A december 11-i gyakorlat témája
Rövid összefoglaló

Ezen a gyakorlaton is elsősorban azzal foglalkozunk, hogy hogyan lehet megérteni annak a jegyzetnek néhány eredményét, mely a vizsgan kérdezett anyagot tartalmazza. Előtte azonban beszéljük meg röviden a pótdolgozaton feladott első feladatot, mert erre a konvolució segítségével is megoldható feladatra senki nem adott jó megoldást. A feladat a következő:

Legyen ξ_1 és ξ_2 két független, a $\left[-\frac{1}{2}, \frac{1}{2}\right]$ intervallumban egyenletes eloszlású valószínűségi változó azaz legyen ξ_j sűrűségföggvénye $f(x) = 1$, ha $-\frac{1}{2} \le x \le \frac{1}{2}$, és $f(x) = 0$ egyébként, $j = 1, 2$. Számítsuk ki a $\xi_1 + \xi_2$ valószínűségi változó sűrűségfüggvényét.

Megoldás a konvolució segítségével: A $\xi_1+\xi_2$ valószínűségi változó sűrűségfüggvénye

$$f * f(x) = \int_{-\infty}^{\infty} f(y)f(x-y)\,dy = \int_{\left\{y\colon -\frac{1}{2} \le y \le \frac{1}{2}, -\frac{1}{2} \le x-y \le \frac{1}{2}\right\}} 1\,dy$$

$$= \int_{\max\left(-\frac{1}{2}, -\frac{1}{2}+x\right)}^{\min\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}+x\right)} dy,$$

ahonnan

$$f * f(x) = \begin{cases} \displaystyle\int_{-\frac{1}{2}+x}^{\frac{1}{2}} dx = 1 - x, & \text{ha } 0 \le x \le 1 \\ \displaystyle\int_{-\frac{1}{2}}^{\frac{1}{2}+x} dx = 1 + x, & \text{ha } -1 \le x \le 0 \\ 0, & \text{egyébként} \end{cases}$$

Másik megoldás: A (ξ_1, ξ_2) vektor sűrűségfüggvénye 1 a $B = \left[-\dfrac{1}{2}, \dfrac{1}{2}\right] \times \left[-\dfrac{1}{2}, \dfrac{1}{2}\right]$ négyzeten, ezért tetszőleges (a síkon értelmezett) A halmazra $P(\xi_1, \xi_2) \in A) = \lambda(A \cap B)$, ahol λ a Lebesgue mérték, azaz az $A \cap B$ halmaz területét tekintettük. Viszont a $\{\xi_1 + \xi_2 < x\}$ esemény megegyezik a $\{(\xi_1, \xi_2) \in A(x)\}$ eseménnyel, ahol $A(x) = \{(u, v)\colon u + v < x\}$. Innen $P(\xi_1 + \xi_2 < x) = \lambda(A_x \cap B)$, ahonnan $P(\xi_1 + \xi_2 < x) = \frac{1}{2}(1 + x)^2$, ha $-1 \le x \le 0$, és $P(\xi_1 + \xi_2 < x) = 1 - P(\xi_1 + \xi_2 \ge x) = \lambda(B \setminus A(x)) = 1 - \frac{1}{2}(1 - x)^2$, ha $0 \le x \le 1$. Továbbá $P(\xi_1 + \xi_2 < x) = 0$, ha $x < -1$ és $P(\xi_1 + \xi_2 < x) = 1$, ha $x > 1$. Deriválva a $\xi_1 + \xi_2$ fent kiszámítoott eloszlásfüggvényét megkapjuk a $\xi_1 + \xi_2$ sűrűségfüggvényét.

Megjegyzések a függetlenség és Kolmogorov 0–1 törvénye fejezethez.

Először értsék meg σ-algebráknak a 2.1 definicióban megadott függetlenségét. Itt szerepel események illetve valószínűségi változók függetlensége is. Én ez utóbbiakat tételnek nevezném. Nevezetesen arról van szó, hogy a korábban tanult σ-algebrák fogalma nélkül definiált függetlenség definiciója megegyezik az itt megadottal. E fejezet megértéséhez elengedhetetlen a farok σ-algebra 2.2 definiciója. A legérdekesebb eset, és gyakorlatilag az összes tekintett példa olyan, amelyikben bizonyos valószínűségi változók által generált σ-algebrát tekintünk. Ekkor a farok σ-algebra szemléletes tartalma az, hogy ez azon eseményekből áll, melyek be vagy be nem következtét meg tudjuk állapítani akkor is, ha ezen valószínűségi változók közül néhánynak az értékét nem ismerjük. Például az az esemény, hogy valószínűségi változók átlaga kisebb mint egy rögzített x szám nem függ véges sok az átlagban résztvevő valószínűségi változó értékétől. A fejezet fő eredménye a 2.6 Tétel, a Kolmogorov féle 0–1 törvény. Ez azt mondja ki, hogy független σ-algebrák, speciálisan független valószínűségi változók által generált σ-algebrák farok σ-algebrája olyan eseményekből áll, melyeknek valószínűsége nulla vagy egy. A bizonyítás eszméje a következő: Tekintünk egy rögzített A eseményt a farok σ-algebrában, és keresünk olyan (halmaz) algebrákat illetve σ-algebrákat, melyeknek elemei függetlenek ettől az A halmaztól. Belátjuk bizonyos analízis segítségével, hogy a farok σ-algebra eseményei ilyenek. Az állítás bizonyításában kulcsszerepet játszik mértékek kiterjeszteésének egyértelműsége algebráról az általa generált σ-algebrára. Ez azt jelenti speciálisan, hogy az A halmaz önmagától is független, azaz $P(A) = P(A)^2$, és innen következik a kívánt állítás. Feltétlenül javaslom, hogy nézzék át a fejezetben szereplő példákat is. A fejezetben szereplő Borel–Cantelli lemma megegyezik a bevezető valószínűségszámításban szereplő hasonló nevű eredménnyel.

Konvergenciafogalmak a valószínűségszámításban. A nagy számok erős és gyenge törvénye.

Először idézzük fel e fogalmak definicióit. Megjegyzem, hogy az egy valószínűségű konvergenciát majdnem mindenütt való a sztochasztikus konvergenciát pedig mértékben való konvergenciának is hívják.

Az egy valószínűségű konvergencia definiciója: *Valószínűségi változók ξ_n, $n = 1, 2, \dots$, sorozata akkor konvergál egy valószínűséggel egy ξ valószínűségi változóhoz, ha (egyrészt ezek a valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn vannak definiálva, másrészt)*

$$P\left(\omega\colon \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right) = 1.$$

A sztochasztikus konvergencia definiciója: *Valószínűségi változók ξ_n, $n = 1, 2, \dots$, sorozata akkor konvergál sztochasztikusan egy ξ valószínűségi változóhoz, ha (egyrészt ezek a valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn vannak definiálva, másrészt) minden $\varepsilon > 0$ számra*

$$\lim_{n\to\infty} P\left(|\xi_n(\omega) - \xi(\omega)| > \varepsilon\right) = 0.$$

Az eloszlásban való konvergencia definiciója: *Valószínűségi változók ξ_n, $n = 1, 2, \dots$, sorozata akkor konvergál eloszlásban egy $F(u)$ eloszlásfüggvényhez vagy az ezen* *eloszlásfüggvény által meghatározott eloszláshoz, ha* $\lim_{n\to\infty} P(\xi_n < u) = F(u)$ minden olyan u számra, ahol az $F(\cdot)$ eloszlásfüggvény függvény folytonos. *(Azt mondjuk, hogy a* ξ_n*,* $n = 1, 2, \dots$*, valószínűségi változók sorozata eloszlásban konvergál egy* ξ *valószínűségi változóhoz, ha ez a sorozat eloszlásban konvergál az* $F(u) = P(\xi < u)$ *eloszlásfüggvényhez.)*

Eloszlásfüggvények $F_n(u)$ *sorozata akkor és csak akkor konvergál egy* $F(u)$ *eloszlásfüggvényhez, ha valószínűségi változók olyan* ξ_n*,* $n = 1, 2, \dots$*, sorozata, melyekre* ξ_n *eloszlása* $F_n(u)$ *eloszlásban konvergálnak az* $F(u)$ *eloszlásfüggvényhez. Más szavakkal ez azt jelenti, hogy* $\lim_{n\to\infty} F_n(u) = F(u)$ *az* $F(\cdot)$ *függvény minden folytonossági pontjában.*

Igaz a következő kapcsolat: Egy valószínűségi konvergencia $\Rightarrow$ Sztochasztikus konvergencia $\Rightarrow$ Eloszlásban való konvergencia.

Megmutatjuk hogy, ha $\xi_n(\omega) \to \xi(\omega)$ egy valószínűséggel, akkor definiálva az $A_n = A_n(\varepsilon) = \left\{\omega\colon \sup_{k\ge n} |\xi_k(\omega) - \xi(\omega)| < \varepsilon\right\}$ halmazokat kapjuk, hogy az egymásba skatulyázott A_n halmazokra, (azaz $A_1(\omega) \subset A_2 \subset \cdots$), $P\left(\bigcup_{n=1}^{\infty} A_n\right) = 1$. Ezért $\lim_{n\to\infty} P(A_n) = 1$. Mivel $\{|\xi_n(\omega) - \xi(\omega)| < \varepsilon\} \supset A_n$, $P(|\xi_n(\omega) - \xi(\omega)| < \varepsilon) \to 1$, azaz $P(|\xi_n(\omega) - \xi(\omega)| \ge \varepsilon) \to 0$, ha $n \to \infty$. Ez azt jelenti, hogy az egy valószínűségű konvergenciából következik a sztochasztikus konvergencia. A jegyzet tartalmazza a következő eredményt is (Lemma 5.7 (i):

Valószínűségi változók ξ_n, $n = 1, 2, \dots$, sorozata, akkor és csak akkor konvergál egy valószínűséggel egy ξ valószínűségi változóhoz, ha az $\eta_n = \sup_{k\ge n} |\xi_k - \xi|$ valószínűségi változók sorozata sztochasztikusan konvergál nullához, azaz minden $\varepsilon > 0$ számra $P\left(\lim_{n\to\infty} \sup_{k\ge n} |\xi_k - \xi| > \varepsilon\right) = 0$.

Lássunk példát arra, hogy lehetséges olyan ξ_n, $n = 1, 2, \dots$, és ξ valószínűségi változókat konstruálni, melyekre a ξ_n, $n = 1, 2, \dots$, sorozat sztochasztikusan tart ξ-hez, de a ξ_n sorozat nem konvergál egy valószínűséggel a ξ valószínűséggel a ξ valószínűségi változóhoz.

Tekintsük a következő $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt: Ω a $[0, 1]$ intervallum, $\mathcal{A}$ a Borel mérhető halmazok σ-algebrája a $[0.1]$ intervallumon, a P valószínűségi mérték a Lebesgue mérték. Legyen

$$\xi_n(x) = \begin{cases} 1 & \text{ha } x \in \left[(n-2^k)2^{-k}, (n+1-2^k)2^{-k}\right] \\ 1 & \text{ha } x \in \left[(n-2^k)2^{-k}, (n+1-2^k)2^{-k}\right] \end{cases}$$
$$\text{akkor ha } 2^k \le n < 2^{k+1}, \quad k = 1, 2, \dots,$$

és $\xi(x) = 0$ minden $0 \le x \le 1$ számra. Ekkor $P(|\xi_n - \xi| > \varepsilon) = 2^{-k}$ minden $\varepsilon > 0$ számra, ha $2^k \le n < 2^{k+1}$. Tehát a ξ_n, $n = 1, 2, \dots$, sorozat sztochasztikusan konvergál a ξ valószínűségi változóhoz. Viszont mivel $\limsup_{n\to\infty} \xi_n(x) = 0$ minden $0 \le x \le 1$ számra, ezért a ξ_n sorozat nem konvergál egy valószínűséggel a ξ valószínűségi változóhoz.

Azt mondjuk, hogy $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók teljesítik a nagy számok gyenge törvényét, ha ezek $S_n = \xi_1 + \cdot + \xi_n$ összegeire $\frac{S_n}{n} \to E\xi_1$ sztochasztikusan, teljesítik a nagy számok erős törvényét, ha $\frac{S_n}{n} \to E\xi_1$ egy valószínűséggel, ha $n \to \infty$. Valójában, olyan becsléseket bizonyítanak, melyek független (esetleg csak korrelálatlan), de nem feltétlenül egyforma eloszlású valószínűségi változók normalizált összegéről szólnak.

Az 5.1 illetve 5.2 Tétel tartalmazza a nagy számok gyenge törvényének a bizonyítását. A bizonyítás lényege a Csebisev egyenlőtlenség.

Az itt szereplő eredményekben fontos szerepet játszik az, hogy milyen momentum feltételek teljesülnek az egyes valószínűségi változókra. Az, hogy egy valószínűségi változónak léteznek a magas momentumai azt jelenti, hogy viszonylag kis valószínűséggel vesz fel nagyon nagy értékeket, ezért nem kell attól tartani, hogy néhány tag extrém viselkedése rontja el a nagy számok erős törvényének az érvényességét. Így az 5.5 Tételben négy momentum létezése van feltéve. A nagy számok törvényének az igazán éles és nehezebben bizonyítható alakja a nagy számok Kolmogorov féle erős törvénye, 5.11 tétel. Eszerint az eredmény szerint a nagy számok erős törvényének szükséges és elégséges feltétele az, hogy a valószínűségi változók abszolut értékének legyen véges várható értéke. A bizonyításban szereplő elégségesség bizonyításának alaplépése a Kolmogorov egyenlőtlenség. Ez az egyenlőtlenség ugyanazt a becslést adja arra, hogy független egyforma eloszlású valószínűségi változók részletösszegeinek maximuma nagyobb mint egy x szám, mint amit a Csebisev egyenlőtlenség ad arra, hogy a maximumban résztvevő összegek közül az utolsó nagyobb mint ez az x szám. Egy másik fontos technikai lépés a bizonyításban a (nem valószínűségszámítás tartalmú) Kronecker Lemma ((5.9) Lemma). Ennek bizonyításában az úgynevezett Abel féle átrendezés játszik fontos szerepet, melyet a jegyzetben "parciális szummázás"-nak neveznek.

Érdemes röviden elmagyarázni azt, hogy a nagy számok erős törvényének miért szükséges feltétele az, hogy az összeadandók abszolut értékének legyen véges várható értéke. Azaz tegyük fel, hogy $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók, $S_n = \xi_1 + \cdots + \xi_n$. Azt állítjuk, hogy amennyiben $\lim_{n\to\infty} \frac{S_n}{n} \to c$ valamilyen konstanshoz (vagy valamilyen valószínűségi változóhoz), akkor $E|\xi_1| < \infty$.

Be lehet látni, hogy $E|\xi| < \infty$ akkor és csak akkor, ha $\sum_{n=0}^{\infty} P(|\xi| > n) < \infty$. Jelölje ugyanis $F(\cdot)$ a ξ valószínűségi változó eloszlásfüggvényét. Ezzel a jelöléssel $E|\xi| < \infty$ akkor és csak akkor, ha $\int |u|\, dF(u) < \infty$. Továbbá ez az integrál akkor és csak akkor véges, ha $\sum_{n=0}^{\infty} nP(n < |\xi| \le n+1) < \infty$. (Miért?) Ezt az összeget átrendezve

$$\sum_{n=0}^{\infty} nP(n < |\xi| \le n+1)) = \sum_{n=0}^{\infty} n\left(P(|\xi| > n) - P(|\xi| > n+1)\right)$$

$$= \sum_{n=0}^{\infty} P(|\xi| > n)[(n+1) - n] = \sum_{n=0}^{\infty} P(|\xi| > n)$$

Innen következik az állítás.

Belátjuk az előző eredmény és a Borel–Cantelli lemma segítségével, hogy ha a $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók sorozata olyan, hogy $E|\xi_1| = \infty$, akkor az $\frac{S_n(\omega)}{n}$ sorozat egy valószínűséggel divergens, ahol $S_n = \sum\limits_{k=1}^{\infty} \xi_k$. Ez azt jelenti, hogy ebben az esetben nem teljesül a nagy számok erős törvénye.

Valóban, ha $E|\xi_1| = \infty$, akkor $\sum\limits_{n=0}^{\infty} P(|\xi_n| > n) = \sum\limits_{n=0}^{\infty} P(|\xi_1| > n) = \infty$. Továbbá az $\{\omega\colon |\xi_n(\omega)| > n\}$ események függetlenek, ezért a Borel-Cantelli lemma szerint majdnem minden $\omega \in \Omega$ esetén $|\xi_n(\omega)| > n$ végtelen sok n indexre.

Ha $\omega \in \Omega$ olyan pont, melyben az $\frac{S_n(\omega)}{n}$, $n = 1, 2, \dots,$ sorozat konvergens, akkor ebben az ω pontban $\sup\limits_{1 \le n < \infty} \left| \frac{S_n(\omega)}{n} \right| < K(\omega)$, és $\lim\limits_{n \to \infty} \left(\frac{S_{n+1}(\omega)}{n+1} - \frac{S_n(\omega)}{n} \right) = 0$. Viszont, $\frac{S_{n+1}(\omega)}{n+1} - \frac{S_n(\omega)}{n} = \frac{\xi_{n+1}(\omega)}{n+1} + \frac{S_n(\omega)}{n(n+1)}$, és $\frac{S_n(\omega)}{n(n+1)} \to 0$, $|\xi_{n+1}(\omega)| < n+1$ minden $n = 1, 2, \dots$ számra, ha az $\frac{S_n(\omega)}{n}$ sorozat konvergál. Ez viszont csak nulla valószínűségű $\omega \in \Omega$ halmazra érvényes. Megfogalmazom bizonyítás nélkül, hogy érvényes ennek az állításnak a megfordítása is, nevezetesen a következő tétel, amelyik szerepel a jegyzetben.

Tétel. *Legyen $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók sorozata, és definiáljuk az $S_n = \sum\limits_{k=1}^{n} \xi_k$, $n = 1, 2, \dots,$ részletösszegeket. Az $\frac{S_n(\omega)}{n}$, $n = 1, 2, \dots,$ sorozat akkor és csak akkor konvergál pozitív valószínűséggel, ha $E|\xi_1| < \infty$. Ha $E|\xi_1| < \infty$, akkor ez a sorozat teljesíti a nagy számok erős törvényét, azaz*

$$\lim_{n \to \infty} \frac{S_n(\omega)}{n} = E\xi_1 \quad \textit{majdnem minden } \omega \in \Omega\textit{-ra.}$$

Az iterált logaritmus tétel.

Ez az eredmény felfogható úgy mint a nagy számok erős törvényének az élesítése. Ez pontosabban kifejezi azt, hogy független, egyforma eloszlású valószínűségi változók összegeinek a $\limsup$-ja mekkora. Ez akkor érvényes, ha a valószínűségi változóknak létezik második momentuma. Az iterát logaritmustétel a következőt mondja ki:

Iterált logaritmus tétel. *Legyenek $\xi_1(\omega), \xi_2(\omega), \dots,$ független, egyforma eloszlású valószínűségi változók valamilyen $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyekre $E\xi_1(\omega) = 0$, $\operatorname{Var} \xi_1(\omega) = \sigma^2 < \infty$. Ekkor*

$$\limsup_{n \to \infty} \frac{\xi_1(\omega) + \dots + \xi_n(\omega)}{\sqrt{2n\sigma^2 \log\log n}} = 1, \quad \textit{majdnem minden } \omega \in \Omega \textit{ elemi eseményre,}$$

és

$$\liminf_{n\to\infty} \frac{\xi_1(\omega)+\cdots+\xi_n(\omega)}{\sqrt{2n\sigma^2\log\log n}} = -1, \quad \textit{majdnem minden } \omega\in\Omega \textit{ elemi eseményre,}$$

sőt igaz a következő némileg élesebb állítás is: A $\dfrac{\xi_1(\omega)+\cdots+\xi_n(\omega)}{\sqrt{2n\sigma^2\log\log n}}$, $n=1,2,\dots,$ *sorozat torlodási pontjainak halmaza a* $[-1,1]$ *intervallum egy valószínűséggel.*

Nem nehéz belátni a Csebisev egyenlőtlenség vagy a centrális határeloszlástétel segítségével, hogy tetszőleges $\alpha(n)\to\infty$, ha $n\to\infty$ esetén

$$\lim_{n\to\infty} P\left(\frac{|\xi_1(\omega)+\cdots+\xi_n(\omega)|}{\sqrt{n}\alpha(n)}<\varepsilon\right)=1 \quad \text{minden } \varepsilon>0 \text{ számra,}$$

azaz a $\dfrac{\xi_1(\omega)+\cdots+\xi_n(\omega)}{\sqrt{n}\alpha(n)}$ valószínűségi változók sztochaszikusan konvergálnak nullához.

A jegyzetben az állítás csak normális valószínűségi változók összegére van bizonyítva. Az bizonyos fokig standard, de meglehetősen fárasztó lépéseken alapul. Egyrészt a normális eloszlásfüggvény végtelenben való viselkedését kell leírni, (Lemma (6.2) ezután részben a Borel–Cantelli lemmát alkalmzzuk illetve a

$$P\left(\frac{\xi_1(\omega)+\cdots+\xi_n(\omega)}{\sqrt{2n\sigma^2\log\log n}}>x\right)$$

és

$$P\left(\sup_{1\le k\le n}\frac{\xi_1(\omega)+\cdots+\xi_k(\omega)}{\sqrt{2n\sigma^2\log\log n}}>x\right)$$

alakú eseményekre kell jó becslést adni. Ez utóbbi célból szerepel a bizonyításban a tükrözési elv (Lemma 6.3).

Megjegyzések a gyenge konvergencia című fejezethez.

A legfontosabb azt megérteni, hogy a gyenge konvergencia fogalma a számegyenesen már tárgyalt eloszlásban való konvergencia általánosítása (szeparábilis) metrikus tereken értelmezett valószínűségi mértékekre. Azaz, arról van szó, hogy amennyiben olyan valószínűségi változók (mérhető függvények) sorozatát vesszük, melyek értéküket nem feltétlenül a számegyenesen, hanem esetleg egy általános téren veszik fel, akkor mikor mondhatjuk, hogy ezen valószínűségi változók eloszlásban (azaz gyengén) konvergálnak egy ugyanazon a téren értelmezett valószínűségi változóhoz. Az előbb mondottak speciálisan azt jelentik, hogy valós értékű valószínűségi változók eloszlásban való, illetve gyenge konvergenciája megegyezik.

Megbeszéltük, hogy a számegyenesen is az eloszlásban való konvergencia esetén az eloszlásfüggvények konvergenciáját csak a határeloszlás folytonossági pontjaiban követeljük meg. Nem meglepő, hogy ennek a megszorításnak valamilyen módon öröklődni kell általános terekben is. Ráadásul általános esetben nem beszélhetünk eloszlásfüggvényekről. A fejezet fő eredménye a 7.2 Lemma, amelyik ekvivalens megfogalmazásait adja a gyenge konvergenciának. Ezek közül az (i) jellemzés a számegyenesen megegyezik az eloszlásban való konvergencia általunk megadott jellemzésével. Annak érdekében, hogy megértsük a (iii) és (iv) jellemzéseket, azt hogy miért van $\limsup$ a lim helyett zárt halmazok, illetve $\liminf$ a lim helyett nyílt halmazok esetén, tekintsük a következő egyszerű példát. Tekintsük a számegyenest, és legyen rajta x_n, $n = 1, 2, \dots$, olyan számsorozat, melyre $x_n \to x_0$ valamilyen x_0 számra Legyen továbbá az x_n sorozat olyan, hogy $x_n \neq x_0$, ha $n \geq 1$. Definiáljuk azon μ_n, $n = 0, 1, 2, \dots$, valószínűségi mértékeket, melyek az x_n pontba vannak koncentrálva, $n = 0, 1, 2, \dots$, azaz legyen $\mu_n(A) = 1$, ha $x_n \in A$, és $\mu_n(A) = 0$, ha $x_n \notin A$. Ekkor mint azt természetes várni, a μ_n mértékek gyengén konvergálnak a μ_0 mértékhez. Tekintsük az $F = \{x_0\}$ egyelemű halmazt. Ekkor F zárt halmaz, $\mu_n(F) = 0$, ha $n \geq 1$, és $\mu_0(F) = 1$. Ez azt jelenti, hogy $\limsup \mu_n(F) \leq \mu_0(F)$ erre a zárt halmazra, ami konzisztens a (iii) tulajdonsággal, de nem írhatunk limeszt és egyenlőséget a $\limsup$ és $\leq$ helyett. Hasonlóan, ha a $G = R^1 \setminus \{x\}$ halmazt tekintjük, azaz az F halmaz komplementerét, akkor a G halmaz nyílt, $\mu_n(G) = 1$, ha $n \geq 1$, és $\mu_0(G) = 0$. Ez azt jelenti, hogy a (iv) tulajdonság teljesül ebben az esetben, de nem írhatunk egyenlőséget és limeszt ebben a formulában sem.

A feltételes valószínűségről és feltételes várható értékről

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 óra. A második probléma 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 változókat 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

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

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

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

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 egy fontos eredményéből, a Banach–Hahn 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 $f_A(x_1,\dots,x_k)=\bar f_A(x_1,\dots,x_k)$ az F eloszlás szerint meghatározott mérték szerint majdnem minden $(x_1,\dots,x_k)$ pontra.

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

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

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 Banach–Hahn tétel segítségével láthatjuk.

Mielőtt megtárgyalnánk a fent említett Banach–Hahn 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 ezek az események egy σ-algebrát alkotnak. Ezért az általános kérdés úgy fogalmazható meg, hogy adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn $\mathcal{F} \subset \mathcal{A}$ σ-algebra és egy η valószínűségi változó, melyre $E|\eta| < \infty$, illetve A esemény, akkor definiáljuk a $P(A|\mathcal{F})(\omega)$ feltételes valószínűségi változót, 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.

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

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

A feltételes valószínűség és feltételes várható érték definiciója. *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi térnek 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 Banach–Hahn 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 Banach–Hahn tételt megfogalmazhassuk előbb be kell vezetni a következő definiciót.

Definició: Egy mérték abszolut folytonossága egy másik 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.*

Banach–Hahn 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 $\int_C f(\omega)\,d\mu(\omega) = \nu(C)$ 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.*

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

A Banach–Hahn 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 Banach–Hahn 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 Banach–Hahn 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 Banach–Hahn 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 Banach–Hahn 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 Banach–Hahn 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. A fent kimondott tétel tulajdonképpen a jegyzetben szereplő 1.13 Állítás némileg általánosabb formában.)

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. Ezek és néhány egyéb állítás szerepel a 9.5 tételben. Rögzítsünk egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt. Az alábbi tulajdonságokban szereplő valószínűségi változók ezen a valószínűségi mezőn vannak értelmezve.

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

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

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

A hatodik 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).

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

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

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

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

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