A Valószínűségszámítás II. előadássorozat tízennegyedik előadása.
2002. december 17.

Feltételes valószínűség és várható érték *(második rész)*

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

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

Tétel a feltételes várható érték tulajdonságairól.

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

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

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

A Tétel bizonyítása. Az 1. tulajdonság bizonyításához azt kell belátni, hogy minden

$F \in \mathcal{F}$ halmazra

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

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

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

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

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

mert a feltételes várható érték definiciója alapján

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

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

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

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

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

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

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

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

A 4. tulajdonság bizonyítása triviális, ezért azt elhagyom.

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

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

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

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

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

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

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

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

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

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

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

1. megjegyzés: A fenti tétel azt fejezi ki, hogy ha a ξ valószínűségi változót az $\mathcal{F}$ σ-algebra eseményeitől függő valószínűségi változóval, azaz $\mathcal{F}$ mérhető valószínűségi változóval akarjuk becsülni, és a becslés jóságát úgy mérjük, hogy véve a becsült ξ valószínűségi változó és a becslésre használt valószínűségi változó különbségének a négyzetét, e kifejezés várható értéke legyen minél kisebb, akkor az $E(\xi|\mathcal{F})$ feltételes várható érték az optimális becslés. Jegyezzük meg, hogy a bevezető valószínűségszámítás előadásban szerepelt egy eredmény, amelyik a fenti állítás speciális esetének tekinthető. Nevezetesen beláttuk, hogy $\operatorname{Var}\xi = E(\xi - E\xi)^2 \le E(\xi - a)^2$ tetszőleges a valós számra. Ha definiáljuk azt a triviális $\mathcal{F}_0 = \{\emptyset, \Omega\}$ σ-algebrát, amelyik csak az üres halmazból és a biztos eseményből áll, akkor erre a σ-algebrára csak a konstans függvények mérhetők, és erre a $\mathcal{F}_0$ σ-algebrára nézve $E(\xi|\mathcal{F}_0) = E\xi$. Ezért az előbb megfogalmazott tétel az ebben a megjegyzésben megfogalmazott állítást jelenti ebben a speciális esetben.

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

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

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

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

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

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

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

mert $E(\eta - E\xi|\mathcal{F}))(E(\xi|\mathcal{F}) - \xi)) = E\eta(E(\xi|\mathcal{F}) - \xi)) - EE\xi|\mathcal{F})(E(\xi|\mathcal{F}) - \xi)) = 0$, és ezt kellett belátni.

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

Feladat:

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

Megoldás: Láttuk, (lásd a 9. előadáson szereplő 6. következményt a normális eloszlású véletlen vektorokról, hogy $\xi=a\eta+\zeta$ alakban írható, ahol az a konstans alkalmas választásával (nevezetesen az $a=\dfrac{\operatorname{Cov}(\xi,\eta)}{\operatorname{Var}\eta}$ választással) elérhető, hogy a $\zeta=\xi-a\eta$ és η valószínűségi változók függetlenek legyenek. Ezzel az a választással viszont

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

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

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

Amikor valószínűségi változók függvényeinek (nem feltételes) várható értékét vizsgáltuk, nagy segítséget jelentett az, hogy e várható értékeket ki tudtuk fejezni e valószínűségi változók eloszlásfüggvényei, pontosabban ezen eloszlásfüggvények által meghatározott Lebesgue–Stieltjes mérték szerinti (Lebesgue) integrál segítségével. Felmerül a kérdés, nem lehet-e megfogalmazni és bebizonyítani ezen eredménynek valamely természetes megfelelőjét a feltételes várható értékre. Erre a kérdésre pozitív választ lehet adni, de a vizsgálatban bizonyos nehézségek lépnek fel. E nehézségek azzal kapcsolatosak, hogy a Lebesgue-féle integrált csak (σ-additív) mértékek szerint tudjuk definiálni, és felmerül a kérdés, tudjuk-e definiálni ezeket a σ-additív feltételes eloszlásokat. Be lehet látni, hogy ez lehetséges, de a bizonyítás az általános esetben mély gondolatokat igényel, és amellett a kapott eredmény egzisztencia tétel, tehát nem ad lehetőséget arra, hogy a kívánt feltételes várható értékeket effektíve kiszámoljuk. Azt a tételt, mely az ilyen vizsgálatokban alapvető a Kiegészítésben ismertetem bizonyítás nélkül.

Viszont van egy olyan speciális eset, amelyikben mindent explicit módon ki tudunk számolni. Ráadásul ez a speciális eset tartalmazza a klasszikus statisztikai vizsgálatokban legfontosabb eseteket.

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

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

Annak igazolásához, hogy a fenti feltételes valószínűségeket az adott módon lehet kiszámolni 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

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

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

$$h(\xi_1, \dots, \xi_k, \xi_{k+1}, \dots, \xi_{k+l}) | \xi_{k+1} = x_{k+1}, \dots, \xi_{k+l} = x_{k+l})$$

feltételes várható értéket ki lehet számolni a következő módon:

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

Ennek a formulának a bizonyítását, mely úgy történhet, hogy azt előbb a legegyszerűbb h függvényekre, bizonyítjuk, olyan h függvényekre, melyek

$$h(x_1, \dots, x_k, x_{k+1}, \dots, x_{k+l}) = I_A(x_1, \dots, x_k) I_B(x_{k+1}, \dots, x_{k+l})$$

alakban írhatóak, ahol $I_A(\cdot)$ és $I_B(\cdot)$ egy k illetve l dimenziós mérhető halmaz indikátor függvénye. Ezután bizonyos mértékelméleti eredmények felhasználásával bizonyítható az állítás általános függvényekre alkalmas határátmenet segítségével. Ennek részleteit elhagyom. A fent tárgyalt formulák segítségével be lehet bizonyítani néhány érdekes formulát a feltételes eloszlásokra, mint például a következő két feladat megoldását.

1. feladat:

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

2. feladat:

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

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

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

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

$$\int_A EG(\zeta,y)H(\,dy) = EG(\zeta,\eta)I_A(\eta) = \int I_A(y)\int G(u,y)F(\,du)H(\,dy),$$

ahol $I_A(\cdot)$ az A halmaz indikátorfüggvényét jelöli. Ez az azonosság viszont igaz, mert alkalmazva a Fubini tételt a bizonyítandó azonosság jobboldalán szereplő integrálra, és először az x változó szerint integrálva a bizonyítandó azonosság azt kapjuk, hogy a belső integrál $EG(\zeta,y)$-nal egyenlő. Innen látható, hogy a jobboldalon szereplő integrál egyenlő a baloldalon szereplő kifejezéssel.

A 2. feladat bizonyítása. Alkalmazzuk 9. előadáson szerepelt 6. következményt, és tekintsük a ξ koordinátájának $\xi = a\eta+\zeta$ előállítását $a = \dfrac{\mathrm{Cov}\,(\xi,\eta)}{\mathrm{Var}\,\eta}$ választással. Ekkor a $\zeta = \xi - a\eta$ és az η valószínűségi változók függetlenek. Ezért $P(\xi < x|\eta = y) = P(a\eta+\zeta < x|\eta = y) = P(ay+\zeta < x)$ az 1. feladat eredménye alapján. (Ez az 1. feladat eredményéből következik, ha azt a következő $G(u,v) = G_x(u,v)$ függvényre alkalmazzuk: $G(u,v) = 1$, ha $u+av < x$, és $G(u,v) = 0$, ha $u+av \ge 0$, ahol $a = \dfrac{\mathrm{Cov}\,(\xi,\eta)}{\mathrm{Var}\,\eta}$.) Viszont ζ normális eloszlású valószínűségi változó $E\xi - aE\eta = E\xi - \dfrac{\mathrm{Cov}\,(\xi,\eta)E\eta}{\mathrm{Var}\,\eta}$ várható értékkel és $\mathrm{Var}\,\xi+a^2\mathrm{Var}\,\eta-2a\mathrm{Cov}\,(\xi,\eta) = \mathrm{Var}\,\xi-\dfrac{\mathrm{Cov}\,(\xi,\eta)^2}{\mathrm{Var}\,\xi}$ szórásnégyzettel. Másrészt $Eay+\zeta = ay+E\zeta$, és $\mathrm{Var}\,(ay+\zeta) = E\zeta$, ahonnan következik a feladat b) része is.

Kiegészítés. A következő tétel bizonyítását elhagyom. Ez nehezebb, mint a korábbi eredmények vizsgálata, viszont 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.*

E tétel bizonyítása nem triviális. Itt nem elég 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:*

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