Az október 17-i szeminárium témája
Rövid összefoglaló

Feltételes eloszlások

A feltételes valószínűség és feltételes várható érték nulla valószínűségi feltételek mellett is definiálják, de az így definiált fogalmak a valószínűségszámítás legnehezebb fogalmai közé tartoznak. Ahhoz, hogy e fogalmakat jobban megértsük, próbáljuk először megérteni azt, hogy milyen szemléletes képet akarunk ennek a definiciónak a segítségével megfogalmazni. Ennek érdekében 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 10×25=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 felvetni a következő kérdést. Adva egy esemény vagy egy valószínűségi változó, akkor érdekelhet minket ennek az eseménynek a valószínűsége vagy valószínűségi változónak várható értéke. Előfordulhat, hogy más eseményekenek bekövetkezését vagy be nem következését más valószínűségi változók felvett értékét meg tudjuk figyelni, és ezen plusz információ ismeretében akarjuk megbecsülni a minket érdeklő esemény valószínűségét vagy valószínűségi változó várható értékét. Ekkor természetes olyan becslést adni, amelyik ezeket a plusz információkat is figyelembe veszi. Az előző paragrafusban is ilyen kérdést fogalmaztunk meg. Ott meg akartuk becsülni tíz valószínűségi változó összegének a várható értékét azon feltétel mellett, hogy az első két változó összegének az értéke ismert. Ennek a kérdésnek a természetes általánosítása vezet a feltételes valószínűség és feltételes várható érték fogalmához.

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, hogy

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

Ez az azonosság azonban semmitmondó, mert az azonosság mindkét oldala nulla. Viszont ez egy tartalmas állítássá válik, ha ezt az azonosságot kiintegráljuk. Ez azt sugallja, hogy teljesülnie kell a

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

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 Radon–Nikodym tételből következik, hogy létezik olyan $f_A(\cdot,\cdot,\cdot)$ függvény mely teljesíti a $(*)$ azonosságot minden mérhető B halmazra, és ezek az azonosságok lényegében egyértelműen meghatározzák ezt az f_A feltételes valószínűség függvényt. Annak érdekében, hogy megértsük a lényegében egyértelműen kitétel értelmét jegyezzük meg, hogy ha egy f_A függvény teljesíti a $(*)$ azonosságok rendszerét, akkor megváltoztatva ezt az f_A függvényt egy a $(\xi_1,\dots,\xi_k)$ valószínűségi változók F eloszlása által meghatározott mérték szerint null mértékű halmazon, akkor olyan új függvényt kapunk, mely szintén teljesíti a fenti egyenletrendszert. Az, hogy a $(*)$ azonosságot minden B halmazra teljesítő f_A függvény lényegében egyértelműen meg van határozva azt jelenti, hogy ha két f_A és $\bar f_A$ függvény teljesíti a $(*)$ azonosságot a k-dimenziós euklideszi tér minden B Borel mérhető halmazára, akkor $f_A(x_1,\dots,x_k) = \bar f_A(x_1,\dots,x_k)$ a $(\xi_1,\dots,\xi_k)$ véletlen vektor F eloszlása szerint meghatározott mérték szerint majdnem minden $(x_1,\dots,x_k)$ pontban.

Hasonló gondolatmenet segítségével kapjuk meg egy η valószínűségi változó, $E|\eta| < \infty$, $g_\eta(x_1,\dots,x_k) = E(\eta|\xi_1 = x_1,\dots,\xi_k = x_k)$ feltételes várható érték definicióját feltéve a $\xi_1 = x_1,\dots,\xi_k = x_k$ feltételeket. Ezt az

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

relációk definiálják, ahol $I(B)$ jelöli egy $B \subset \Omega$ halmaz indikátor függvényét, A a k-dimenziós R^k euklideszi tér tetszőleges Borel mérhető részhalmaza, $F(x_1,\dots,x_k)$ 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 mérték szerint egy valószínűséggel meg van határozva szintén a Radon–Nikodym tétel segítségével láthatjuk be.

Mielőtt megtárgyalnánk a fent említett Radon–Nikodym tételt, megfogalmazzuk a feltételes valószínűség és várható érték fogalmát némileg általánosabb esetben. Előfordulhat ugyanis, hogy az előzetes információink, melyek alapján egy halmaz valószínűségére vagy egy valószínűségi változó értékére becslést akarunk adni nem tömöríthető véges sok valószínűségi változó megfigyelt értékébe. Ahhoz, hogy a feltételes valószínűség és feltételes várható érték fogalmát természetes módon meg tudjuk fogalmazni ebben az 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 és egy A esemény vagy egy η valószínűségi változó, melyre $E|\eta| < \infty$, akkor definiáljuk a $P(A|\mathcal{F})(\omega)$ feltételes valószínűséget illetve $E(\eta|\mathcal{F})(\omega)$ feltételes várható értéket feltéve a $\mathcal{F}$ σ-algebrát. Az, hogy a $\mathcal{F}$ σ algebra ismeretében akarjuk megbecsülni az A halmaz valószínűségét illetve η valószínűségi változó értékét a $P(\mathbf{A}|\mathcal{F})(\omega)$ és $E(\eta|\mathcal{F})(\omega)$ feltételes valószínűség definiciójában azt jelenti, hogy

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

i$'$). Az $E(\eta|\mathcal{F})(\omega)$ feltételes várható érték $\mathcal{F}$ mérhető függvény.

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.

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 mely 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, mely teljesíti az i.$'$) és ii$'$.) tulajdonságokat.*

Tisztázni kell, hogy a fenti definició tényleg meghatározza a $P(A|\mathcal{F})(\omega)$ feltételes valószínűséget és $E(\xi|\mathcal{F})(\omega)$ feltételes várható értéket. Ez az alább megfogalmazandó Radon–Nikodym tétel következménye. 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.

Annak érdekében, hogy a Radon–Nikodym tételt megfogalmazhassuk előbb be kell vezetni a következő definiciót.

Definició: Egy előjeles mérték abszolut folytonossága egy mérték szerint. *Legyen μ 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 $C \in \mathcal{F}$ halmazra, melyre $\mu(C) = 0$, a $\nu(C) = 0$ reláció is teljesül.

Radon–Nikodym tétel. *Legyen adva egy Ω tér, rajta egy $\mathcal{F}$ σ-algebra, továbbá ezen a $\mathcal{F}$ σ-algebrán egy μ (véges) mérték és ν (véges) előjeles mérték. 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, melyre $\int |f(\omega)|\,d\mu(\omega) < \infty$, és $\int_C f(\omega)\,d\mu(\omega) < \infty$ minden $C \in \mathcal{F}$ halmazra. Továbbá ez az $f(\omega)$ függvény egyértelmű 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.*

Megjegyzés: A Radon–Nikodym tétel tipikus existencia tétel. Az általános esetben nem ad módszert arra, hogyan lehet a keresett $f(\omega)$ Radon–Nikodym deriváltat effektive kiszámolni. Ez a probléma a feltételes eloszlás és feltételes várható érték fogalmában is megjelenik, és ez teszi a feltételes várható érték fogalmát nehézzé és népszerűtlenné. Szerencsére a legfontosabb statisztikai problémákban, ahol ez előkerül effektív módon tudunk vele számolni.

Jegyezzük meg, hogy az a kitétel, hogy a Radon–Nikodym tételben meghatározott $f(\omega)$ függvény csak μ majdnem mindenütt van meghatározva természetes. Ha egy $f(\omega)$ függvény teljesíti a Radon–Nikodym tételt, és a μ mérték szerint null mértékű halmazon megváltoztatjuk ezt a függvényt, akkor ez a megváltoztatott függvény is teljesíti a Radon–Nikodym tételben megkövetelt tulajdonságokat.

A Radon–Nikodym tételből egyszerűen következik a feltételes várható érték létezése. Valóban, ha adva van egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, azon egy $\mathcal{F} \subset \mathcal{A}$ σ-algebra valamint egy $\xi(\omega)$ valószínűségi változó, melyre $E|\xi| < \infty$, akkor alkalmazzuk a Radon–Nikodym tételt a következő választással: Legyen a μ mérték a P valószínűségi mérték, pontosabban annak megszorítása az $\mathcal{F}$ σ-algebrára, és definiáljuk a ν előjeles mértéket az $\mathcal{F}$ σ-algebrán a következő formula segítségével: $\nu(F) = \int_F \eta(\omega)\,dP(\omega)$ minden $F \in \mathcal{F}$ halmazra. Ekkor a Radon–Nikodym tételt alkalmazhatjuk a μ mértékre és a ν előjeles mértékre, mert a ν előjeles mérték abszolut folytonos a μ mérték szerint. E tétel szerint létezik olyan $f(\omega)$ *$\mathcal{F}$ mérhető függvény,* melyre $\nu(F) = \int_F f(\omega)\mu(\,d\omega)$ minden $F \in \mathcal{F}$ halmazra. Ez pedig azt jelenti, hogy a Radon–Nikodym tétel segítségével konstruált $f(\omega)$ függvény az $E(\eta|\mathcal{F})(\omega)$ feltételes várható érték.

A $g_\eta(x_1, \dots, x_k) = E(\eta|\xi_1 = x_1, \dots, \xi_k = x_k)$ feltételes várható értéket hasonlóan definiálhatjuk csak ekkor más μ mértékkel és ν előjeles mértékkel 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, melyre $\nu(B) = \int_B g_\eta(x_1, \dots, x_k)\mu(\,dx_1, \dots, \,dx_k)$. Ekkor 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.

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 $P(A|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k) = E(\chi_A(\omega)|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k)$. A továbbiakban egyrészt meg kívánjuk tárgyalni az $E(\eta|\mathcal{F})(\omega)$ és $E(\eta|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k)$ 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 az a legszűkebb σ-algebra, melyben az összes $\xi_1(\omega), \dots, \xi_k(\omega)$ valószínűségi változó mérhető függvény. Ezenkívül felsoroljuk a feltételes várható érték legfontosabb tulajdonságait, és azt, hogy hogyan lehet számolni a feltételes várható értékkel. Ez azért is fontos, mivel a feltételes várható értéket csak implicit módon (a Radon–Nikodym tétel segítségével) tudjuk definiálni. Ez a fő oka annak, hogy csak nehezebben tudunk a feltételes várható érték segítségével számolni.

Az első kérdés megértéséhez szükségünk van a következő (nem triviális) mértékelméleti eredményre.

Tétel. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, 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, melyre $\eta(\omega) = f(\xi_1(\omega), \dots, \xi_k(\omega))$.*

(Megjegyezzük, hogy a tételben tekintett $f(x_1, \dots, x_k)$ Borel mérhető függvény nincs egyértelműen meghatározva.)

Legyen $\mathcal{F}$ egy $\xi_1(\omega), \dots, \xi_k(\omega)$ valószínűségi változók által generált σ-algebra, és $\eta(\omega)$, $E|\eta| < \infty$ tetszőleges valószínűségi változó. Ekkor az $E(\eta|\mathcal{F})(\omega)$ $\mathcal{F}$ mérhető valószínűségi változó az előző tétel alapján felírható $E(\eta(\omega)|\mathcal{F}) = g_\eta(\xi_1(\omega), \dots, \xi_k(\omega))$ alakban, ahol $g_\eta(x_1, \dots, x_k)$ k-változós Borel mérhető függvény. Azt állítjuk, hogy ekkor $E(\eta|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k) = g_\eta(x_1, \dots, x_k)$. Ehhez azt kell ellenőrizni integráltranszformáció segítségével, hogy a ii$'$.) relációból következik a $(**)$ reláció, ha a ii$'$.) relációban $B = \{\omega\colon (\xi_1(\omega), \dots, \xi_k(\omega)) \in A\}$ halmazt választunk. Hasonlóan be lehet látni, hogy ha $E(\eta|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k) = g_\eta(x_1, \dots, x_k)$, akkor $E(\eta|\mathcal{F})(\omega) = g_\eta(\xi_1(\omega), \dots, \xi_k(\omega))$. Ehhez, azt kell észrevenni, hogy az előbb kimondott tétel alapján egy $B \in \mathcal{F}$ halmazra létezik olyan A Borel mérhető halmaz az R^k k-dimenziós téren, melyre $B = \{\omega\colon (\xi_1(\omega), \dots, \xi_k(\omega)) \in A\}$. Ezután be lehet látni az integráltranszformáció segítségével, hogy a $(**)$ relációból következik a ii$'$.) relációból.

Felsoroljuk a feltételes várható érték legfontosabb tulajdonságait. Ezek bizonyítását, ami viszonylag egyszerű elhagyjuk. Ez valószínűleg szerepelt az előadáson. Rögzítsünk egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt. Az alábbi tulajdonságokban szereplő valószínűségi változók ezen a valószínűségi mezőn vannak értelmezve.

1. Ha $\sum\limits_{k=1}^{\infty} |a_k| E|\xi_k| < \infty$, $\mathcal{F} \subset \mathcal{A}$ tetszőleges σ-algebra, akkor
$$E\left(\left.\sum_{k=1}^{\infty} a_k \xi_k \right| \mathcal{F}\right)(\omega) = \sum_{k=1}^{\infty} a_k E(\xi_k|\mathcal{F})(\omega) \quad \text{majdnem minden } \omega \in \Omega \text{ pontban.}$$

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, melyre $P(a \le \xi \le b) = 1$ 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.
4. $E(\xi|\mathcal{A})(\omega) = \xi(\omega)$. 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.

A fenti állítások egyik érdekes következménye a következő tulajdonság. Tekintsünk egy ξ valószínűségi változót, melyre $E\xi^2 < \infty$. Ekkor a

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

felbontás olyan, hogy az összegben szereplő két valószínűségi változó egymásra merőleges az $(\Omega, \mathcal{A}, P)$ téren négyzetesen integrálható függvények (Hilbert) terében, azaz

$$E(\xi|\mathcal{F})(\omega)\left(\xi(\omega) - E(\xi|\mathcal{F})(\omega)\right) = 0.$$

Ebből a tényből, illetve a Hilbert terek néhány (egyszerű) alapvető tulajdonsága segítségével látható, hogy

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

A hetedik pontban megfogalmazott tulajdonság fontos mind a valószínűségszámításban mind a statisztikában. Úgy lehet interpretálni, hogy az $E(\xi|\mathcal{F})(\omega)$ feltételes várható érték a ξ valószínűségi változó legjobb közelítése $\mathcal{F}$ mérhető valószínűségi változóval (az $L_2(\Omega, \mathcal{A}, P)$ térben).

$* * *$ Ez a rész egyelőre elhagyható. $* * *$

A következő tételnek a bizonyítását is elhagyjuk. Ennek a bizonyítása nehezebb, mint a fent felsorolt tulajdonságok bizonyítása. Ez az eredmény néhány finomabb vizsgálatban nagyon hasznos.

Tétel a feltételes reguláris eloszlás 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 $F(B,\omega)$, $F(B,\omega)\colon \mathcal{B}\times\Omega\to R^1$, függvényt, melyre

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 tekinthető mint a csak majdnem minden $\omega\in\Omega$ pontban meghatározott $P\left((\xi_1(\omega),\dots,\xi_k(\omega))\in B|\mathcal{F}\right)(\omega)$ feltételes valószínűség egyik verziója.*

Megjegyezzük, hogy az előző tétel bizonyítása nem triviális, nem elegendő pusztán a feltételes valószínűség definicióját kihasználni, hanem a bizonyításban fontos szerepe van az euklideszi tér szép topológiai tulajdonságainak is. A fő probléma az, hogy bár a feltételes várható érték 1.) pontban megfogalmazott tulajdonsága alapján a ii.) pontban megkövetelt σ-additivitás érvényes rögzített diszjunkt $B_n\in\mathcal{B}$, $n=1,2,\dots$, halmazokra majdnem minden $\omega\in\Omega$ pontban ahhoz, hogy a ii.) pontban megfogalmazott állítást belássuk szükségünk van kontinum sok ilyen feltételt biztosítani (majdnem) minden $\omega\in\Omega$ pontban. Ez viszont sokkal erősebb állítás mint az 1.) pontban megfogalmazott tulajdonság.

Egy véletlen vektor függvényének a várható értékét ki lehet számolni mint ennek a függvénynek a véletlen vektor eloszlása szerinti integrált. A reguláris feltételes eloszlásról szóló tétel azért hasznos, mert lehetővé teszi ennek az eredménynek a természetes általánosítását a feltételes várható érték kiszámiításáról. Ez a tartalma a következő tételnek, melyet itt nem bizonyítunk, bár a bizonyítás, amelyik standard módon elvégezhető nem túl nehéz. Ebben az eredményben nagyon fontos az, hogy a reguláris feltételes eloszlás minden $\omega\in\Omega$ pont esetén mérték, ezért minden $\omega\in\Omega$ pontban definiálhatunk az $F(\cdot,\omega)$ reguláris eloszlás szerinti Lebesgue integrálokat.

Tétel. *Legyen adva egy $(\Omega,\mathcal{A},P)$ σ-algebra, 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 $(\eta_1,\dots,\eta_k)$ véletlen vektor $\mathcal{F}$ mérhető, és jelölje $F(B,\omega)$, $B\subset R^k$ Borel mérhető függvény, $\omega\in\Omega$ a $(\xi_1,\dots,\xi_k)$ véletlen vektor reguláris feltételes eloszlása 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, melyre $E|h(\xi_1,\dots,\xi_k,\eta_1,\dots,\eta_l)|<\infty$. 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:*

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

* * *

A matematikai statisztikában bizonyos vizsgálatokban szükség van feltételes eloszlásokkal való számolásra. Az itt felmerülő kérdések azonban egyszerűbbek. Olyan tipusú kérdések merülnek fel, melyekben adott egy $(\xi_1,\dots,\xi_k,\xi_{k+1},\dots,\xi_{k+l})$ véletlen vektor, melynek létezik $f(x_1,\dots,x_{k+l})$ sűrűségfüggvénye, és ki akarjuk számítani a

$$(\xi_1,\dots,\xi_k,\xi_{k+1},\dots,\xi_{k+l})$$

véletlen vektor egy $h(\xi_1,\dots,\xi_k,\xi_{k+1},\dots,\xi_{k+l})$ függvényének feltételes várható értékét a $\xi_{k+1}=x_{k+1},\dots,\xi_{k+l}=x_{k+l}$ feltételek mellett. Ez azért egyszerűbb az előzőleg vizsgált kérdéseknél, mert ebben az esetben a felmerülő feltételes eloszlásokat explicit módon kiszámíthatjuk. Azt állítjuk, hogy ebben az esetben a $(\xi_1,\dots,\xi_k)$ feltételes sűrűségfüggvénye feltéve a $\xi_{k+1}=x_{k+1},\dots,\xi_{k+l}=x_{k+l}$ feltételeket

$$f(x_1,\dots,x_k|x_{k+1},\dots,x_{k+l})=\frac{f(x_1,\dots,x_k,x_{k+1},\dots,x_{k+l})}{g(x_{k+1},\dots,x_{k+l})},$$

ahol

$$g(x_{k+1},\dots,x_{k+l})=\int f(x_1,\dots,x_k,x_{k+1},\dots,x_{k+l})\,dx_1,\dots\,dx_k,$$

azaz a $(\xi_{k+1},\dots,\xi_{k+l})$ véletlen vektor sűrűségfüggvénye. Ez azt jelenti, hogy tetszőleges Borel mérhető $A\subset R^k$ halmazra

$$\begin{aligned}&P\left(\{\omega\colon(\xi_1(\omega),\dots,\xi_k(\omega))\in A\}|\xi_{k+1}(\omega)=x_{k+1},\dots,\xi_{k+l}(\omega)=x_{k+l}\right)\\&\qquad=\int_A f(x_1,\dots,x_k|x_{k+1},\dots,x_{k+l})\,dx_1\dots\,dx_k.\end{aligned}$$

Illetve a $(*)$ formula alapján azt kell ellenőrizni, hogy minden $A\subset R^k$ és $B\in R^l$ Borel mérhető halmazpárra

$$\begin{aligned}&P\left(\{\omega\colon(\xi_1(\omega),\dots,\xi_k(\omega))\in A\}\cap\{\omega\colon(\xi_{k+1}(\omega),\dots,\xi_{k+l}(\omega))\in B\}\right)\\&=\int\limits_B\left[\int\limits_A f(x_1,\dots,x_k|x_{k+1},\dots,x_{k+l})\,dx_1\dots\,dx_k\right]g(x_{k+1},\dots,x_{k+l})\,dx_{k+1}\dots dx_{k+l}.\end{aligned}$$

Ez az azonosság viszont érvényes, mert

$$\begin{aligned}&\int_B\left[\int_A f(x_1,\dots,x_k|x_{k+1},\dots,x_{k+l})\,dx_1\dots\,dx_k\right]g(x_{k+1},\dots,x_{k+l})\,dx_{k+1}\dots\,dx_{k+l}\\&\qquad=\int_{A\times B}f(x_1,\dots,x_k,x_{k+1},\dots,x_{k+l})\,dx_1\dots\,dx_k\,dx_{k+1}\dots\,dx_{k+l}\\&\qquad=P\left(\{\omega\colon(\xi_1(\omega),\dots,\xi_k(\omega))\in A\}\cap\{\omega\colon(\xi_{k+1}(\omega),\dots,\xi_{k+l}(\omega))\in B\}\right).\end{aligned}$$

Be lehet látni, hogy egy $h(x_1,\dots,x_k,x_{k+1},\dots,x_{k+l})$ függvényre

$$\begin{aligned}&h(\xi_1,\dots,\xi_k,\xi_{k+1},\dots,\xi_{k+l})|\xi_{k+1}=x_{k+1},\dots,\xi_{k+l}=x_{k+l})\\&\qquad=\int h(x_1,\dots,x_k,x_{k+1},\dots,x_{k+l})\frac{f(x_1,\dots,x_k,x_{k+1},\dots,x_{k+l})}{g(x_{k+1},\dots,x_{k+l})}\,dx_1\dots\,dx_k.\end{aligned}$$

Ennek a formulának a bizonyítását, mely úgy történhet, hogy azt előbb a legegyszerűbb alakú h függvényekre majd alkalmas határátmenettel általános függvényekre bizonyítjuk, elhagyjuk. A fent tárgyalt formulák segítségével be lehet bizonyítani néhány érdekes formulát a feltételes eloszlásokra, ezt azonban időhiány miatt nem tesszük.