A Valószínűségszámítás II. előadássorozat tízenharmadik előadása.
2002. december 10.

Feltételes valószínűség és várható érték *(első rész)*

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 $10\times25=250$ 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 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. Figyeljünk fel arra, hogy az első két lámpa összélettartama folytonos eloszlású valószínűségi változó, ezért nulla annak a valószínűsége, hogy 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 tárgyalt feltételes valószínűség (és feltételes várható érték) fogalma csak abban az esetben volt értelmes, ha a feltétel nem nulla valószínűségű. Szeretnénk a feltételes valószínűség és feltételes várható érték fogalmának olyan általánosítását adni, amelyik lehetővé teszi, hogy a feltételes valószínűségről és feltételes várható értékről beszélhessünk nulla valószínűségi feltételek mellett is, és a bevezetett fogalmak megfeleljenek szemléletes képünknek, melyet első pillanatban csak homályosan tudunk megfogalmazni. Lehetséges ilyen definiciót adni, de ehhez szükségünk van a mértékelmélet néhány mély eredményére. Mostani előadásom célja e feltételes valószínűség és várható érték fogalmának ismertetése az általános esetben, illetve legfontosabb tulajdonságainak 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 (formálisan a

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

és az A halmaz választásával), hogy

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

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

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

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, az alább ismertetendő 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 Lebesgue–Stieltjes 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 Lebesgue–Stieltjes mérték szerint majdnem minden $(x_1,\dots,x_k)$ pontban.

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 értékének a definicióját feltéve a $\xi_1 = x_1, \dots, \xi_k = x_k$ feltételeket. Ez olyan (Borel-mérhető) $g_\eta(x_1, \dots, x_k)$ függvény, melyre az

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

relációk teljesülnek a k-dimenziós R^k euklideszi tér tetszőleges A 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.

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 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 $P(\mathbf{A}|\mathcal{F})(\omega)$ 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 (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 $C \in \mathcal{F}$ 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 (σ-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.*

1. megjegyzés: A Radon–Nikodym tétel tipikus egzisztencia 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. (Ehhez a kérdéshez később még visszatérünk.)

2. megjegyzés: Legyen adva egy $(\Omega, \mathcal{A})$ mérhető tér, azon egy (σ-additív μ mérték, és egy előjeles ν mérték. (Tegyük fel az egyszerűség kedvéért, hogy μ σ-véges és ν véges. Az az integrál tulajdonságaiból azonnal látszik, hogy amennyiben a ν mérték előáll a Radon–Nikodym tételben leírt módon, akkor az abszolut folytonos a μ mértékre nézve. Létezik a két mérték kapcsolatának részletesebb leírása is, mely arról az esetről is szól, amikor a ν mérték nem abszolut folytonos a μ mértékre nézve. Nevezetesen, mindig igaz, hogy az Ω halmaz felbontható $\Omega = B \cup (\Omega \setminus B)$ alakba, $B \in \mathcal{A}$ úgy, hogy

$\mu(\Omega \setminus B) = 0$ és minden $C \subset B$ halmazra $\nu(C) = 0$, ha $\mu(C) = 0$. Ez azt jelenti, hogy, ha a μ és ν mértéket, illetve előjeles mértéket megszorítjuk a B halmazra, akkor ott ν abszolut folytonos a μ mértékre, míg a komplementer halmaznak a μ mértéke nulla, míg a ν mértékre ez a tulajdonság nem feltétlenül teljesül. Ez azt jelenti, hogy be lehet vezetni a ν_a és $\nu_{\rm sz}$ előjeles mértékeket a $\nu_a(A) = \nu(A \cap B$ és $\nu_{\rm sz} = \nu(A \cap (\Omega \setminus B))$, $A \in \mathcal{A}$ képlettel. Ekkor $\nu = \nu_a + \nu_{\rm sz}$, a ν_a mérték abszolut folytonos a μ mértékre nézve, ezért alkalmazható rá a Radon–Nikodym tétel. A $\nu_{\rm sz}$ mérték szinguláris a μ mértékre nézve, ami azt jelenti, hogy létezik egy olyan $B \in \mathcal{A}$ halmaz, melyre $\mu(B) = 0$ viszont $\nu_{\rm sz}(C) = 0$ minden $C \subset \Omega \setminus B$ halmazra. Egy ν előjeles mérték felbontása egy abszolut folytonos és szinguláris komponensre egyértelmű.

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.

Közbevetett megjegyzés: Annak érdekében, hogy jobban megértsük a Radon–Nikodym tételt, felidézek néhány eredményt, ami része a Mértékelmélet anyagának, és segíthet jobban megérteni ezt az eredményt. A tétel (klasszikus) bizonyítása, melyet nem dolgozunk ki az alábbi, az irodalomban Hahn-féle felbontási tételnek nevezett eredményen alapul. Ez olyan tényt fejez ki, hogy tetszőleges előjeles mértéknek van egy pozitív és negatív része, és az előjeles mérték ezek összegeként írható fel.

Hahn-féle felbontási tétel. *Legyen adva egy $(X, \mathcal{A})$ mértéktér, azaz egy X halmaz, és annak „mérhető” halmazai, melyek egy $\mathcal{A}$ σ-algebrát alkotnak, és egy ν σ-additív előjeles mérték az $\mathcal{A}$ σ-algebrárán. (Ez azt jelenti, hogy minden $A \in \mathcal{A}$ halmaz előállítható véges sok vagy megszámlálható sok $A_n \in \mathcal{A}$ halmaz uniójaként úgy, hogy $A = \bigcup A_n$, $|\nu(A_n)| < \infty$ és vagy $\nu(B) < \infty$ vagy $\nu(C) > -\infty$ minden $C \in \mathcal{A}$ halmazra. Ekkor létezik olyan $B \in \mathcal{A}$ halmaz, melyre $\nu(C) \geq 0$ minden $C \subset B$, $C \in \mathcal{A}$ és $\nu(D) \leq 0$ minden $D \subset X \setminus B$, $D \in \mathcal{A}$ halmazra.*

Megjegyzés. Tetszőleges $A \in \mathcal{A}$ halmazt felírhatunk $A = (A \cap B) \cup (A \setminus B)$ alakban. Ekkor $\nu(A) = \nu(A \cap B) + \nu(A \setminus B$, $\nu(A \cap B) \geq 0$ és $\nu(A \setminus B) \leq 0$. Ezért bevezethetjük a $\nu^+(A) = \nu(A \cap B)$ és $\nu^-(A) = \nu(A \setminus B)$ mennyiségeket minden $A \in \mathcal{A}$ halmazra. A ν^+ mértéket a ν előjeles mérték pozitív részének, a ν^- (negatív) mértéket pedig a ν előjeles mérték negatív részének nezezzük.

Legyen adva egy μ véges mérték és egy ν előjeles mérték ugyanazon az $(X, \mathcal{A})$ mértéktéren. Tegyük fel, hogy ν abszolut folytonos a μ mértékre nézve. Hogyan tudjuk megtalálni a Banach–Hahn tételben szereplő a ν előjeles mértéknek a μ mérték szerinti Radon–Nikodym deriváltját? A ν mérték Hahn felbontása megadja, hol kell ezt a Radon–Nikodym deriváltat pozítívnak és hol kell negatívnak választani. Hasonlóan véve egy tetszőleges c valós számot, és tekintve a $\nu - c\mu$ mérték Hahn-felbontását, meg tudjuk mondani, hol lesz a Radon–Nikodym derivált nagyobb vagy egyenlő, mint c. A bizonyítás ennek a gondolatnak a következetes végigviteléből áll. A fő probléma az, hogy a Hahn-felbontás csak egzisztecia tétel, ezért nem teszi lehetővé, hogy egyszerű módszert adjunk a Radon–Nikodym derivált kiszámítására.

Különösen érdekes, és a mértékelméletben külön vizsgált eset az, amikor a számegyenesen definiált ν mértékeket tekintjük, és a μ mérték a λ Lebesgue mérték. Ekkor egy ν abszolut folytonos előjeles mérték előállítható $\nu([a,b]) = \int_a^b f(u)\,du$ alakban, és csak az abszolut folytonos mértékek írhatóak ilyen alakban. Ha egy ν mérték felírható a fenti alakban, akkor érvényes az $\int g(u)\nu(\,du) = \int g(u)f(u)\,du$ azonosság is minden mérhető $g(\cdot)$ függvényre. (Ez az azonosság úgy értendő, hogy a két oldalon szereplő integrál egyszerre létezik vagy nem létezik.) Tekintsünk a továbbiakban olyan előjeles mértékeket a számegyenesen, melyeknek mind a negatív mind a pozitív része véges. Jegyezzük meg, hogy ezen megszorítás után kölcsönösen egyértelmű megfeleletetés létesíthető a számegyenesen értelmezett ν előjeles mértékek és a számegyenesen definiált korlátos változású $G(\cdot)$ függvények között, ha azonosítunk két olyan függvényt, melyek különbsége konstans. (Ezt a megfeletetést megkaphatjuk, hogy egy függvényhez hozzárendeljük az általa meghatározott Lebesgue–Stieltjes mértéket, egy ν mértékhez pedig a $G(x) = \nu([-\infty, x))$ függvényt rendeljük hozzá. Igaz az a tétel is, mely szerint egy függvény akkor és csak akkor korlátos változású, ha előáll két korlátos monoton függvény különbségeként. Ez az eredmény megfelel a számegyenesen definiált előjeles mértékek Hahn-féle felbontásának.) Egy (korlátos változású) függvényt akkor és csak akkor nevezünk abszolut folytonosnak, ha az általa definiált előjeles mérték abszolut folytonos. Természetes módon felmerül a kívánság, hogy adjunk lehetőleg minél áttekinthetőbb jellemzést az abszolut folytonos mértékekről (függvényekről) a számegyenesen, és jellemezzük a különböző lehetséges tulajdonságú mértékeket.

Be lehet látni, hogy egy G korlátos változású függvény akkor és csak akkor abszolut folytonos, ha minden $\varepsilon > 0$ intervallumra létezik olyan $\delta > 0$ szám, melyre igaz, hogy tetszőleges olyan diszjunkt $[a_k, b_k)$ intervallumokra $1 \le k \le n$, (az n szám tetszőleges), melyekre $\sum\limits_{k=1}^{n} (b_k - a_k) < \delta$, teljesül a $\left|\sum\limits_{k=1}^{n} (G(b_k) - G(a_k))\right| < \varepsilon$ egyenlőtlenség is, vagy ami ezzel ekvivalens, $\left|\nu_G\left(\bigcup\limits_{k=1}^{n}([a_k, b_k)\right)\right| < \varepsilon$, ha $\lambda\left(\bigcup\limits_{k=1}^{n}([a_k, b_k)\right)\varepsilon$, ahol ν_G jelöli a G függvény által indukált Lebesgue–Stieltjes mértéket. Egy előjeles mértéket diszkrétnek nevezzünk, ha véges vagy megszámlálható pontba van koncentrálva, azaz egy véges vagy megszámlálható halmaz komplementerének a mértéke nulla. Bevezetnek még egy harmadik fogalmat is: Egy előjeles mértéket szingulárisnak neveznek, ha minden pontjának a mértéke nulla, de van a számegyenesnek olyan mérhető részhalmaza, melynek előjeles mértéke nem nulla. Be lehet látni, hogy ilyen szinguláris előjeles mérték valóban létezik, sőt az is igaz, hogy tetszőleges előjeles mérték egyértelműen felbontható egy abszolut folytonos, egy diszkrét és egy abszolut folytonos előjeles mérték összegeként. A szinguláris mértékek létezése okozza, hogy általános problémák tárgyalásában szükséges Lebesgue-Stieltjes integrálokkal is számolni. Egy abszolut folytonos előjeles mérték szerinti integrál átírható mint alkalmas Lebesgue mérték szerinti integrál, a diszkrét mértékek szerinti integrálok valójában véges vagy megszámlálható összeget jelentenek, de egy szinguláris mérték szerinti integrál esetében nincsen hasonló egyszerűsítési lehetőség.

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 alkalmazhatjuk a Radon–Nikodym tételt a μ mértékre és a ν előjeles mértékre, mert a ν előjeles mérték abszolut folytonos a μ mérték szerint. Ugyanis, $\mu(A) = P(A) = 0$ esetén $\mu(A) = \int_A \eta(\omega) P(\,d\omega) = 0$. A Radon–Nikodym 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. (Az állítás jobb megértése érdekében jegyezzük meg, hogy az $\eta(\omega)$ valószínűségi változó azért nem választható mindig az $f(\omega)$ valószínűségi változónak, mert igaz ugyan, hogy ez a függvény teljesíti a kívánt azonosságot minden $F \in \mathcal{F}$ halmazra teljesíti, de nem feltétlenül $\mathcal{F}$ mérhető. (Eredetileg csak annyit tudunk, hogy η $\mathcal{A}$ mérhető, és $\mathcal{F} \subset \mathcal{A}$.)

Példa: Tekintsük a következő példát a definició jobb megértése érdekében: Legyen az $(\Omega, \mathcal{A}, P)$ valószínűségi mező az $\Omega = [0,1] \times [0,1]$ egyszégnégyzet, rajta a szokásos A Borel σ-algebra és $P = \lambda$ a Lebesgue mérték az egységnégyzet Borel-mérhető részhalmazain. 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 integrálható) $f(x,y)$ függvényt az egységnégyzeten, (az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn) és számoljuk ki az $E(f(x,y)|\mathcal{F})$ valószínűégi 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 mérhető), ezért $\mathcal{F}$) mérhető. Azt állítom, hogy $E(f(x,y)|\mathcal{F}) = g(x,y)$. Ehhez azt kell még 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.

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 $P(A|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k) = E(I_A(\omega)|\xi_1(\omega) = x_1, \dots, \xi_k(\omega) = x_k)$. 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 $\eta(\omega) = f(\xi_1(\omega), \dots, \xi_k(\omega))$.*

Világos, hogy amennyiben adva van egy $g(x_1, \dots, x_k)$ k változós függvény, és $\xi_1, \dots, \xi_k$ valószínűségi változók egy valószínűségi mezőn, akkor ezek egyértelműen meghatározzák a $\eta = g(\xi_1, \dots, \xi_k)$ valószínűségi változót. Megfordítva, ha adott egy a $(\xi_1, \dots, \xi_k)$ valószínűségi változók által generált σ-algebra, és egy erre a σ-algebrára mérhető η valószínűségi változó, akkor az előbb megfogalmazott tétel szerint ez felírható $\eta = g(\xi_1, \dots, \xi_k)$ alakban. Felmerül a kérdés, hogy az η valószínűségi változó meghatározza-e egyértelműen a g függvényt a fenti reprezentációban. Erre a kérdésre a válasz nyílván nemleges. Például, ha mindegyik ξ_j valószínűségi változó 0 és 1 közé esik, akkor a $g(x_1, \dots, x_k)$ tetszőleges olyan $(x_1, \dots, x_k)$ pontban megváltoztathatjuk, ahol legalább az egyik koordináta nagyobb mint 1 vagy kisebb mint nulla. De ez nem a teljes igazság. A g függvény ugyanis egyértelműen meg van határozva a következő gyengített értelemben. Jelölje $F(x_1, \dots, x_k)$ a $(x_1, \dots, x_k$ eloszlásfüggvényét, és jelölje μ_F az $F(x_1, \dots, x_k)$ eloszlásfüggvény által meghatározott Lebesgue–Stieltjes mértéket. Ha $g(x_1, \dots, x_k)$ és $\bar{g}(x_1, \dots, x_k)$ két olyan függvény, melyre $g(\xi_1, \dots, \xi_k) = \bar{g}(\xi_1, \dots, \xi_k)$, akkor a g és $\bar{g}$ függvény a μ_F mérték szerint majdnem mindenütt megegyezik. Viszont a feltételes valószínűség és feltételes várható érték csak egy valószínűséggel van meghatározva, ezért mint látni fogjuk a g függvénynek a fenti értelemben vett egyértelműségére van szükségünk.

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 integráltranszformációk tulajdonságainak ismeretében, hogy a $(**)$ reláció következik a ii$'$.) relációból.