sok (nyílt) gömböt, amelyek generálják a $\mathcal{B}$ σ-algebrát. (Vehetünk egy mindenütt sűrű megszámlálható halmazt az (X,ρ) metrikus téren és az e halmaz pontjai körüli racionális sugarú gömböket.) Az e halmazokat és az X halmazt tartalmazó legszűkebb algebra szintén megszámlálható sok elemből áll az 1. lemma szerint, és ezt fogjuk választani az A tulajdonságot teljesítő rendszerben a $\mathcal{C}$ algebrának.

Vezessük be a $\mu(B) = P(\xi \in B)$, $B \in \mathcal{B}$, képlettel definiált μ mértéket az (X,ρ) metrikus tér $\mathcal{B}$ Borel σ-algebráján. Ezután, felhasználva a 2. lemma eredményét és azt a tényt, hogy két kompakt halmaz uniója is kompakt halmaz választhatunk minden $C \in \mathcal{C}$ halmazhoz kompakt halmazok olyan $K_{1,C} \subseteq K_{2,C} \subseteq \cdots$ monoton sorozatát, amelyre $K_{n,C} \subseteq C$ minden $1 \le n < \infty$ indexre, és $\lim\limits_{n\to\infty} \mu(K_{n,C}) = \mu(C)$. Ezután vezessük be az összes $C \in \mathcal{C}$ és $K_{n,C}$, $n = 1,2,\dots$, $C \in \mathcal{C}$ halmazt tartalmazó legszűkebb $\mathcal{C}_1$ algebrát. A $\mathcal{C}_1$ algebra szintén csak megszámlálható sok elemet tartalmaz. Azt állítom, hogy létezik olyan $\Omega_0 \subset \Omega$, $P(\Omega_0) = 1$, $\Omega_0 \in \mathcal{F}$ halmaz és egy $Q(C,\omega)$, $C \in \mathcal{C}_1$, $\omega \in \Omega_1$ halmazfüggvény, amely teljesíti a következő tulajdonságokat.

i. $Q(C,\omega) = P(\xi_1(\omega) \in C|\mathcal{F})(\omega)$ minden $C \in \mathcal{C}_1$ halmazra, azaz $Q(C,\omega)$ tekinthető, mint e feltételes valószínűség (csak 1 valószínűséggel meghatározott) verziója.

ii. $Q(C,\omega)$ nem-negatív, additív halmazfüggvény, és $Q(X,\omega) = 1$ a $\mathcal{C}_1$ algebrán minden $\omega \in \Omega_0$ elemi eseményre.

iii. $\lim\limits_{n\to\infty} Q(K_{n,C},\omega) = Q(C,\omega)$ minden $C \in \mathcal{C}$ halmazra és $\omega \in \Omega_0$ elemi eseményre.

Valóban, tekintsük a $Q(C,\omega) = P(\xi_1(\omega) \in C|\mathcal{F})(\omega)$ feltételes valószínűség egyik verzióját minden $C \in \mathcal{C}_1$ halmazra és $\omega \in \Omega$ elemi eseményre. Belátom, hogy el lehet hagyni egy Ω_1, $P(\Omega_1) = 0$, $\Omega_1 \in \mathcal{F}$, halmazt úgy, hogy a (ii) és (iii) tulajdonságok teljesülnek minden $\omega \in \Omega \backslash \Omega_1$ halmazra. Egy ilyen választás teljesíti az (i) tulajdonságot is. A (ii) tulajdonság biztosítása kapcsán vegyük észre, hogy ez megszámlálható sok olyan feltétel biztosítását jelenti, amelyek mindegyike teljesül egy 1 valószínűségű és $\mathcal{F}$ mérhető halmazon. Ezért létezik olyan $\Omega_2 \subset \Omega$, $P(\Omega_2) = 0$, $\Omega_2 \in \mathcal{F}$ halmaz úgy, hogy a (ii) tulajdonság teljesül minden $\omega \notin \Omega_2$ elemi eseményre.

Ha $\omega \in \Omega \setminus \Omega_2$ akkor a $Q(K_{n,C},\omega)$ függvénysorozat monoton nővekszik az n változóban minden $C \in \mathcal{F}$ halmazra, és $Q(K_{n,C},\omega) \le Q(C)$, $n = 1,2,\dots$. Másrészt a feltételes várható érték tulajdonsága miatt

$$\lim_{n\to\infty}\int Q(K_{n,C},\omega)\,dP(\omega)=\lim_{n\to\infty}\mu(K_{n,C})=\mu(C)=\int Q(C,\omega)\,dP(\omega)$$

minden $C\in\mathcal{C}$ halmazra. Innen következik, hogy $\lim\limits_{n\to\infty}Q(K_{n,C},\omega)=Q(C,\omega)$ egy null mértékű $\mathcal{F}$ mérhető halmazt kivéve minden $\omega\in\Omega$ elemi eseményre. Mivel a (iii) feltételben megszámlálható sok ilyen tulajdonság teljesülését írtuk elő, ezért létezik olyan $\Omega_3\subset\Omega$, $P(\Omega_3)=0$, $\Omega_3\in\mathcal{F}$ halmaz úgy, hogy a (iii) tulajdonság teljesül minden $\omega\in\Omega\setminus\Omega_3$ elemi eseményre. Ezért a kívánt tulajdonságok mindegyike teljesül a $Q(C,\omega)$ függvényre az $\Omega_0=\Omega\setminus(\Omega_2\cup\Omega_3)$ halmazon.

Ezután elég megmutatni, hogy az (i), (ii) és (iii) tulajdonságokból következik, hogy a $\mathcal{C}$ algebra teljesíti az A tulajdonságot. A 4. lemma alapján ehhez elég megmutatni azt, hogy minden olyan $C_n\in\mathcal{C}$, $n=1,2,\dots$, $C_1\subseteq C_2\subseteq\cdots$ halmazsorozatra, amelyre $\bigcap\limits_{n=1}^{\infty}C_n=\emptyset$ és $\omega\in\Omega_1$ elemi eseményre $\lim\limits_{n\to\infty}Q(C_n,\omega)=0$.

Ezen állítás igazolása érdekében rögzítsünk egy kis $\varepsilon>0$ számot és válasszunk mindegyik C_n halmazhoz egy olyan $\bar n(n)=\bar n(n,\varepsilon,\omega)$ indexet, amelyre $Q(K_{\bar n(n),C_n},\omega)\ge Q(C_n,\omega)-2^{-(n+1)}\varepsilon$. A $\bigcap\limits_{n=1}^{\infty}C_n=\emptyset$ reláció és a 3. lemma alapján $\bigcap\limits_{n=1}^{\infty}K_{\bar n(n),C_n}=\emptyset$, és létezik olyan n_0 index, amelyre $\bigcap\limits_{n=1}^{n_0}K_{\bar n(n),C_n}=\emptyset$. Mivel

$$C_{n_0}\subset\bigcup_{n=1}^{n_0}(C_{n_0}\setminus K_{\bar n(n),C_n})\cup\left(\bigcap_{n=1}^{n_0}K_{\bar n(n),C_n}\right)$$

innen következik, hogy

$$Q(C_{n_0},\omega)\le\sum_{n=1}^{n_0}Q(C_{n_0}\setminus K_{\bar n(n),C_n},\omega)\le\sum_{n=1}^{n_0}(Q(C_n,\omega)-Q(K_{\bar n(n),C_n},\omega)),$$

és ezért $Q(C_{n_0},\omega)\le\sum\limits_{n=1}^{\infty}\varepsilon 2^{-(n+1)}=\varepsilon$. (Ebben a lépésben kihasználtuk, hogy a $Q(\cdot,\omega)$ halmazfüggvény a $\mathcal{C}_1$ algebrán is additív.) Ezért $Q(C_{n_0},\omega)\le\varepsilon$, és $\limsup\limits_{n\to\infty}Q(C_n,\omega))\le\varepsilon$ minden $\omega\in\Omega_0$ elemi eseményre. Ez minden $\varepsilon>0$ számra igaz, így $\lim\limits_{n\to\infty}Q_n(C_n,\omega)=0$.

A feltételes várható értékek kiszámolásáról szóló tételnek reguláris eloszlások segítségével szintén létezik általánosítása általános terekben definiált valószínűségi változókra. Ennek az alább megfogalmazott eredménynek a bizonyítását fogom ismertetni.

A feltételes várható érték reguláris feltételes eloszlások segítségével való kiszámolásáról szóló tétel általánosítása. *Legyen adva két $(X,\mathcal{U})$ és $(Y,\mathcal{V})$ mérhető tér, egy $(\Omega,\mathcal{A},P)$ valószánűségi mező mező és azon egy $\mathcal{F}\subset\mathcal{A}$ σ-algebra. Legyen* *ezenkívül adva egy ξ, értékeit az $(X,\mathcal{U})$ egy η, értékeit az $(Y,\mathcal{V})$ térben felvevő valószínűségi változó, amelyek közül η $\mathcal{F}$ mérhető, és egy olyan $h(x,y)$ mérhető függvény az $(X\times Y,\mathcal{U}\times\mathcal{V})$ szorzattéren, amelyre $E|h(\xi,\eta)|<\infty$. Tegyük fel ezenkívül, hogy a ξ valószínűségi változónak létezik egy $F(B,\omega)$, $P(\xi(\omega)\in B|\mathcal{F})(\omega)=F(B,\omega)$, $B\subset\mathcal{U}$, $\omega\in\Omega$, reguláris feltételes eloszlása az $\mathcal{F}$ σ-algebrára nézve. Ekkor felírható az*

$$E(h(\xi,\eta)|\mathcal{F})(\omega)=\int h(x,\eta(\omega))F(dx,\omega) \tag{1}$$

azonosság.

A tétel bizonyítása. Azt kell belátni, hogy az (1) formula jobboldalán szereplő integrál $\mathcal{F}$ mérhető valószínűségi változó, és teljesíti a feltételes várható érték definiciójában megadott ii.′) feltételt. Lássuk be a mérhetőségről szóló állítást először abban az esetben, ha $h(x,y)$ egy az $X\times Y$ térben mérhető C halmaz indikátor függvénye.

Ha $C=A\times B$, $A\in\mathcal{U}$, $B\in\mathcal{V}$ alakú, akkor ez az állítás nyilvánvaló. Ugyancsak nyilvánvaló akkor, ha C véges sok ilyen diszjunkt téglalap uniója. Viszont az ilyen halmazok algebrát alkotnak. Ezután viszonylag egyszerűen ellenőrizhető, hogy azon C halmazok, amelyek indikátor függvénye teljesíti a kívánt mérhetőségi feltételt monoton halmazosztályt alkotnak. Ezért az 5. lemma alapján az ilyen halmazok családja tartalmazza az előbb tekintett $A\times B$ alakú halmazok által generált, azaz az $\mathcal{U}\times V$ σ-algebrát.

A mérhető halmazok indikátorfüggvényeire már bizonyított állítást felhasználva megkapjuk, hogy az (1) formula jobboldalán szereplő integrál $\mathcal{F}$ mérhető akkor, ha $h(x,y)$ lépcsős függvény, majd az általános függvényeket lépcsős függvények e függvényhez konvergáló sorozatával közelítve megkapjuk a kívánt $\mathcal{F}$ mérhetőségi állítást minden olyan mérhető $h(x,y)$ függvényre, amelyre $E|h(\xi,\eta)|<\infty$.

A ii.′) feltételt is először mérhető halmazok indikátor függvényeire látjuk be. Ezen belül is először azt az esetet tekintjük, amikor $h(x,y)=I_A(x)I_B(u)$, $A\in\mathcal{U}$, $B\in\mathcal{V}$. Ekkor

$$E(h(\xi,\eta)|\mathcal{F})=E(I_A(\xi)|\mathcal{F})I_B(\eta)=F(A,\omega)I_B(\eta(\omega)$$

és ez nyilván egyenlő az (1) formula jobboldalán szereplő integrállal ebben az esetben.

Egy általános $C\in\mathcal{U}\times\mathcal{V}$ halmaz indikátorfüggvényére a következő relációt kell ellenőrizni.

Rögzítsünk egy tetszőleges $G\in\mathcal{F}$ eseményt, és definiáljuk először a

$$\mu_{1,G}(C)=P(\{\omega\colon\ (\xi(\omega),\eta(\omega))\in A\}\cap F),\quad C\in\mathcal{U}\times\mathcal{V},$$

mértéket az $(X\times Y,\mathcal{U}\times\mathcal{V})$ téren. Ezután vezessük be minden $y\in Y$ pontra és $C\in\mathcal{U}\times\mathcal{V}$ halmazra a C halmazra a következő $C(y)\in\mathcal{U}$ (metszet)halmazt: $C(y)=\{x\colon\ (x,y)\in C\}$. Vezessük be továbbá a $H(y,\omega)=H_C(y,\omega)=F(C(y),\omega)$ valószínűségi változót, ahol $F(\cdot,\omega)$ a ξ valószínűségi változó feltételes eloszlásfüggvénye, feltéve a $\mathcal{F}$ σ-algebrát. (A $H(y,\omega)$ valószínűségi változót úgy definiáltam, hogy ez az (1) függvény jobboldalán definiált integrállal egyenlő akkor, ha $h(\cdot)$ a C halmaz indikátorfüggvénye. Ezután definiáljuk a

$$\mu_{2,G}(C) = \int_G H(\eta(\omega))\, dP(\omega), \quad A \in \mathcal{R}^{k+l}$$

mértéket az $(X \times Y, \mathcal{U} \times \mathcal{V})$ szorzattéren. Azt kell igazolni, hogy $\mu_{1,G}(C) = \mu_{2,G}(C)$ minden $C \in \mathcal{U} \times \mathcal{V}$ halmazra, ami azt jelenti, hogy a ii$'$) azonosság igaz minden ilyen halmaz indikátorfüggvényére.

Ezt az azonosságot már ellenőriztük abban a speciális esetben, amikor $C = A \times B$, $A \in \mathcal{U}$, $B \in \mathcal{V}$, és a kívánt azonosság következik ebből tetszőleges $C \in \mathcal{U} \times \mathcal{V}$ halmazra is a $\mu_{1,G}$ és $\mu_{2,G}$ mértékek kiterjesztésének egyértelműsége miatt a $\mathcal{U} \times \mathcal{V}$ σ-algebrára.

Az általános eset visszavezetése a már bebizonyított állításhoz hasonló, csak egyszerűbb érvelésen alapul. Rögzítsünk egy $G \in \mathcal{F}$ eseményt és definiáljuk minden olyan $h(x,y)$ mérhető függvényre, amelyre $E|h(\xi,\eta)| < \infty$ a

$$K_1(G,h) = \int_G h(\xi(\omega), \eta(\omega))\, dP(\omega)$$

és

$$K_2(G,h) = \int_G \left[\int h(x, \eta(\omega)) F(dx, \omega) \right] dP(\omega)$$

integrálokat. Azt kell belátni, hogy $K_1(G,h) = K_2(G,h)$ minden olyan h függvényre, amelyre $E|h(\xi,\eta)| < \infty$. Ezt az azonosságot tudjuk abban az esetben, ha $h(\cdot)$ az $X \times Y$ tér egy mérhető halmazának indikátor függvénye. Mivel mind $K_1(G,h)$ mind $K_2(G,h)$ a h függvénynek lineáris függvénye (rögzített $G \in \mathcal{F}$ eseményre) ezért ez az állítás igaz minden lépcsős függvényre. Ezután limeszeléssel megkapjuk azt nem-negatív majd tetszőleges függvényekre. Az (egyszerű) részletek kidolgozását elhagyom.

A *feltételes várható érték kiszámolásáról szóló tétel egy változata* eredmény hasonlóan bizonyítható, ezért ennek tárgyalását elhagyom.