A Valószínűségszámítás I. előadássorozat hetedik témája.

Valószínűségi változók eloszlása, várható értéke és függetlensége.

Megbeszéltük, hogy egy ξ (valós értékű) valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn értelmezett (mérhető) függvény. De konkrét problémákban általában nem magukat a valószínűségi változókat adjuk meg (amihez előbb be kell vezetni azt a valószínűségi mezőt, ahol azok definiálva vannak), hanem azok alább definiált eloszlásfüggvényét, és kérdéseinket ezen eloszlásfüggvény segítségével fogalmazzuk meg.

Valószínűségi változó eloszlásfüggvényének a definiciója. *Legyen adva egy $\xi(\omega)$ (valós értékű) valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ennek $F(x)$ eloszlásfüggvényén az $F(x) = P(\{\omega\colon\ \xi(\omega) < x\})$, $-\infty < x < \infty$, függvényt értjük.*

E fogalom jobb megértésének érdekében tekintsük a következő példákat.

Feladatok:

1.) Tekintsük egy szabályos dobókocka feldobását, illetve azt a ξ valószínűségi változót, amely megadja, hogy mi a dobás eredménye. Mi a ξ valószínűségi változó $F(x)$ eloszlásfüggvénye?

Megoldás: A ξ valószínűségi változó az 1, 2, 3, 4, 5 és 6 értékek valamelyikét veszi fel, mindegyiket $\frac{1}{6}$ valószínűséggel. Ezért annak a valószínűsége, hogy $\xi(\omega) < x$ nullával egyenlő, ha $x < 1$. Sőt $x = 1$ esetében is teljesül a $P(\xi < 1) = 0$ azonosság, mert annak valószínűségét nézzük, hogy a ξ valószínűségi változó szigorúan kisebb mint az $x = 1$ szám. Ezért $F(x) = 0$, ha $-\infty < x \le 1$. Ha $1 < x < 2$, akkor az $\{\omega\colon\ \xi(\omega) < x\}$ esemény azt jelenti, hogy a dobás eredménye 1. Ez az állítás igaz $x = 2$ esetében is. Ezért $F(x) = \frac{1}{6}$, ha $1 < x \le 2$. Ha $2 < x \le 3$, akkor az $\{\omega\colon\ \xi(\omega) < x\}$ esemény azt jelenti, hogy a dobás eredménye 1 vagy 2. Ezért $F(x) = \frac{2}{6}$, ha $2 < x \le 3$. Ezt a gondolatot végigkövetve kapjuk, hogy $F(x) = 0$, ha $-\infty < x \le 1$, $F(x) = \frac{j}{6}$, ha $j < x \le j+1$, $1 \le j \le 5$, és $F(x) = 1$, ha $6 < x < \infty$,

2.) Feldobunk egy pénzdarabot, amely p valószínűséggel esik a fej, $1-p$ valószínűséggel az írás oldalra kétszer egymástól függetlenül. Jelölje ξ azt a valószínűségi változót, amely a fejdobások számát adja meg. Adjuk meg a ξ valószínűségi változó $F(x)$ eloszlásfüggvényét.

Megoldás: A ξ valószínűségi változó eloszlását a $P(\xi = 0) = (1-p)^2$, $P(\xi = 1) = 2p(1-p)$ és $P(\xi = 2) = p^2$ képletek adják meg. Ez azt jelenti, hogy a ξ valószínűségi változó $F(x) = P(\xi < x)$ eloszlásfüggvényére $F(x) = 0$, ha $x \le 0$, $F(x) = (1-p)^2$, ha $0 < x \le 1$, $F(x) = P(\xi = 0) + P(\xi = 1) = (1-p)^2 + 2p(1-p) = 1 - p^2$, ha $1 < x \le 2$, és $F(x) = 1$, ha $x > 2$.

3.) Ledobunk egy pontot véletlenül az egységintervallumra, azaz annak a valószínűsége, hogy a ledobott pont valamely $[a, b] \subset [0, 1]$ intervallumba esik legyen $b - a$. Jelölje ξ a ledobott pont helyének a nagyságát. Adjuk meg a ξ valószínűségi változó eloszlásfüggvényét.

Megoldás: Mivel $P(\xi < x) = 0$, ha $x \le 0$, $P(\xi < x) = P(\xi \in [0, x)) = x$, ha $0 \le x \le 1$, és $P(\xi < x) = 1$, ha $x > 1$, ezért az $F(x) = P(\xi < x)$ eloszlásfüggvényre $F(x) = 0$, ha $x < 0$, $F(x) = x$, ha $0 \le x \le 1$, és $F(x) = 1$, ha $x > 1$.

4.) Ledobunk egy pontot véletlenül a $[0,2]$ intervallumra, azaz annak a valószínűsége, hogy a ledobott pont valamely $[a,b]\subset[0,2]$ intervallumba esik legyen $\frac{b-a}{2}$. Valaki felírja a ledobott pont helyét egy jegyzőkönyvbe, ha a ledobott pont a $[0,1]$ intervallumba esik, és a 0 értéket írja be, ha ez a pont az $[1,2]$ intervallumba esik. Jelölje η a jegyzőkönyvbe írt szám értékét. Adjuk meg az η valószínűségi változó eloszlásfüggvényét.

Megoldás: Jelölje ξ a ledobott pont helyének az értékét. Ekkor $P(\eta<0)=0$, $P(\eta=0)=P(\xi\in[1,2])=\frac{1}{2}$, $P(\eta<x)=P(\xi<x)+P(\xi\in[1,2])=\frac{1}{2}+\frac{x}{2}$, ha $0\le x\le 1$, és $P(\eta<x)=1$, ha $x>1$. Innen az η valószínűségi változó $F(x)$ eloszlásfüggvénye a következő: $F(x)=0$, ha $x\le 0$, $F(x)=\frac{1}{2}+\frac{x}{2}$, ha $0\le x\le 1$, és $F(x)=1$, ha $x\ge 1$.

Az első két példában egy diszkrét eloszlású valószínűségi változó eloszlásfüggvényét adtuk meg. A harmadik példában szereplő eloszlásfüggvény folytonos függvény volt. Az ilyen eloszlásokat folytonos eloszlásnak hívják az irodalomban. A negyedik példában szereplő eloszlásfüggvény nem volt folytonos, mert a nulla pontban ugrása van, és ugyancsak nem volt egy diszkrét eloszlású valószínűségi változó eloszlásfüggvénye.

Miért fontos az eloszlásfüggvény fogalma? Általában nem tudjuk, hogy milyen véletlen hatások eredményeként jelenik meg egy $\xi(\omega)$ valószínűségi változó értéke, csak azt, hogy milyen valószínűséggel vesz fel ez a valószínűségi változó bizonyos értékeket. Ezért természetes, hogy csak ezeket a valószínűségeket adjuk meg. Viszont a $\xi(\omega)$ valószínűségi változó $F(x)$ eloszlásfüggvénye csak az $\{\omega\colon\ \xi(\omega)<x\}$ alakú események valószínűségét adja meg. A következő kérdés az, hogy nem jelent-e ez megszorítást, hiszen minket az összes $\{\omega\colon\ \xi(\omega)\in B\}$ alakú esemény valószínűsége érdekel, ahol B „szép" halmaz. De bizonyos mértékelméleti eredményekből következik, hogy az $F(x)$ eloszlásfüggvény meghatározza az összes ilyen esemény valószínűségét. Az, hogy egy halmaz „szép" pontosabban megfogalmazva azt jelenti, hogy ez a halmaz Borel mérhető. Lássuk, hogyan lehet néhány ilyen esemény valószínűségét meghatározni az eloszlásfüggvény segítségével.

Mivel $\{\omega\colon\ a\le\xi(\omega)<b\}=\{\omega\colon\ \xi(\omega)<b\}\setminus\{\omega\colon\ \xi(\omega)<a\}$, ezért

$$P\left(\{\omega\colon\ a\le\xi(\omega)<b\}\right)=P\left(\{\omega\colon\ a\le\xi(\omega)<b\}\right)-P\left(\{\omega\colon\ \xi(\omega)<a\}\right)=F(b)-F(a).$$

Mivel $\{\omega\colon\ a\le\xi(\omega)\le b\}=\bigcap\limits_{n=1}^{\infty}\left\{\omega\colon\ a\le\xi(\omega)<b+\frac{1}{n}\right\}$, és a valószínűség σ-additivitásából következnek annak folytonossági tulajdonságai, ezért

$$\begin{aligned}P\left(\{\omega\colon\ a\le\xi(\omega)\le b\}\right)&=\lim_{n\to\infty}P\left(\left\{\omega\colon\ a\le\xi(\omega)<b+\frac{1}{n}\right\}\right)\\&=\lim_{n\to\infty}\left[F\left(b+\frac{1}{n}\right)-F(a)\right].\end{aligned}$$

Hasonlóan, ha adva van diszjunkt zárt $[a_k,b_k]$, $1\le k\le n$, intervallumok halmaza, akkor

$$P\left(\bigcup_{k=1}^{n}\{\omega\colon\ a_k\le\xi(\omega)\le b_k\}\right)=\sum_{k=1}^{n}\lim_{N\to\infty}\left[F\left(b_k+\frac{1}{N}\right)-F(a_k)\right].$$

Feladat:

Legyen adva egy $\xi(\omega)$ valószínűségi változó $F(x)$ eloszlásfüggvénye. Határozzuk meg ennek segítségével az $\{\omega\colon\, a < \xi(\omega) < b\}$ alakú események, $-\infty < a < b < \infty$, valószínűségét, valamint annak valószínűségét, hogy $\xi(\omega)$ valamilyen páros egész értéket vesz fel.
Megoldás:

$$\begin{aligned} P\left(\{\omega\colon\, a<\xi(\omega)<b\}\right) &= \lim_{n\to\infty} P\left(\left\{\omega\colon\, a+\frac{1}{n} \le \xi(\omega) < b\right\}\right) \\ &= \lim_{n\to\infty}\left[F(b) - F\left(a+\frac{1}{n}\right)\right]. \end{aligned}$$

Egy tetszőleges u pontra $P(\xi = u) = \lim\limits_{n\to\infty}\left[F\left(u+\frac{1}{n}\right) - F(u)\right]$. Ezért annak valószínűsége, hogy a ξ valószínűségi változó valamely páros egész értéket vesz fel $\sum\limits_{k=-\infty}^{\infty} \lim\limits_{n\to\infty}\left[F\left(2k+\frac{1}{n}\right) - F(2k)\right]$.

Megfogalmazom azokat eredményeket, amelyek segítségével jellemezni tudjuk az eloszlásfüggvényeket, illetve, amelyek kimondják, hogy az eloszásfüggvények meghatározzák a minket érdeklő valószínűségeket.

Lemma eloszlásfüggvények tulajdonságairól. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, és legyen $F(x) = P(\{\omega\colon\, \xi(\omega) < x\})$, $-\infty < x < \infty$, e valószínűségi változó eloszlásfüggvénye. Az $F(x)$ eloszlásfüggvény teljesíti a következő tulajdonságokat.*

a) *$F(x)$ monoton növekvő függvény.*
b) *$F(x)$ balról folytonos függvény, azaz $\lim\limits_{h\to 0,\, h>0} F(x-h) = F(x)$ minden $-\infty < x < \infty$ számra.*
c) $\lim\limits_{x\to\infty} F(x) = 1$.
d) $\lim\limits_{x\to-\infty} F(x) = 0$.

Tétel valószínűségi változók eloszlásának jellemzéséről. *Ha egy $F(x)$ függvény teljesíti az eloszlásfüggvények tulajdonságairól szóló lemmában megfogalmazott a)—d) tulajdonságokat, akkor az eloszlásfüggvény. Részletesebben megfogalmazva ez azt jelenti, hogy ha az $F(x)$ függvény teljesíti az a)—d) feltételeket, akkor létezik egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és azon olyan $\xi(\omega)$ valószínűségi változó, amelynek az eloszlásfüggvénye ez az $F(x)$ függvény.*

A fent megfogalmazott Tétel bizonyítása felhasznál egy fontos mértékelméleti eredményt, amelyet megfogalmazok. De e mértékelméleti eredmény bizonyítása a mértékelmélet anyaga, ezért azt itt nem tárgyaljuk.

Tétel eloszlásfüggvények által meghatározott mértékekről. *Ha egy $F(x)$ függvény teljesíti az eloszlásfüggvények tulajdonságairól szóló lemmában megfogalmazott a)—d) tulajdonságokat akkor létezik egy olyan az F függvény által egyértelműen meghatározott $\mu_F(\cdot)$ úgynevezett Stieltjes mérték, amely teljesíti a következő tulajdonságokat.*

a) A számegyenes minden B Borel halmazának létezik $\mu_F(B)$ Stieltjes mértéke.

b) A $\mu_F(\cdot)$ halmazfüggvény σ-additív a számegyenes Borel-halmazaiból álló (Borel) σ-algebrán, és $\mu_F(R^1) = 1$.

c) $\mu_F([a,b)) = F(b) - F(a)$ minden $-\infty < a < b < \infty$ számra. Innen az is következik, hogy $\mu_F((-\infty, b)) = F(b)$ minden $-\infty < b < \infty$ számra.

Az eloszlásfüggvények által meghatározott mértékekről szóló tételből és a mértékelmélet egy fontos eredményéből, a második téma leírásának 2. kiegészítésében ismertetett Carathéodory-féle kiterjesztési tételből az is következik, hogy az $F(x)$ eloszlásfüggvény egyértelműen meghatározza a $P(\{\omega\colon\ \xi(\omega) \in B\}) = F_\mu(B)$ valószínűségeket minden Borel–mérhető B halmazra.

Az eloszlásfüggvények tulajdonságairól szóló lemma bizonyítása: Mivel $\{\omega\colon\ \xi(\omega) < a\} \subset \{\omega\colon\ \xi(\omega) < b\}$, ha $a < b$, ezért $P(\{\omega\colon\ \xi(\omega) < a\}) \le P(\{\omega\colon\ \xi(\omega) < b\})$ ebben az esetben, és ez az a) tulajdonság.

Az a) tulajdonság érvényessége miatt a b) tulajdonság bizonyítása érdekében elég megmutatni azt, hogy ha h_n, $n = 1, 2, \dots$, olyan monoton csökkenő sorozat, amelyre $\lim\limits_{n\to\infty} h_n = 0$, akkor $\lim\limits_{n\to\infty} F(x - h_n) = \lim\limits_{n\to\infty} P(\{\omega\colon\ \xi(\omega) < x - h_n\}) = F(x) = P(\{\omega\colon\ \xi(\omega) < x\})$. Ez viszont következik az $\bigcup\limits_{n=1}^{\infty} \{\omega\colon\ \xi(\omega) < x - h_n\} = \{\omega\colon\ \xi(\omega) < x\}$ relációból és a valószínűségi mérték folytonosságából. A c) és d) tulajdonság bizonyítása hasonló.

A valószínűségi változók eloszlásának jellemzéséről szóló tétel bizonyítása az eloszlásfüggvények által meghatározott mértékekről szóló tétel segítségével. Tekintsük a következő $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt: $\Omega = R^1$, a számegyenes, $\mathcal{A} = \mathcal{B}$, a számegyenes Borel mérhető halmazainak σ-algebrája, $P = \mu_F$, a Borel mérhető halmazok σ-algebráján az F függvény által definiált Stieltjes mérték, amelynek létezését az eloszlásfüggvények által meghatározott mértékekről szóló tétel állítása fogalmazza meg. Legyen $\xi(x) = x$, azaz jelen példában az ω elemi események az x valós számok, és a $\xi(x)$ valószínűségi változó az x helyen az x számmal egyenlő. Ekkor $P(\{\omega\colon\ \xi(\omega) < x\}) = \mu_F(u\colon\ u < x) = F(x)$ minden x számra. Ez azt jelenti, hogy a definiált $\xi(\omega)$ valószínűségi változó eloszlása az $F(x)$ eloszlásfüggvény.

1. megjegyzés. Az előző bizonyítás egyetlen nehezebb lépése annak igazolása, hogy a konstruált μ_F halmazfüggvény (valószínűségi) mérték, azaz σ-additív. Megjegyzem, hogy ennek az állításnak a bizonyítása elkerülhetetlen. Ugyanis az eloszlásfüggvények által meghatározott mértékekről szóló tétel állítása könnyen levezethető az előző tétel eredményéből.

2. megjegyzés. Felmerülhet a kérdés, hogy mi a jelentősége számunkra a valószínűségi változók eloszlásának jellemzéséről szóló tételnek. Ezen eredmény fontosságának oka a következő. Gyakran megfogalmazunk olyan feladatot, amelyben egy olyan valószínűségi változóról akarunk tudni valamit, amelyiknek az eloszlásfüggvényét egy képlettel megadtuk. Felmerülhet a kérdés, hogy értelmes feladatot fogalmaztunk-e meg, azaz létezik-e ilyen eloszlású valószínűségi változó. A valószínűségi változók eloszlásának jellemzéséről szóló tétel azt állítja, hogy amennyiben az általunk megadott képlet teljesít néhány olyan feltételt, amelyet minden eloszlásfüggvénynek teljesítenie kell, akkor ilyen valószínűségi változó létezik.

Feladat:

Lássuk be, hogy létezik olyan diszkrét eloszlású ξ valószínűségi változó, amelynek $F(x) = P(\xi < x)$ eloszlásfüggvényére igaz, hogy minden $-\infty < a < b < \infty$ számpárra $F(a) < F(b)$ szigorú egyenlőtlenséggel.

Megoldás: Egy lehetséges konstrukció a következő. Legyen $r_1, r_2, \dots,$ a racionális számok sorozata valamilyen sorrendben felsorolva. Definiáljunk olyan ξ valószínűségi változót, amelyre $P(\xi = r_j) = 2^{-j}$, $j = 1, 2, \dots$. Ekkor ξ diszkrét eloszlású valószínűségi változó, és ennek $F(x)$ eloszlására $F(b) - F(a) = P(a < \xi < b) > 0$, mert $P(a < \xi < b) > P(\xi = r_j) > 0$, ahol r_j egy az $[a, b]$ intervallum belsejében lévő racionális szám.

A következő témánk valószínűségi változók várható értékének a definiciója és legfontosabb tulajdonságai az általános esetben.

Valószínűségi változók várható értéke.

A valószínűségszámítás egy nagyon fontos fogalma a valószínűségi változók várható értéke. Ezt a fogalmat részletesen tárgyaltuk diszkrét eloszlású valószínűségi változók esetében. Az általános eset tárgyalása visszavezethető erre a speciális esetre alkalmas határátmenet segítségével. Ezt a határátmenet eljárást kissé általánosabb formában elvégezték a mértékelméletben az úgynevezett Lebesgue integrál bevezetésénél. A valószínűségszámítás tárgyalásában a várható érték vizsgálatát egyszerűen és gyorsan el tudják végezni, ha szabad használni a Lebesgue integrál fogalmát.

Ebben az előadásban a következő köztes megoldást választom. Elmagyarázom azt a képet, amely természetessé teszi a használt Lebesgue mérték definicióját, és szemléletesen megmutatom, hogy miért érvényesek a legfontosabb eredmények. A formális részletek kidolgozása viszont nem ennek az előadásnak a témája.

Valószínűségi változó várható értékének *formális* **definiciója.** *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. E valószínűségi változó várható értékét úgy definiáljuk, mint a $\xi(\omega)$ függvénynek az*

$$E\xi(\omega) = \int \xi(\omega) dP(\omega)$$

Lebesgue integrálját a P valószínűségi mérték szerint, feltéve, hogy $\int |\xi(\omega)| dP(\omega) < \infty$. Ha ez a feltétel nem teljesül, akkor nem definiáljuk a ξ valószínűségi változó várható értékét.

A következő eredmény lehetővé teszi egy valószínűségi változó várható értékének kiszámítását csak a valószínűségi változó eloszlásfüggvényének az ismeretében.

Tétel várható érték kiszámításáról az eloszlásfüggvény segítségével. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn egy $F(x) = P(\xi < x)$, $-\infty < x < \infty$, eloszlásfüggvénnyel. Ekkor*

$$E\xi(\omega) = \int_{-\infty}^{\infty} xF(\,dx),$$

ahol a fenti integrál Lebesgue–Stieltjes integrált jelöl az F eloszlás által meghatározott μ_F Stieltjes mérték szerint. Ez a formula akkor érvényes, ha $\int_{-\infty}^{\infty} |x|F(\,dx) < \infty$. Ellenkező esetben az $E\xi$ várható értéket nem definiáltuk.

Sőt, igaz a következő általánosítása a várható érték kiszámításáról szóló tételnek az eloszlásfüggvény segítségével.

A várható értéknek az eloszlásfüggvény segítségével való kiszámításáról szóló tétel általánosítása. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelynek $F(x) = P(\xi < x)$, $-\infty < x < \infty$, az eloszlásfüggvénye. Legyen $g(x)$ tetszőleges (mérhető) függvény a számegyenesen, és definiáljuk az $\eta(\omega) = g(\xi(\omega))$ valószínűségi változót. Ekkor*

$$E\eta(\omega) = Eg(\xi(\omega)) = \int_{-\infty}^{\infty} g(x)F(\,dx),$$

ahol a fenti integrál Lebesgue–Stieltjes integrált jelöl az F eloszlás által meghatározott μ_F Stieltjes mérték szerint. Ez a formula akkor érvényes, ha $\int_{-\infty}^{\infty} |g(x)|F(\,dx) < \infty$. Ellenkező esetben az $E\eta$ várható értéket nem definiáltuk.

1. megjegyzés: Az előbb megfogalmazott eredményekben szerepelt az $\int_{-\infty}^{\infty} |x|F(\,dx) < \infty$, illetve $\int_{-\infty}^{\infty} |g(x)|F(\,dx) < \infty$ feltétel. E feltételek természetes megfelelői annak a diszkrét valószínűségi változók esetében szereplő feltételnek, hogy egy diszkrét valószínűségi változó várható értékét definiáló összeg legyen abszolut konvergens.

2. megjegyzés: Ismertettem egy modellt, amelyben egy adott $F(x)$ eloszlásfüggvényhez konstruáltunk egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt és azon megadtunk egy olyan ξ valószínűségi változót, amelynek $F(x)$ az eloszlásfüggvénye. Az $(\Omega, \mathcal{A}, P)$ valószínűségi mező az $(R^1, \mathcal{B}, \mu_F)$ rendszer volt ebben a modellben, és a $\xi(x) = x$ képlet adta meg a kívánt tulajdonságú valószínűségi változót, ahol R^1 a számegyenes, $\mathcal{B}$ a Borel σ-algebra és μ_F az F eloszlás által indukált Stieltjes mérték. Ebben a modellben az előző két a várható érték kiszámításáról szóló tétel következik a várható érték definiciójából. A tételek tartalma az, hogy a fenti azonosságok modell függetlenek, azaz tetszőleges $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált $F(x)$ eloszlású ξ valószínűségi változóra érvényesek.

A várható érték definiciójában szerepelt a Lebesgue integrál fogalma. Ennek a fogalomnak a jelentése némi magyarázatra szorul. Számunkra elég lesz egy olyan kép magyarázatként, amelyik szemléletesen megmutatja ennek értelmét. Ilyen magyarázatot adok a jegyzet végén egy rövid kiegészítésben. Ott azonban nem törekszem teljességre. Külön érdemes megjegyezni, hogy maga az $\int \xi(\omega)\, P(\, d\omega)$ integrál azon a valószínűségi mezőn van definiálva, ahol a $\xi(\omega)$ valószínűségi változó, (emlékezzünk, hogy egy valószínűségi változó egy valószínűségi mezőn értelmezett (mérhető) függvény), és magát az integrálást is ugyanezen a mezőn hajtjuk végre. Mégis a várható értéknek az eloszlásfüggvény segítségével történő kiszámításáról szóló tételben, illetve annak általánosításában olyan eredményt fogalmaztam meg, amely lehetővé teszi egy ilyen integrál kiszámolását anélkül, hogy explicite megadnánk azt a valószínűségi mezőt, ahol dolgozunk. A tétel talán legfontosabb állítása az, hogy elég tudnunk a valószínűségi változó eloszlásfüggvényét, és akkor ki tudjuk számolni a valószínűségi változó, sőt a valószínűségi változó egy tetszőleges függvényének a várható értékét.

Az itt szereplő integrál általánosabb, mint az analízisben tanult Riemann illetve Riemann–Stieltjes integrál. Viszont, ha az utóbbiak léteznek, akkor azok megegyeznek a minket érdeklő integrállal. Végül megjegyzem, hogy tanultuk korábban, hogyan lehet kiszámolni diszkrét eloszlású valószínűségi változók, illetve azok függvényeinek várható értékét. Ezt érdemes ideírni, hogy lássuk hasonlóságukat a várható érték kiszámításáról szóló tétellel az eloszlásfüggvény segítségével, illetve ezen ezen eredmény általánosításával. Valójában a diszkrét eloszlású valószínűségi változók várható értékének kiszámításáról szóló eredmények természetes megfelelőit fogalmaztam meg a fenti állításokban.

Legyen ξ diszkrét értékű valószínűségi változó, amely valamely $x_1, x_2, \dots,$ valós értékeket vesznek fel. Ekkor

$$E\xi = \sum_i x_i P(\xi = x_i)$$
$$Eg(\xi) = \sum_i g(x_i) P(\xi = x_i),$$

feltéve, hogy a jobboldalon szereplő összegek abszolut konvergensek.

A Lebesgue integrál teljesíti a Riemann integrálra érvényes additivitási tulajdonság következő természetes általánosítását:

$$\int (c_1\xi_1(\omega) + c_2\xi_2(\omega))\, dP(\omega) = c_1 \int \xi_1(\omega) dP(\omega) + c_2 \int \xi_2(\omega) dP(\omega)$$

minden (integrálható) ξ_1 és ξ_2 valószínűségi változóra és valós c_1, c_2 számra. Ennek az azonosságnak a következménye (valójában átfogalmazása) a következő

Tétel a várható érték additivitásáról. *Ha két ξ_1, ξ_2 valószínűségi változónak (amelyek ugyanazon a valószínűségi mezőn vannak definiálva) létezik várható értéke, c_1 és c_2 két valós szám, akkor a $c_1\xi_1 + c_2\xi_2$ valószínűségi változónak is létezik várható értéke, és*

$$E\,(c_1\xi_1 + c_2\xi_2) = c_1 E\xi_1 + c_2 E\xi_2.$$

Megjegyzés: Ez az eredmény azt állítja, hogy az általános esetben is érvényes az a diszkrét valószínűségi változókról szóló eredmény, amely véletlen összegek várható értékének additivitását fejezi ki. Ebben az esetben sem kell megkövetelni a valószínűségi változók (általános esetben még nem tárgyalt) függetlenségét. A fenti eredmény a várható értéknek eredeti definiciójából adódott, amely a várható értéket mint egy "absztrakt" Lebesgue integrált adta meg. Sokszor hasznosabb a várható érték eredeti definicója helyett azt az eredményt használni, amely lehetővé teszi egy valószínűségi változó várható értékének kiszámítását akkor is, ha csak a valószínűségi változó eloszlásfüggvényét ismerjük. Mégis, mint a fenti eredmény mutatja, néha érdemes a várható érté eredeti definiciójára hivatkozni.

A várható érték (Lebesgue integrál formájában megadott definiciójából) könnyen látható, hogy ha egy $\xi(\omega)$ valószínűségi változóra teljesül a $P(\xi(\omega) \geq 0) = 1$ feltétel, akkor $E\xi(\omega) \geq 0$. Innen következik, hogy amennyiben két $\xi(\omega)$ és $\eta(\omega)$ valószínűségi változóra teljesül, hogy $P(\xi(\omega) \geq \eta(\omega)) = 1$, akkor $E\xi(\omega) \geq E\eta(\omega)$. Valóban, ekkor $E\xi(\omega) - E\eta(\omega) = E(\xi(\omega) - \eta(\omega)) \geq 0$. Ennek az észrevételnek egyik következménye az alábbi Markov egyenlőtlenségnek nevezett állítás.

Markov egyenlőtlenség. *Ha $\xi(\omega)$ nem negatív valószínűségi változó, azaz ha $P(\xi(\omega) \geq 0) = 1$ akkor*

$$P(\xi(\omega) \geq x) \leq \frac{E\xi(\omega)}{x} \quad \textit{minden } x > 0 \textit{ számra.}$$

Bizonyítás: Definiáljuk a ξ valószínűségi változó alábbi $\eta(\omega)$ függvényét:

$$\eta(\omega) = \begin{cases} x & \text{ha } \xi(\omega) \geq x \\ 0 & \text{ha } 0 \leq \xi(\omega) < x \end{cases}$$

Ekkor $P(\xi(\omega) \geq \eta(\omega)) = 1$, ezért $E\xi(\omega) \geq E\eta(\omega) = xP(\eta(\omega) \geq x) = xP(\xi(\omega) \geq x)$, ahonnan következik az állítás.

Definiáljuk általános esetben is valószínűségi változók szórásnégyzetét.

Szórásnégyzet definiciója. *Legyen ξ olyan valószínűségi változó, amelyre $E\xi^2 < \infty$. Ekkor e valószínűségi változó szórásnégyzete*

$$\operatorname{Var} \xi = E(\xi - E\xi)^2$$

(Ha $E\xi^2 = \infty$, akkor vagy nem definiáljuk a ξ valószínűségi változó szórásnégyzetét, vagy azt mondjuk, hogy $\operatorname{Var} \xi(\omega) = \infty$.

Lemma.

$$\operatorname{Var} \xi = E\xi^2 - (E\xi)^2 .$$

Bizonyítás:

$$\operatorname{Var} \xi = E(\xi - E\xi)^2 = E\left(\xi^2 - 2\xi E\xi + (E\xi)^2\right) = E\xi^2 - 2E\xi E\xi + (E\xi)^2 = E\xi^2 - (E\xi)^2 .$$

Megjegyzés: Az előző lemmából és a $\mathrm{Var}\,\xi \ge 0$ egyenlőtlenségből következik, hogy $(E\xi)^2 \le E\xi^2$. Ennek az állításnak igaz a következő Cauchy–Schwarz egyenlőtlenségnek nevezett egyenlőtlenség: $(E\xi\eta)^2 \le E\xi^2 E\eta^2$. Sőt, ez az állítás érvényesebb az alábbi élesebb formában is: Adva egy $(X, \mathcal{A}, \mu)$ mértéktér, azon két (mérhető) $f(x)$ és $g(x)$ függvény, $x \in X$, akkor $\left(\int f(x)g(x)\mu(\,dx)\right)^2 \le \int f(x)^2\mu(\,dx) \int g(x)^2\mu(\,dx)$. Ez azt jelenti, hogy a fenti egyenlőtlenség érvényes tetszőleges (σ-véges) és nemcsak egyre normált, azaz valószínűségi mérték szerinti integrálra.

Bár a fenti egyenlőtlenségre nem lesz szükségünk később, és annak bizonyítását nem kell tudni a valószínűségszámítás vizsgán, mégis ismertetem annak bizonyítását, mivel az tanulságos. Ezt a bizonyítást megkaphatjuk az előző egyenlőtlenség bizonyításának a finomításával. Nevezetesen, tetszőleges λ valós számra felírhatjuk, hogy $0 \le \int (f(x) + \lambda g(x))^2 \mu(\,dx) = \lambda^2 \int g(x)^2\mu(\,dx) + 2\lambda \int f(x)g(x)\mu(\,dx) + \int f(x)^2\mu(\,dx)$. Válasszuk a λ számot a fenti egyenlőtlenségben, mint $\lambda = -\frac{\int f(x)g(x)\mu(\,dx)}{\int g(x)^2\mu(\,dx)}$. (Jegyezzük meg, hogy az általánosság megszorítása nélkül feltehetjük, hogy $\int g(x)^2\mu(\,dx) > 0$. Ellenkező esetben ugyanis $\int g(x)^2\mu(\,dx) = 0$, ahonnan $g(x) = 0$ (majdnem) minden $x \in X$ számra (a μ mérték szerint). A λ szám fenti választása nagyon természetes, mivel λ-nak egy olyan másodfokú polinomját tekintettük, amely minden λ paraméterre pozitív, és mi azt a paramétert választottuk, ahol ez a másodfokú polinom felveszi a minimumát.) Ezzel a választással azt kapjuk, hogy

$$-\frac{\left(\int f(x)g(x)\mu(\,dx)\right)^2}{\int g(x)^2\mu(\,dx)} + \int f(x)^2\mu(\,dx) \ge 0,$$

ami ekvivalens a Cauchy–Schwarz egyenlőtlenséggel.

Általában a szórásnégyzet tulajdonságainak a bizonyítása csak a várható érték tulajdonságait használja. Ezért a diszkrét valószínűségi változók esetében érvényes

$$\mathrm{Var}\,(a\xi + b) = a^2 \mathrm{Var}\,\xi$$

azonosságot hasonlóan lehet bizonyítani az általános esetben.

Feladat:

Általános valószínűségi változók esetében is érvényes a

$$\inf_{-\infty < M < \infty} E(\xi - M)^2 = \mathrm{Var}\,\xi$$

azonosság.

Megoldás: Az ötödik téma ismertetésének 1. feladatában bebizonyítottuk ezt az állítást diszkrét eloszlású valószínűségi változókra. Az ott ismertetett megoldás változatás nélkül alkalmazható a most tárgyalt általános esetben is.

Megfogalmazom és bebizonyítom az alábbi Csebisev egyenlőtlenségnek nevezett állítást.

Csebisev egyenlőtlenség. *Minden ξ valószínűségi változóra, amelyre $E\xi^2 < \infty$*

$$P(|\xi - E\xi| \ge x) \le \frac{\operatorname{Var}\xi}{x^2} \quad \textit{minden } x \ge 0 \textit{ számra.}$$

Bizonyítás:

$$P(|\xi - E\xi| \ge x) = P((\xi - E\xi)^2 \ge x^2) \le \frac{E(\xi - E\xi)^2}{x^2} = \frac{\operatorname{Var}\xi}{x^2}$$

a Markov egyenlőtlenség alapján.

Megjegyzés: A Csebisev egyenlőtlenség azért hasznos, mert a benne szereplő szórásnégyzet sok érdekes esetben könnyen kiszámolható. Felmerülhet az a kérdés, hogy a Csebisev egyenlőtlenség mennyire éles. Erre a kérdésre később még visszatérek.

Többváltozós eloszlásfüggvények, és valószínűségi változók függetlensége.

Megtárgyaljuk az eloszlásfüggvény többdimenziós változatát. E fogalom bevezetése után lehet természetes módon definiálni valószínűségi változók függetlenségét. Az itt szereplő fogalmak és bizonyítások hasonlóak az egydimenziós esethez. Ezen eredmények bizonyítását csak vázlatosan ismertetem. Megfogalmazom azokat a mértékelméleti eredményeket, amelyeket a bizonyítások felhasználnak. Elsősorban a fogalmak és eredmények tisztességes leírására fogok törekedni.

Többdimenziós eloszlásfüggvény definiciója. *Legyen adva k valós értékű $\xi_1, \dots, \xi_k$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ezek (együttes) eloszlásfüggvénye az*

$$F(x_1, \dots, x_k) = P(\xi_1 < x_1, \dots, \xi_k < x_k),$$

k változós függvény, ahol $-\infty < x_j < \infty$ minden $1 \le j \le k$ indexre.

Megfogalmaztam az egydimenziós eloszlásfüggvények jellemzését egy tételben és az azt megelőző lemmában. Érvényes ezeknek az eredményeknek a természetes többdimenziós változata is. A többváltozós eloszlásfüggvényeket jellemző eredmény megfogalmazásához a következő kérdést kell tisztázni. Az előbb említett lemmában felsoroltuk az egyváltozós függvényeknek a tulajdonságait. Mik e tulajdonságok többváltozós megfelelői? Többek között meg kell találnunk a *Lemma eloszlásfüggvények tulajdonságairól* eredményben szereplő az eloszlásfüggvény monotonítását kimondó a) feltételnek a többdimenziós megfelelőjét.

Ezt a következőképp találhatjuk meg. Azt, hogy az $F(x)$ eloszlásfüggvény monoton az egyvátozós esetben, azaz $F(b) \ge F(a)$, ha $b > a$ úgy is megfogalmazható, hogy $P(\xi \in [a, b) = F(b) - F(a) \ge 0$. Ennek a tulajdonságnak a többváltozós megfelelője azt a tényt fejezi ki, hogy egy véletlen vektor nem negatív valószńűséggel esik egy téglatestbe. Ezen állításnak az eloszlásfüggvények segítségével való megfogalmazása érdekében vezessük be a következő jelöléseket.

Tekintsünk egy tetszőleges $\mathbf{K} = [a_1, b_1) \times \cdots \times [a_k, b_k)$ alakú téglatestet, $-\infty < a_j < b_j < \infty$, $j = 1, \ldots, k$, a k dimenziós térben. Világos, hogy tetszőleges $(\xi_1, \ldots, \xi_k)$ k-dimenziós valószínűségi változóra teljesül a

$$P\left((\xi_1(\omega), \ldots, \xi_k(\omega)) \in \mathbf{K}\right) = P\left(\bigcap_{j=1}^{k} \{\omega\colon\ a_j \le \xi_j(\omega) < b_j\}\right) \ge 0$$

tulajdonság. Azt állítom, hogy az ilyen alakú események valószínűsége viszonylag egyszerűen kifejezhető a $(\xi_1, \ldots, \xi_k)$ véletlen vektor $F(x_1, \ldots, x_k) = P(\xi_1 < x_1, \ldots, \xi_k < x_k)$ eloszlásfüggvényének a segítségével, és az így kapott kifejezés nem-negatív volta az eloszlásfüggvények monotonitásáról szóló a) feltétel megfelelője a többdimenziós esetben.

Annak érdekében, hogy ezt a feltételt megfogalmazzuk, a következő állítást kell igazolni. Tekintsük egy $\mathbf{K} = [a_1, b_1) \times \cdots \times [a_k, b_k)$ alakú téglatestet, és ennek csúcspontjait, azaz az olyan $(u_1, \ldots, u_k)$ pontokat, amelyek koordinátái az a_j vagy b_j számok, $1 \le j \le k$. Annak valószínűsége, hogy egy $F(x_1, \ldots, x_k)$ eloszlású $(\xi_1, \ldots, \xi_k)$ véletlen vektor a $\mathbf{K} = [a_1, b_1) \times \cdots \times [a_k, b_k)$ téglatestbe esik kifejezhető mint az $F(\cdot)$ eloszlásfüggvénynek a téglatest csúcspontjaiban felvett értékeinek alkalmas lineáris kombinációja. Pontosabban,

$$P((\xi_1, \ldots, \xi_k) \in \mathbf{K}) = \sum_{\substack{u_j = \, a_j \text{ vagy } b_j \\ j=1,\ldots,k}} (-1)^{\chi(u_1, \ldots, u_k)} F(u_1, \ldots, u_k),$$

ahol $\chi(u_1, \ldots, u_k)$ jelöli az a_j-k számát az $u_1, \ldots, u_k$ sorozatban. Az $F(\cdot)$ függvény monotonitásának az többdimenziós esetben az felel meg, hogy a fent felírt azonosság jobboldalán szereplő kifejezés nem-negatív. A jobb megértés érdekében ellenőrizzük ezt az azonosságot először a $k = 2$ speciális esetben.

Ekkor azt kell megmutatni, hogy $P(a_1 \le \xi_1 < b_1, a_2 \le \xi_2 < b_2) = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2)$. Viszont $F(b_1, b_2) - F(b_1, a_2) = P(\xi_1 < b_1, a_2 \le \xi_2 < b_2)$, és $F(a_1, b_2) - F(a_1, a_2) = P(\xi_1 < a_1, a_2 \le \xi_2 < b_2)$. Ezt a két azonosságot kivonva egymásból megkapjuk a kívánt állítást.

Feladat:

11.) Legyen egy (ξ_1, ξ_2) véletlen vektor eloszlása az $F(x_1, x_2) = P(\xi_1 < x_1, \xi_2 < x_2)$ függvény, $-\infty < a_1 < b_1 < \infty$, $-\infty < a_2 < b_2 < \infty$ valós számok. Fejezzük ki az $F(x_1, x_2)$ eloszlásfüggvény segítségével a $P(a_1 \le \xi_1 < b_1, a_2 \le \xi_2 < b_2)$, a $P(a_1 \le \xi_1 \le b_1, a_2 \le \xi_2 \le b_2)$ és a $P(a_1 < \xi_1 < b_1, a_2 < \xi_2 < b_2)$ valószínűségeket.
Megoldás: Először azt megmutatjuk meg, hogy $P(a_1 \le \xi_1 < b_1, a_2 \le \xi_2 < b_2) = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2)$. Valóban, $F(b_1, b_2) - F(b_1, a_2) = P(\xi_1 < b_1, a_2 \le \xi_2 < b_2)$, és $F(a_1, b_2) - F(a_1, a_2) = P(\xi_1 < a_1, a_2 \le \xi_2 < b_2)$. Ezt a két azonosságot kivonva egymásból megkapjuk a kívánt állítást.

$$P(a_1 \le \xi_1 \le b_1, a_2 \le \xi_2 \le b_2) = \lim_{n \to \infty} P\left(a_1 \le \xi_1 < b_1 + \frac{1}{n}, a_2 \le \xi_2 < b_2 + \frac{1}{n}\right)$$

$$= \lim_{n\to\infty}\left(F\left(b_1+\frac{1}{n}, b_2+\frac{1}{n}\right) - F\left(b_1+\frac{1}{n}, a_2\right) - F\left(a_1, b_2+\frac{1}{n}\right) + F(a_1,a_2)\right)$$
$$= \lim_{n\to\infty} F\left(b_1+\frac{1}{n}, b_2+\frac{1}{n}\right) - \lim_{n\to\infty} F\left(b_1+\frac{1}{n}, a_2\right)$$
$$- \lim_{n\to\infty} F\left(a_1, b_2+\frac{1}{n}\right) + F(a_1,a_2).$$

Megjegyzem, hogy az eloszlásfüggvény tulajdonságaiból következik, hogy a felírt határértékek, például a $\lim\limits_{n\to\infty} F\left(b_1+\frac{1}{n}, b_2+\frac{1}{n}\right)$ kifejezés valóban létezik. Hasonlán,

$$P(a_1<\xi_1<b_1, a_2<\xi_2<b_2) = \lim_{n\to\infty} P\left(a_1+\frac{1}{n} \le \xi_1 < b_1, a_2+\frac{1}{n} \le \xi_2 < b_2\right)$$
$$= \lim_{n\to\infty} F(b_1,b_2) - F\left(b_1, a_2+\frac{1}{n}\right)$$
$$- \lim_{n\to\infty} F\left(a_1+\frac{1}{n}, b_2\right) + \lim_{n\to\infty} F\left(a_1+\frac{1}{n}, a_2+\frac{1}{n}\right).$$

Feladat:

12.) Legyen az $F(x_1,\dots,x_k)=P(\xi_1<x_1,\dots,\xi_k<x_k)$ függvény a $(\xi_1,\dots,\xi_k)$ véletlen vektor eloszlásfüggvénye. Mutassuk meg, hogy

$$I(\{\omega\colon\ a_j \le \xi_j(\omega) < b_j,\ 1\le j\le k\})$$
$$= \sum_{\substack{u_j = \ a_j \text{ vagy } b_j \\ j=1,\dots,k}} (-1)^{\chi(u_1,\dots,u_k)} I(\{\omega\colon\ \xi_1(\omega)<u_1,\dots,\xi_k(\omega)<u_k\}),$$

(a_j,b_k), $j=1,\dots,k$ számpárok minden k-asára, ezért a várható érték additívitásából következik, hogy

$$P((a_j\le\xi_j<b_j,\ 1\le j\le k) = \sum_{\substack{u_j = \ a_j \text{ vagy } b_j \\ j=1,\dots,k}} (-1)^{\chi(u_1,\dots,u_k)} F(u_1,\dots,u_k).$$

A fenti formulákban $\chi(u_1,\dots,u_k)$ jelöli az a_j-k számát az $u_1,\dots,u_k$ sorozatban, $I(A)$, $A\subset\Omega$ pedig egy A halmaz indikátorfüggvénye, azaz $I(A)(\omega)=1$, ha $\omega\in A$, és $I(A)(\omega)=0$, ha $\omega\notin A$.

Megoldás: Adva valamely $u_1,\dots,u_k$ számok, definiáljuk a

$$B(u_1,\dots,u_k)=\{(\omega\colon\ \xi_j(\omega)<u_j,\ 1\le j\le k\}\subset\Omega$$

eseményt, és legyen $I(B(u_1,\dots,u_k))(\omega)$ ezen esemény indikátorfüggvénye. Legyen továbbá $I(\mathbf{K})(\omega)$ a $\mathbf{K}=\{\omega\colon\ a_j\le\xi_j<b_j,\ 1\le j\le k\}$ halmaz indikátorfüggvénye. Elég megmutatni, hogy

$$I(\mathbf{K})(\omega) = \sum_{\substack{u_j = \ a_j \text{ vagy } b_j \\ j=1,\dots,k}} (-1)^{\chi(u_1,\dots,u_k)} I(B(u_1,\dots,u_k))(\omega),$$

mert véve ezután mind a két oldal várható értékét megkapjuk a kívánt azonosságot. Az is világos, hogy $\{\omega\colon\ a_j \le \xi_j < b_j,\ 1 \le j \le k\}$ esetében a bizonyítandó azonosság mind a két oldala 1-gyel egyenlő. Azt kell észrevenni, hogy ekkor a jobboldalon az $I(B(b_1,\dots,b_k))(\omega)$ tag 1, és a jobboldal összes többi tagja nulla.
Ha $\omega \in \Omega$ olyan, hogy az $a_j \le \xi_j(\omega) < b_j$ egyenlőtlenségek nem teljesülnek minden $1 \le j \le k$ indexre, akkor a baloldal nulla. Azt kell belátni, hogy ugyanez érvényes a jobboldalon is. Világos, hogy a jobboldal nulla, ha $\xi_j(\omega) \ge b_j$ valamilyen j indexre. Azt az esetet kell még nézni, amikor $\xi_j(\omega) < b_j$ minden $1 \le j \le k$ indexre, bizonyos j indexekre $a_j \le \xi_j < b_j$, de létezik legalább egy olyan l index, amelyre $\xi_l < a_l$. Legyen L a legkisebb ilyen index. Ha párosítjuk az olyan $(-1)^{\chi(u_1,\dots,u_k)} I(B(u_1,\dots,u_k))(\omega)$ kifejezéseket, amelyeknek u_j koordinátátái megegyeznek $j \ne L$ esetben, de $u_L = a_L$ a pár egyik és $u_L = b_L$ a pár másik tagjára, akkor az összes ilyen pár hozadéka zéró. Ezért az azonosság ekkor is érvényes, mert mind a bal mind a jobboldal zéró ebben az esetben.

Be lehet látni, hogy az egyváltozós valószínűségi változók eloszlásának jellemzéséről szóló tételnek és az azt megelőző lemmának igaz a következő többdimenziós általánosítása.

Tétel többváltozós valószínűségi változók eloszlásának jellemzéséről. *Egy $F(u_1,\dots,u_k)$ függvény akkor és csak akkor eloszlásfüggvénye alkalmas $\xi_1,\dots,\xi_k$ valószínűségi változóknak valamely $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, ha ez az F függvény teljesíti a következő (i)—(iv) tulajdonságokat.*

(i) $F(u_1,\dots,u_k)$ minden változójának balról folytonos függvénye.

(ii) $\displaystyle\lim_{\substack{u_j\to\infty \\ \text{minden } j=1,\dots,k \text{ indexre}}} F(u_1,\dots,u_k) = 1.$

(iii) $\displaystyle\lim_{\substack{u_j\to-\infty \\ \text{valamely } 1\le j\le k \text{ indexre}}} F(u_1,\dots,u_k) = 0.$ *(Ez úgy értendő, hogy az összes u_s, $1 \le s \le k$, $s \ne j$ koordinátát rögzítjük, és $u_j \to -\infty$.)*

Végül definiáljuk egy az R^k téren definiált F függvényre és egy $\mathbf{K} = [a_1,b_1) \times \cdots \times [a_k,b_k)$ téglatestre a

$$\mu(\mathbf{K}) = \mu_F(\mathbf{K}) = \sum_{\substack{u_j = a_j \text{ vagy } b_j \\ j=1,\dots,k}} (-1)^{\chi(u_1,\dots,u_k)} F(u_1,\dots,u_k)$$

mennyiséget, ahol $\chi(u_1,\dots,u_k)$ jelöli az a_j-k számát az $u_1,\dots,u_k$ sorozatban. Ekkor

(iv) $\mu_F(\mathbf{K}) \ge 0$ minden $\mathbf{K}$ téglatestre.

A fenti tétel bizonyításának legnehezebb része annak megmutatása, hogy ha az $F(x_1,\dots,x_k)$ függvény teljesíti az (i)–(iv) tulajdonságokat, akkor léteznek olyan $\xi_1,\dots$, ξ_k valószínűségi változók alkalmas $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, amelyek (együttes) eloszlásfüggvénye az $F(x_1,\dots,x_k)$ függvény. Ennek bizonyítása az egyváltozós esetben bizonyított az eloszlásfüggvények által meghatározott mértékekről szóló tétel következő többváltozós általánosításán alapul.

Az eloszlásfüggvények által meghatározott mértékekről szóló tétel többváltozós általánosítása. *Legyen $F(x_1,\dots,x_k)$ olyan k változós függvény, amely teljesíti az (i)–(iv) tulajdonságokat. Ekkor létezik és egyértelműen meghatározott egy olyan μ_F Stieltjes mérték az R^k k-dimenziós tér Borel mérhető halmazainak $\mathcal{B}^k$ σ-algebráján, amelyik nem negatív σ-additiv halmazfüggvény ezen a σ-algebrán, $\mu_F(R^k)=1$, és*

$$\mu_F(\{u_1,\dots,u_k\}\colon\, u_j<x_j,\, 1\le j\le k\})=F(x_1,\dots,x_k)$$

minden $x_1,\dots,x_k$ valós számra.

E tétel segítségével a következő módon konstruálhatunk $F(x_1,\dots,x_k)$ eloszlásfüggvénnyel rendelkező valószínűségi változókat. Legyen $(\Omega,\mathcal{A},P)=(R^k,\mathcal{B}^k,\mu_F)$, ahol $\mathcal{B}^k$ jelöli a Borel σ-algebrát az R^k k-dimenziós térben, és μ_F az előző tételben jellemzett az F eloszlásfüggvény által meghatározott Stieltjes mérték. Ebben a valószínűségi mezőben az $(x_1,\dots,x_k)\in R^k$ pontok alkotják az ω elemi eseményeket. A $(\xi_1,\dots,\xi_k)$ valószínűségi változókat a következő módon definiáljuk. Legyen $\xi_j(x_1,\dots,x_k)=x_j$, $1\le j\le k$. Ezen valószínűségi változók együttes eloszlása az $F(x_1,\dots,x_k)$ függvény.

Érvényes továbbá a várható értéknek az eloszlásfüggvény segítségével kiszámításáról szóló eredmény alábbi többdimenziós változata, amely segít akkor, ha véletlen vektorok függvényeinek a várható értékét kívánjuk kiszámolni.

Tétel véletlen vektorok várható értékének kiszámításáról e vektor eloszlásfüggvényének a segítségével. *Legyenek $\xi_1(\omega),\dots,\xi_k(\omega)$ valószínűségi változók egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, amelyeknek $F(x_1,\dots,x_k)=P(\xi_1(\omega)<x_1,\dots,\xi_k(\omega)<x_k)$ az eloszlásfüggvénye. Legyen $g(x_1,\dots,x_k)$ tetszőleges (mérhető) k-változós függvény, és definiáljuk az $\eta(\omega)=g(\xi_1(\omega),\dots,\xi_k(\omega))$ valószínűségi változót. Ekkor*

$$E\eta(\omega)=Eg(\xi_1(\omega),\dots,\xi_k(\omega))=\int g(x_1,\dots,x_k)\mu_F(\,dx_1,\dots,\,dx_k),$$

ahol a fenti integrál Lebesgue–Stieltjes integrált jelöl az F függvény által meghatározott μ_F Stieltjes mérték szerint. Ez a formula akkor érvényes, ha

$$\int |g(x_1,\dots,x_k)|\mu_F(\,dx_1,\dots,\,dx_k)<\infty.$$

Ellenkező esetben az $E\eta$ várható értéket nem definiáltuk.

Független valószínűségi változók és szorzatuk várható értéke.

Valószínűségi változók függetlenségének a definiciója. *Legyenek $\xi_1,\dots,\xi_n$ valószínűségi változók egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi változók függetlenek, ha minden $x_1,\dots,x_n$ valós számra*

$$P(\xi_1<x_1,\dots,\xi_n<x_n)=P(\xi_1<x_1)\cdots P(\xi_n<x_n).$$

A teljesség kedvéért megadom e fogalom definicióját általánosabb, vektorértékű valószínűségi változókra is, bár valószínűleg ebben az előadássorozatban nem jutunk el addig a pontig, ahol erre a fogalomra szükség van. (Mindenekelőtt a többváltozós centrális határeloszlástételre gondolok, mint olyan témakörre, ahol ez a fogalom megjelenik.)

Vektorértékű valószínűségi változók függetlenségének a definiciója. *Legyenek* $\xi^{(1)} = (\xi_{1,1},\dots,\xi_{1,k})$, $\dots$, $\xi^{(n)} = (\xi_{n,1},\dots,\xi_{n,k})$, *értékeiket az* R^k k*-dimenziós euklideszi térben felvevő valószínűségi változók (vektorok) egy* $(\Omega,\mathcal{A},P)$ *valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi vektorok függetlenek, ha minden* $x^{(1)} = (x_{1,1},\dots,x_{1,k})$, $\dots$, $x^{(n)} = (x_{n,1},\dots,x_{n,k})$, k*-dimenziós vektorra*

$$\begin{aligned}&P\left(\xi_{1,1}<x_{1,1},\dots,\xi_{1,k}<x_{1,k},\ \dots,\ \xi_{n,1}<x_{n,1},\dots,\xi_{n,k}<x_{n,k}\right)\\&\qquad=P\left(\xi_{1,1}<x_{1,1},\dots,\xi_{1,k}<x_{1,k}\right)\cdots P\left(\xi_{n,1}<x_{n,1},\dots,\xi_{n,k}<x_{n,k}\right).\end{aligned}$$

Megjegyzés: Vektorértékű valószínűségi változók függetlenségének a definiciójában semmilyen függetlenségi feltevést nem tettünk az egyes $\xi^{(j)} = (\xi_{j,1},\dots,\xi_{j,k})$, $1\le j\le k$, vektorok koordinátái között. Ezzel a definicióval egyelőre nem foglalkozunk.

Független valószínűségi változókkal való számolást megkönnyíti a mértékelmélet egyik alapvető eredménye, az alább ismertetendő Fubini tétel. Ennek megfogalmazása előtt teszek néhány megjegyzést.

A Fubini tétel a következő, a területi (Riemann-)integrálról szóló eredménynek az általánosítása Lebesgue integrálokra:

$$\int\!\!\int f(x,y)\,dx\,dy=\int\left(\int f(x,y)\,dx\right)dy \tag{1}$$

minden integrálható $f(\cdot,\cdot)$ függvényre. Ez az eredmény azt jelenti, hogy egy kétváltozós függvény területi integrálját úgy is kiszámolhatjuk, hogy először rögzítjük az egyik paramétert, (amelyet itt y-nal jelöltünk,) és kiszámítjuk az így kapott egyváltozós integrált. Ezután az első lépésben rögzített paraméter szerint integrálva az így kapott függvényt, megkapjuk a területi integrál értékét. Ez azt jelenti, hogy a területi integrál kiszámítható két egyszeres integrál szukcessziv alkalmazásának a segítségével. Érdemes megfogalmazni ezt az eredményt abban a speciális esetben, amikor az $f(x,y)$ függvény $f(x,y)=g_1(x)g_2(y)$ alakú. Ekkor az (1) formula a következő alakú:

$$\int\!\!\int g_1(x)g_2(y)\,dx\,dy=\int g_1(x)\,dx\int g_2(y)\,dy.$$

Valóban, az (1) képlet jobboldala ebben az esetben így írható fel:

$$\int\left(\int g_1(x)g_2(y)\,dx\right)dy=\int g_2(y)\left(\int g_1(x)\,dx\right)dy=\int g_1(x)\,dx\int g_2(y)\,dy.$$

A Fubini tétel az (1) formulában megfogalmazott azonosság általánosítása arra az esetre, amikor (a Lebesgue mérték szerinti) Riemann integrál helyett általános szorzat mértékek szerinti Lebesgue integrálokat tekintünk. Továbbá nem csak kétváltozós, hanem tetszőleges k, $k \ge 2$, változós integrálokat vizsgálunk. Ennek ismertetése előtt megfogalmazom a következő eredményt.

Legyenek $F_1(\cdot)$, … , $F_k(\cdot)$ eloszlásfüggvények a számegyenesen, és definiáljuk az

$$F(x_1, \dots, x_k) = F_1(x_1) \cdots F_k(x_k)$$

függvényt. Az így definiált $F(x_1, \dots, x_k)$ függvény egy k-változós eloszlásfüggvény.

Részletesebben kifejtve ez a következőt jelenti. E jegyzetben megadtam annak szükséges és elégséges feltételét, hogy egy egy illetve egy többváltozós függvény eloszlásfüggvény legyen. A fenti állítás tartalma az, hogy ha az $F_j(\cdot)$, $1 \le j \le k$, függvények teljesítik az (egyváltozós) eloszlásfüggvény jellemzését leíró tulajdonságokat, akkor az $F(x_1, \dots, x_k) = F_1(x_1) \cdots F_k(x_k)$ függvény teljesíti a többváltozós eloszlásfüggvény eloszlását leíró tulajdonságokat.

Nem nehéz belátni, hogy a fenti állításban az összes ellenőrizendő feltétel teljesül. Mivel e részletek kidolgozása egyszerű, és ennek különösebb tanulságai nincsenek, ezért ezt elhagyom.

Ha $F_j(\cdot)$, $1 \le j \le k$, egydimenziós eloszlásfüggvények, akkor tekinthetjük az $F(x_1, \dots, x_k) = F_1(x_1) \cdots F_k(x_k)$ k-változós eloszlásfüggvényt, illetve az általa meghatározott $\mu_F = \mu_{(F_1, \dots, F_k)}$ Stieltjes mértéket a k-dimenziós téren. Egy $F(x_1, \dots, x_k) = F_1(x_1) \cdots F_k(x_k)$ szorzat alakú k-változós eloszlásfüggvény által meghatározott Stieltjes mértéket szorzatmértéknek szokás nevezni. Ez ugyanis rendelkezik a következő tulajdonsággal. Tetszőleges $A = A_1 \times A_2 \times \cdots \times A_k$ alakú halmazra (azaz 1-dimenziós halmazok direkt szorzatára) $\mu_F(A) = \mu_{F_1}(A_1)\mu_{F_2}(A_2) \cdots \mu_{F_k}(A_k)$. Az alább megfogalmazandó Fubini tétel azt mondja ki, hogy az ilyen szorzatmértékek szerinti integrál a területi integrálhoz hasonló módon kiszámolható megfelelő szukcesszív integrálás segítségével. Független valószínűségi változók függvényeinek várható értékéről szóló legfontosabb eredmények megkaphatóak a Fubini tétel egyszerű következményeként, ha a vizsgált várható értéket kifejezzük az eloszlás (pontosabban az eloszlás által meghatározott Stieltjes mérték) szerinti Lebesgue integrálként. Ez a *Tétel véletlen vektorok várható értékének kiszámításáról e vektor eloszlásfüggvényének a segítségével* eredménye alapján lehetséges.

Az irodalomban általános szokás, hogy az előbb definiált μ_F k-dimenziós téren értelmezett mérték szerinti integrált $F_1(\,dx_1) \dots F_k(dx_k)$ vagy $dF_1(x_1) \dots dF_k(x_k)$-val jelölik, azaz

$$\begin{aligned}\int g(x_1, \dots, x_k)\mu_F(\,dx_1, \dots, \,dx_k) &= \int g(x_1, \dots, x_k)\mu_{(F_1, \dots, F_k)}(\,dx_1, \dots, \,dx_k) \\ &\overset{\text{jel.}}{=} \int g(x_1, \dots, x_k)F_1(\,dx_1) \dots F_k(\,dx_k) = \int g(x_1, \dots, x_k)\,dF_1(x_1) \dots \,dF_k(x_k)\end{aligned}$$

tetszőleges integrálható $g(x_1, \dots, x_k)$ (integrálható) fügvényre, ahol „jel." a jelölés szó rövidítése. A továbbiakban mi is ezt a jelölést követjük.

Fubini tétel. *Legyenek $F_j(\cdot)$, $1 \le j \le k$, eloszlásfüggvények a számegyenesen, és legyen $g(x_1,\dots,x_k)$ (mérhető) k-változós függvény. Ekkor*

$$\begin{aligned}&\int g(x_1,x_2,\dots,x_k)F_1(\,dx_1)F_2(\,dx_2)\dots F_k(\,dx_k)\\&\qquad=\left(\int\dots\left(\int\left(\int g(x_1,\dots,x_k)F_1(\,dx_1)\right)F_2(\,dx_2)\right)\dots F_k(\,dx_k)\right).\end{aligned}$$

Ez az azonosság úgy értendő, hogy az azonosság két oldalán lévő integrál illetve szukcesszív integrál egyszerre létezik.

Érdemes külön megfogalmazni ennek az azonosságnak a következő fontos speciális esetét. Ha $g(x_1,\dots,x_k)=g_1(x_1)\cdots g_k(x_k)$ alakú speciális függvények integrálját tekintjük, akkor a következő azonosságot írhatjuk fel:

$$\begin{aligned}&\int g_1(x_1)g_2(x_2)\cdots g_k(x_k)F_1(\,dx_1)F_2(\,dx_2)\dots F_k(\,dx_k)\\&\qquad=\int g(x_1)F_1(\,dx_1)\int g_2(x_2)F_2(\,dx_2)\cdots\int g_k(x_k)F_k(\,dx_k)=\prod_{j=1}^{k}\int g_j(x)F_j(\,dx).\end{aligned}$$

1. megjegyzés. A Fubini tétel valójában egy általánosabb eredmény, mint a most kimondott tétel. A Fubini tétel az itt megfogalmazottakhoz hasonló eredményt mond ki általános (és nemcsak az Euklideszi terekben definiált) szorzatmértékek szerinti integrálokra. Számunkra viszont elegendő csak a fent leírt speciális esetet tekinteni.

2. megjegyzés. Az előbb kimondott tétel és a további eredmények röviden, informálisan úgy foglalhatók össze, hogy mindazok az eredmények, amelyeket független, diszkrét eloszlású valószínűségi változókról tanultunk, érvényesek általános, nem feltétlenül diszkrét eloszlású valószínűségi változókra is.

Az alábbiakban a Fubini tételt fogjuk használni, és azokat a eredményeket, amelyek lehetővé teszik, hogy valószínűségi változók függvényeinek várható értékét kiszámoljuk e valószínűségi változók eloszlásfüggvényének a segítségével. Ezen eredmények felhasználásával bebizonyítok néhány alapvető eredményt független valószínűségi változókról. Ezelőtt egy megjegyzést teszek az itt követett tárgyalásmódról.

3. megjegyzés. Az alábbiakban független valószínűségi változók legfontosabb tulajdonságait ismertetem. Ezeket a tulajdonságokat meg lehet fogalmazni az eloszlások nyelvén is. Íly módon kiderül, hogy ezek a tulajdonságok ekvivalensek valamilyen integrálokra megfogalmazható azonosságokkal, és ezen azonosságok mindegyike a Fubini-tétel következményeként kezelhető.

Tétel független valószínűségi változók szorzatának a várható értékéről. *Legyenek $\xi_1,\dots,\xi_k$ független valószínűségi változók, amelyek mindegyikének létezik várható értéke, azaz $E|\xi_j|<\infty$. Ekkor a $\xi_1\cdots\xi_k$ szorzatnak is létezik várható értéke, és*

$$E\xi_1\cdots\xi_k=E\xi_1\cdots E\xi_k.$$

Megjegyzés: Ezt az eredményt már tárgyaltam az 4. téma ismertetésében a *Tétel független valószínűségi változók szorzatáról* eredményében, de ott csak diszkrét eloszlású valószínűségi változók szorzatát tekintettem.

A tétel bizonyítása. Jelölje $F_j(\cdot)$, $1 \le j \le k$, a ξ_j valószínűségi változó eloszlásfüggvényét, és vezessük be a $g(x_1, \dots, x_k) = x_1 \dots x_k$ függvényt. Ekkor

$$E\xi_1 \cdots \xi_k = Eg(\xi_1, \dots, \xi_k) = \int x_1 \cdots x_k F_1(\,dx_1) \dots F_k(\,dx_k).$$

Továbbá a Fubini tétel szerint

$$\int x_1 \cdots x_k F_1(\,dx_1) \dots F_k(\,dx_k) = \prod_{j=1}^{k} \int x F_j(\,dx).$$

Mivel $E\xi_j = \int x F_j(x)$, ez az azonosság azt jelenti, hogy $E\xi_1 \cdots \xi_k = E\xi_1 \cdots E\xi_k$. Ezenkívül azt is állíthatjuk (felhasználva azt a tényt, hogy a Fubini tétel két oldalán szereplő kifejezés egyszerre értelmes), hogy a Tételben szereplő azonosság két oldala egyszerre értelmes.

Tétel független valószínűségi változók együttes eloszlásáról. *Legyenek $\xi_1, \dots, \xi_k$ független valószínűségi változók, $B_1, \dots, B_k$ a számegyenes Borel mérhető részhalmazai. Ekkor*

$$P(\xi_1 \in B_1, \dots, \xi_k \in B_k) = P(\xi_1 \in B_1) \cdots P(\xi_k \in B_k).$$

Megjegyzés: A függetlenség definiciója csak azt követeli meg, hogy a tételben kimondott azonosság teljesüljön speciális $B_j = (-\infty, x_j)$ alakú halmazokra. A tétel azt állítja, hogy ha ez az azonosság teljesül ezekre a speciális alakú halmazokra, akkor az teljesül minden ,,szép" azaz Borel mérhető halmazra. E tételnek, illetve e tételnek a 4. téma ismertetésében tárgyalt *Lemma független, diszkrét eloszlású valószínűségi változók tulajdonságairól* néven megfogalmazott megfelelőjének egyik következménye az, hogy a diszkrét, valós szám értékű valószínűségi változók függetlenségének korábban ismertetett definiciója megegyezik a függetlenség általános definiciójával ebben a speciális esetben.

Bizonyítás: Legyen $g_j(\cdot)$ a B_j halmaz indikátorfüggvénye, $1 \le j \le k$, azaz legyen $g_j(x) = 1$, ha $x \in B_j$, és $g_j(x) = 0$, ha $x \notin B_j$. Definiáljuk a $g(x_1, \dots, x_k) = g_1(x_1) \cdots g_k(x_k)$ függvényt a k-dimenziós téren. Ekkor a Fubini tétel alapján

$$\begin{aligned} P(\xi_1 \in B_1, \dots, \xi_k \in B_k) &= Eg(\xi_1, \dots, \xi_k) = Eg_1(\xi_1) \cdots Eg_k(\xi_k) \\ &= P(\xi_1 \in B_1) \cdots P(\xi_k \in B_k). \end{aligned}$$

Feladat:

Legyenek $\xi_1,\dots,\xi_n,\xi_{n+1},\dots,\xi_{n+m}$ független valószínűségi változók, $g(x_1,\dots,x_n)$ n-változós (mérhető) függvény, $\eta=g(\xi_1,\dots,\xi_n)$. Mutassuk meg a Fubini tétel segítségével, hogy $\eta,\xi_{n+1},\dots,\xi_{n+m}$ független valószínűségi változók.
Megoldás: Rögzítsünk valamilyen $x_0,x_1,\dots,x_m$ számokat. Definiáljuk ezek segítségével a $h_j(u)=1$, ha $u<x_j$, $h_j(u)=0$, ha $u\ge x$, $1\le j\le m$, függvényeket, valamint legyen $h_0(u_1,\dots,u_n)=1$, ha $g(u_1,\dots,u_n)<x_0$, $h_0(u_1,\dots,u_n)=0$, ha $g(u_1,\dots,u_n)\ge x_0$. Jelölje $F_j(\cdot)$ a ξ_j valószínűségi változó eloszlásfüggvényét, $1\le j\le n+m$. Ekkor a Fubini tétel alapján

$$
\begin{aligned}
&P(\eta<x_0,\xi_{n+1}<x_{n+1},\dots,\xi_{n+m}<x_{n+m})\\
&\quad=\int h_0(u_1,\dots,u_n)h_{n+1}(u_{n+1})\dots h_{n+m}(u_{n+m})F_1(\,du_1)\dots F_{n+m}(\,du_{n+m})\\
&\quad=\int h_0(u_1,\dots,u_n)F_1(\,du_1)\dots F_n(\,du_n)\\
&\qquad\quad\int h_{n+1}(u_{n+1})F_{n+1}(\,du_1)\cdots\int h_{n+m}(u_{n+m})F_{n+m}(\,du_{n+m})\\
&\quad=P(\eta<x_0)P(\xi_{n+1}<x_{n+1})\cdots P(\xi_{n+m}<x_{n+m}),
\end{aligned}
$$

ahonnan következik a feladat állítása.

Független valószínűségi változók összegének a szórásnégyzete.
A diszkrét valószínűségi változókhoz hasonlóan defiálhatjuk általános, nem feltétlenül diszkrét eloszlású valószínűségi változók szórásnégyzetét, (ezt már megtettük), kovarianciáját. Általános valószínűségi változók összegének szórásnégyzetére hasonló formula érvényes mint a diszkrét esetben. Röviden leírom a bizonyítást, bár azt akár el is hagyhatnám arra hivatkozva, hogy e tulajdonságok bizonyításában a diszkrét valószínűségi változók esetében is csak olyan összefüggéseket használtunk, amelyek általános valószínűségi változókra is érvényesek.

Valószínűségi változók kovarianciájának a definiciója. *Legyen ξ és η két valószínűségi változó ugyanazon a valószínűségi mezőn, (amelyekre teljesül az $E\xi^2<\infty$, $E\eta^2<\infty$ feltétel.) A ξ és η valószínűségi változók kovarianciafüggvénye*

$$\operatorname{Cov}(\xi,\eta)=E[(\xi-E\xi)(\eta-E\eta)].$$

Lemma.

$$\operatorname{Cov}(\xi,\eta)=E\xi\eta-E\xi E\eta.$$

Ha ξ és η független valószínűségi változók, akkor $\operatorname{Cov}(\xi,\eta)=0$.

Bizonyítás:

$$
\begin{aligned}
\operatorname{Cov}(\xi,\eta)&=E[(\xi-E\xi)(\eta-E\eta)]=E(\xi\eta-\xi E\eta-\eta E\xi+E\xi E\eta)\\
&=E\xi\eta-E\xi E\eta-E\xi E\eta+E\xi E\eta=E\xi\eta-E\xi E\eta.
\end{aligned}
$$

Ha ξ és η független valószínűségi változók, akkor $E\xi\eta = E\xi E\eta$, ezért $\mathrm{Cov}\,(\xi, \eta) = 0$.

Tétel valószínűségi változók összegének szórásnégyzetéről. *Legyenek ξ_j, $1 \le j \le k$, valószínűségi változók ugyanazon a valószínűségi mezőn, amelyekre teljesül $E\xi_j^2 < \infty$ feltétel. Ekkor*

$$\mathrm{Var}\left(\sum_{j=1}^{k} \xi_j\right) = \sum_{j=1}^{k} \mathrm{Var}\,\xi_j + \sum_{\substack{1\le j,l\le k\\ j\ne l}} \mathrm{Cov}\,(\xi_j, \xi_l) = \sum_{j=1}^{k} \mathrm{Var}\,\xi_j + 2\sum_{j=1}^{k-1}\sum_{l=j+1}^{k} \mathrm{Cov}\,(\xi_j, \xi_l).$$

Speciálisan, ha a ξ_j valószínűségi változók függetlenek, akkor

$$\mathrm{Var}\left(\sum_{j=1}^{k} \xi_j\right) = \sum_{j=1}^{k} \mathrm{Var}\,\xi_j.$$

Bizonyítás. Vezessük be a $\bar{\xi}_j = \xi_j - E\xi_j$ valószínűségi változókat. Ekkor

$$\begin{aligned}\mathrm{Var}\left(\sum_{j=1}^{k} \xi_j\right) &= E\left(\sum_{j=1}^{k} \bar{\xi}_j\right)^2 = E\left(\sum_{j=1}^{k} \bar{\xi}_j^2 + \sum_{\substack{1\le j,l\le k\\ j\ne l}} \bar{\xi}_j, \bar{\xi}_l\right)\\ &= \sum_{j=1}^{k} E\bar{\xi}_j^2 + \sum_{\substack{1\le j,l\le k\\ j\ne l}} E\bar{\xi}_j\bar{\xi}_l = \sum_{j=1}^{k} \mathrm{Var}\,\xi_j + \sum_{\substack{1\le j,l\le k\\ j\ne l}} \mathrm{Cov}\,(\xi_j, \xi_l).\end{aligned}$$

Ha a ξ_j valószínűségi változók függetlenek, akkor $\mathrm{Cov}\,(\xi_j, \xi_l) = 0$, ezért

$$\mathrm{Var}\left(\sum_{j=1}^{k} \xi_j\right) = \sum_{j=1}^{k} \mathrm{Var}\,\xi_j.$$

Következmény. *Legyenek ξ_j, $1 \le j \le k$, valószínűségi változók ugyanazon a valószínűségi mezőn, amelyekre teljesül $E\xi_j^2 < \infty$ feltétel, és legyenek $c_1, \dots, c_n$ tetszőleges valós számok. Ekkor*

$$\begin{aligned}\mathrm{Var}\left(\sum_{j=1}^{k} c_j\xi_j\right) &= \sum_{j=1}^{k} c_j^2\mathrm{Var}\,\xi_j + \sum_{\substack{1\le j,l\le k\\ j\ne l}} c_jc_l\mathrm{Cov}\,(\xi_j, \xi_l)\\ &= \sum_{j=1}^{k} c_j^2\mathrm{Var}\,\xi_j + 2\sum_{j=1}^{k-1}\sum_{l=j+1}^{k} c_jc_l\mathrm{Cov}\,(\xi_j, \xi_l).\end{aligned}$$

Legyenek ξ_j, $j = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók, és legyen $E\xi_1^2 < \infty$. Becsüljük meg a Csebisev egyenlőtlenség segítségével a

$$P\left(\left|\frac{1}{n}\sum_{j=1}^{n}(\xi_j - E\xi_j)\right| > \varepsilon\right)$$

valószínűségeket minden $n = 1, 2, \dots$ és $\varepsilon > 0$ számra. Látni fogjuk, hogy ebből a becslésből adódik a valószínűségszámítás egyik fontos eredménye, a nagy számok (gyenge) törvénye.

$$\begin{aligned}
P\left(\left|\frac{1}{n}\sum_{j=1}^{n}(\xi_j - E\xi_j)\right| > \varepsilon\right) &= P\left(\left|\frac{1}{n}\sum_{j=1}^{n}(\xi_j - E\xi_j)\right| > \varepsilon\right) \\
&= P\left(\left(\sum_{j=1}^{n}(\xi_j - E\xi_j)\right)^2 > n^2\varepsilon^2\right) \le \frac{E\left(\sum\limits_{j=1}^{n}(\xi_j - E\xi_j)\right)^2}{n^2\varepsilon^2} \\
&= \frac{\operatorname{Var}\left(\sum\limits_{j=1}^{n}\xi_j\right)}{n^2\varepsilon^2} = \frac{\sum\limits_{j=1}^{n}\operatorname{Var}\xi_j}{n^2\varepsilon^2} = \frac{\operatorname{Var}\xi_1}{n\varepsilon^2}.
\end{aligned}$$

Kiegészítés: A Lebesgue integrál fogalmáról.

A Lebesgue integrál természetes definiciójában ezt az integrált először a legegyszerűbb függvények esetében definiáljuk természetes módon, majd ezt a definiciót kiterjesztjük általános függvényekre folytonossági meggondolások alapján. Ezt a következőképp tehetjük. Legyen adva egy $(X, \mathcal{A}, \mu)$ mértéktér, azaz egy X alaphalmaz, kijelöljük annak bizonyos részhalmazait, amelyek σ-algebrát alkotnak, és μ σ-véges mérték az $\mathcal{A}$ σ-algebrán. (Egy mérték, azaz egy nem negatív értékű σ-additív halmazfüggvény akkor σ-véges, ha létezik az X alaphalmaz $X = \bigcup\limits_{j=1}^{\infty} B_j$, felbontása megszámlálhatóan sok olyan B_j halmazra uniójára, amelyekre $\mu(B_j) < \infty$.) Az $(X, \mathcal{A}, \mu)$ téren értelmezett (mérhető) $f(\cdot)$ függvényeknek akkor definiáljuk a μ mérték szerinti integrálját, ha azok teljesítenek bizonyos feltételeket, és ezeket az integrálokat $\int f(x)\mu(\,dx)$ formában jelöljük. Jegyezzük meg, hogy egy valószínűségi mező speciális $(X, \mathcal{A}, \mu)$ mértéktér. Bizonyos vizsgálatokban hasznos, ha figyelmünket nem korlátozzuk a valószínűségi mezőkön definiált függvények (azaz valószínűségi változók) integráljára.

Egy $f(\cdot)$ függvényt elemi függvénynek nevezünk, ha csak véges sok különböző értéket vesz fel, azaz létezik véges sok olyan $z_1, x_2, \dots, z_k$ szám úgy, hogy $\bigcup\limits_{j=1}^{k} \{x\colon\ f(x) = z\} = \Omega$. Ennek az elemi függvénynek az integrálját az

$$\int f(x)\mu(\,dx) = \sum_{j=1}^{k} z_j \mu\left(\{x\colon\ f(x) = z_j\}\right)$$

összeg definiálja. (Ez hasonló a diszkrét eloszlású függvények integráljának definiciójához. Az egyetlen különbség az, hogy itt csak véges sok értékeket felvevő függvényeket tekintünk, míg ott megengedtük azt is, hogy egy függvény megszámlálhatóan sok értéket vegyen fel. Definiálhattuk volna az elemi függvényeket is általánosbban, megengedhettük volna, hogy azok megszámlálható sok értéket vegyenek fel, de az itt tekintett definició technikailag kényelmesebb, és ennek használatával semmit sem veszítünk.) Ha adva van egy nem-negatív $f(x)$ függvény az az $(X, \mathcal{A}, \mu)$ mértéktéren, akkor azt közelíthetjük elemi $f_n(x)$ elemi függvényekkel például a következő módon. Legyen $f_n(x) = k2^{-n}$, ha $k2^{-n} \le x \le (k+1)2^{-n}$, $k = 0, 1, \dots, 2^{2n}$, és definiáljuk az $f(x)$ függvény integrálját, az $\int f(x)\mu(\,dx) = \lim\limits_{n\to\infty} \int f_n(x)\mu(\,dx)$ képlet segítségével. Egy általános (mérhető) $f(x)$ függvény esetében definiáljuk annak pozitív és negatív részét, mint $f_+(x) = \max(0, f(x))$, $f_-(x) = \min(0, f(x))$, és $\int f(x)\mu(\,dx) = \int f_+(x)\mu(\,dx) - \int(-f_-(x))\mu(\,dx)$. Be lehet látni, hogy az előbb definiált integrál értelmes (például a felírt limesz valóban létezik, ha $\int |f(x)|\mu(\,dx) < \infty$).

Felmerülhet a kérdés, hogy mi a kapcsolat az így definiált integrál és az analízisben tanult Riemann és Riemann–Stieltjes integrál között, és mik a Lebesgue integrál legfontosabb tulajdonságai.

A hagyományos analízis oktatásban az úgynevezett Riemann integrált vezetik be először. Egy $[a, b]$ intervallumon értelmezett $f(u)$ függvény $\int_a^b f(u)\,du$ integrálját úgy definiálják, hogy először alkalmas közelítő összegeket vezetnek be a következő módon. Felosztják az $[a, b]$ intervallumot $a = x_0 < x_1 < \cdots < x_n = b$ osztópontok segítségével, amelyekre $\sup\limits_{1\le k\le n} (x_k - x_{k-1})$ kicsi, mindegyik $[x_{k-1}, x_k]$ intervallumban kijelölnek egy $\xi_k \in [x_{k-1}, x_k]$ pontot, és a $\sum\limits_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$ közelítő összegeket tekintik. Az $\int_a^b f(u)\,du$ integrált úgy definiálják mint ezen integrálközelítő összegek határértékét, ha $\sup\limits_{1\le k\le n} (x_k - x_{k-1})$ nullához tart.

Egy f függvénynek valamely az $[a, b]$ intervallumon megadott $F(u)$ függvény szerinti $\int_a^b f(u)F(\,du)$ Riemann–Stieltjes intergrálját hasonlóan definiálják. A különbség az, hogy most a közelítő összegekben az $x_k - x_{k-1}$ megváltozásokat az $F(\cdot)$ függvény $F(x_k) - F(x_{k-1})$ megváltozásaival helyettesítik, azaz a $\sum\limits_{k=1}^{n} f(\xi_k)\,(F(x_k) - F(x_{k-1}))$ közelítőösszegeket, illetve azok határértékét tekintik. Ha az $F(x)$ függvény nem folytonos, akkor bizonyos technikai problémákat részletesebben tisztázni kell a fenti közelítő összegek definiciójában, ezek kidolgozásától azonban itt eltekintek.

A fenti definiciót a következő módon is interpretálhatjuk. Nevezzünk egy az $[a, b]$ intervallumon értelmezett $f(x)$ függvényt egyszerű vagy lépcsős függvénynek, ha létezik az $[a, b]$ intervallumnak olyan $a = x_0 < x_1 < \cdots < x_n = b$ felosztása, amelyre az $f(x)$ függvény konstans minden egyes $[x_k, x_{k+1})$ intervallumon, $0 \le k < n$. Az ilyen lépcsős függvények integrálját az $\int_a^b f(x)\,dx = \sum\limits_{k=1}^{n} f(x_k)(x_k - x_{k-1})$, és $\int_a^b f(x)F(\,dx) = \sum\limits_{k=1}^{n} f(x_k)\,(F(x_k) - F(x_{k-1}))$ képletekkel definiáljuk. Az általánosabb függvények integrálját úgy definiáljuk, hogy közelítjük őket a fent leírt módon lépcsős függvényekkel, és a függvény integrálja a megfelelő lépcsős függvények integráljainak a limesze.

Tekintsük a számegyenest, rajta a $\mathcal{B}$ Borel σ-algebrát, és azon a Lebesque mértéket. Vegyük észre, hogy mind a Riemann mind a Lebesue integrál esetében először bizonyos elemi függvényeknek definiáljuk az integrálját alkalmas közelítő összegek segítségével, majd a definiciókat alkalmas határátmenet segítségével kiterjesztjük. Ezért nem meglepő, hogy abban az esetben, amikor mind a két integrál létezik, akkor azok megegyeznek.

Legyen F valamely monoton növekvő függvény az $[a,b)$ intervallumon. A Lebesgue integrálhoz hasonlóan az $\int_a^b f(u)F(\,du)$ Riemann–Stieltjes integrálnak megfelelő Lebesgue–Stieltjes integrált is definiálhatunk. Ebben az esetben is az $(X,\mathcal{A},\mu)$ teret úgy definiáljuk, hogy az X halmaz a számegyenes, (vagy esetleg annak egy részintervalluma), $\mathcal{A}$ a számegyenesen vagy az intervallumon bevezetett Borel σ-algebra. A definició teljessé tételéhez definiálni kell a μ mértéket is. Ezt az $[a,b)$ intervallumon vagy számegyenesen definiált F monoton növekvő függvény határozza meg.

A mértékelmélet egyik eredménye szerint az F függvény egyértelműen meghatároz az $[a,b)$ intervallum Borel–mérhető halmazainak σ-algebráján egy olyan σ-additív μ_F az F függvény Stieltjes mértékének nevezett halmazfüggvényt, amelyre $\mu_F([u,v)) = F(v)-F(u)$ minden $a\le u<v\le b$ (u,v) rárra. (Egy részletesebb tárgyalásban ezen állítás megfogalmazásában figyelmesebbnek kell lennünk. Külön vizsgálni kell azt az esetet, amikor az $F(\cdot)$ függvény a tekintett $[u,v)$ intervallum u vagy v végpontjában nem folytonos, azaz $\lim\limits_{u_n\to u-0} F(u_n) \neq \lim\limits_{u_n\to u+0} F(u_n)$, illetve, ha a v pontban teljesül hasonló reláció. Kiderült, hogy elég csak balról folytonos (monoton) függvényeket tekintenünk, azaz olyan $F(\cdot)$ függvényeket, amelyekre $\lim\limits_{x_n\to x-0} F_n(x_n) = F(x)$ minden x számra. Az eloszlásfüggvények teljesítik ezt a tulajdonságot. Ezen megszorítás esetén minden (u,v), $u<v$, párra érvényes az $\mu_F([u,v)) = F(v)-F(u)$ azonosság.) A Lebesgue–Stieltjes integrál definiciójának legnehezebb lépése annak megmutatása, hogy ilyen módon egy mértéket azaz egy σ-additív μ_F halmazfüggvényt definiálunk. Az $\int_a^b f(u)F(\,du)$ Lebesue–Stieltjes integrál az $\int f(x)\mu(\,dx)$ Lebesgue integrál, ha az $[a,b]$ intervallumon és a rajta tekintett Borel σ-algebrán az F függvény által meghatározott μ_F mértéket választjuk μ a mértéknek. A Riemann-Stieltjes és Lebesgue–Stieltjes integrálok megegyznek, ha mind a két integrál létezik.

Láttuk, hogy a Lebesgue és Riemann integrálok definiciójának a logikája nagyon hasonló. Mind a két esetben először szép, egyszerű esetekben definiáljuk az integrált, majd határátmenet segítségével ezt a definiciót kiterjesztjük. A Lebesgue integrál definicióját nem nehezebb megérteni mint a Riemann integrálét. Annak fő oka, hogy az oktatásban a Riemann integrálok tárgyalására helyezik a hangsúlyt — bizonyos tudománytörténeti okokon kívül — az, hogy a Lebesgue integrálok felépítése szükségessé teszi σ-additív mértékek használatát, és ez elkerülhetetlenné tesz bizonyos nem-triviális mértékelméleti vizsgálatokat. A Lebesgue integrálok definiciójában, amikor egy halmaz nívóhalmazait tekintjük nem követeljük meg, hogy ezeknek a nívóhalmazoknak szép geometriai struktúrájuk legyen. Ennek következtében sokkal gazdagabb azon függvények osztálya amelyeknek definiálni tudjuk a Lebesgue integrálját. Ugyanis az integrál definiciójában szükséges határátmeneteketet a Lebesgue integrál definiciójában könnyebb végrehajtani mint a Riemann integrálében. Sőt a Lebesgue integrál definiciója — szemben a Riemann integráléval — nem kötődik a számegyenes geometriájához.

A Lebesgue integrál rendelkezik a Riemann integrálról tanult legfontosabb tulajdonságokkal. Így például lineáris operátor az integrálható függvények terén, azaz tetszőleges f és g (integrálható) függvényre valamint a és b számokra

$$\int (af(x) + bg(x))\mu(\,dx) = a\int f(x)\mu(\,dx) + b\int g(x)\mu(\,dx).$$

Bizonyos fontos, itt nem ismertetett eredmények (Lebesgue tétel, Beppo–Levy tétel Fatou lemma) arról szólnak, hogy konvergens függvénysorok esetén mikor cserélhető fel a limesz és integrálás sorrendje, azaz milyen feltételek mellett érvényes a

$$\lim \int f_n(x)\mu(\,dx) = \int \lim f_n(x)\mu(\,dx)$$

azonosság. A Lebesgue integrálok egy fontos tulajdonsága az ebben a jegyzetben is megfogalmazott Fubini tétel, amely szerint egy szorzatmérték szerinti kétváltozós függvény integrálja szukcesszív integrálással is kiszámolható. Láttuk, hogy ez fontos szerepet játszik független valószínűségi változók vizsgálatában.

Végül megjegyzem, hogy a várható értéknek az eloszlásfüggvény segítségével való kiszámolásáról szóló tételnek illetve ezen eredmény általánosításának a bizonyítása viszonylag egyszerű, ha jól megértjük a Lebesgue integrál definicióját. A Fontos Tétel állítása szemléletesen, de kissé pongyolán fogalmazva azt jelenti, hogy egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált $\int g(\xi(\omega))\,dP(\omega)$ integrál 'átmásolható' egy $\int g(x)\mu_F(\,dx)$ integrállá az $(R^1, \mathcal{B}, \mu_F)$ valószínűségi mezőn, ahol μ_F a ξ valószínűségi változó eloszlása által meghatározott μ_F Lebesgue–Stieltjes mérték. Ez az állítás könnyen látható $g(x)$ lépcsősfüggvények esetében. Általános (mérhető) $g(x)$ függvények jól közelíthetőek lépcsős függvényekkel, és ilyen módon az $\int g(\xi(\omega))\,dP(\omega)$ Lebesgue integrál olyan közelítő összegeit kapjuk, amelyek kifejezhetőek a ξ valószínűségi változó $F(x)$ eloszlásfüggvényét tartalmazó összegek segítségével. Ez utóbbi összegek természetes közelítőösszegei az $\int g(x)F(\,dx)$ Lebesque–Stieltjes integrálnak. Ezután alkalmas határátmenet segítségével megkaphatjuk az említett tételnek, illetve általánosításának a bizonyítását.