

Véletlen folyamatok

Ladóczki Bence

Április 19, 2024

1 Bevezetés

Véletlen (sztochasztikus) folyamatnak nevezzük azon empirikus folyamatok matematikai absztrakcióját, amit valószínűségi törvények vezérelnek. A jelen kéziratban a véletlen folyamatok elméletének csupán egy részhalmazával foglalkozunk. A megértéshez szükség van alapvető valószínűség számítási előismeretekre.

A valószínűség számítás a mértékelmélet egy speciális tárgyalásmódja, annak egy alkalmazott ága. Elviekben van arra is lehetőség, hogy a minta-sorokat és az azokon értelmezett függvényeket a valószínűségi számítástól függetlenül tárgyaljuk és hagyományos sorokkal valamint függvényekkel foglalkozunk és a valószínűség számításra mint az eloszlások tanulmányozására tekintünk. Egy időben a matematika csak így tudta például a nagy számok erős törvényét teljesen egzakt módon értelmezni. Már azonban tudjuk, hogy erre nincs szükség és ez sok helyen túlságosan lerövidítené az anyagot és elferdítené a valóságot.

A valószínűség számítást az teszi egyedivé a matematikai területek között, hogy a legegyszerűbb törvényszerűségei is meglehetősen bonyolult matematikai törvényszerűségekből következnek, viszont a bonyolultabb eredményeket is ettől függetlenül ismerik és alkalmazzák a gyakorlatban (gondoljunk például a statisztikusokra). A valószínűség számítás valóban egzakt tárgyalásmódja ellen gyakran felhozzák azt az érvet, miszerint nincs szükség ilyen szinten belemerülni a részletekbe. A könyv végére mindenki hozza meg magának a saját ítéletét.

2024. április 14.
Budapest

2 Bevezetés és Valószínűség Elméleti Alapok

A jelen kéziratban bonyolult matematikai módszerek tárházát használjuk, ettől függetlenül azonban mindig a valószínűség számítás nyelvénél segítségül hívjuk majd ki az eredményeket. Remélem, hogy azon specialisták számára akik tisztában vannak vele, hogy hogyan kell mondjuk feltételes valószínűséget vagy várható értéket számolni nem fog problémát okozni a megértés. Ebben a fejezetben a legfontosabb matematikai alapokkal ismerkedünk meg.

Amikor magasabb szintű valószínűség számítással foglalkozó könyvet szeretnénk megírni – mivel a valószínűség számítás a mértékelmélet egy ága – két választás előtt áll a szerző. Vagy hagy egy fejezetet a mértékelméletnek, vagy feltételezi, hogy az olvasó a rendelkezésre álló egyéb mértékelméleti szakkönyvekből (előadásokból) már elsajátította a szükséges előismereteket. Praktikussági okokra hivatkozván a második opciót választom, és azt tartom szem előtt, hogy itt egy olyan kézirat készül, amit a mértékelmélet alapos ismerete nélkül is lehet abszolválni. A valószínűség számítás alapvető összefüggéseit, az abban használt módszereket azonban ismertnek feltételezem.

2.1 Milyen Térben Dolgozunk?

Amióta a matematikusok a valószínűség számításra egzakt tudományra tekintenek az olyan fogalmakra mint például a *megjelenés*, *esemény*, *urna*, *kocka* és hasonló nincs többé szükség, amikor valószínűség számításról beszélünk. Ettől függetlenül, ha továbbra is használjuk ezeket, akkor az intuíció a segítségünkre lesz az elkövetkezendő diszkusziók során és analitikus módszereket dolgozhatunk majd ki, valamint új kutatási irányok nyílhatnak meg. Főként emiatt a mai napig gyakran a szóhasználat megmarad még akkor is ha tisztán elméleti valószínűség számításról beszélnek a matematikusok is.

A valószínűség számítás alkalmazása során a *valószínűséget* események bekövetkezésével azonosítják. Olyan kérdéseket szoktunk feltenni, hogy például mi annak a valószínűsége, hogy 5-öt dobunk egy dobókockával? Vagy ha egy részecskre Brown mozgást végez, akkor mi lesz az x irányú elmozdulás centiméterben kifejezett mérőszámának valószínűségi eloszlása t idő múlva ($x(t)$)? Az ilyen eseményekhez rendeljük a valószínűséget. Ahhoz például, hogy 5-öt dobunk egy dobókockával az $\frac{1}{6}$ valószínűséget rendeljük hozzá. Ahhoz, hogy páros számot dobunk, az $\frac{1}{2}$ valószínűséget rendeljük hozzá. Amikor Brown mozgásról beszélünk, akkor az $x(3) > 7$ eseményhez a 9. fejezetben ismertetett szabályok szerint kell valószínűséget rendelni. A kocka esetében a következő módon lehet a matematika nyelvén megfogalmazni a jelenséget. Az összes lehetséges kimenetelhez hozzárendeljük az egész számokat: $1 \dots 6$ és ezekhez a számokhoz az $\frac{1}{6}$ valószínűséget rendeljük hozzá. Ezek tetszőleges osztályához pedig az $\frac{n}{6}$ valószínűséget rendeljük hozzá. Az például hogy páros számot kapunk csak annyit jelent, hogy $n = 3$ -ra kiértékeljük a formulát és $\frac{3}{6} = \frac{1}{2}$ kapunk valószínűségnek. Ezt a gondolatmenetet elvégezni a Brown mozgás esetére kicsivel bonyolultabb és emiatt ezt most nem tesszük meg, viszont megjegyzem itt is csupán arról van szó, hogy események matematikai absztrakciójához számokat kell rendelni. Ezentúl az események matematikai absztrakciójának egy *ponthalmazt* választunk.

A valószínűség számítás elmélete nem más mind a különböző terek mérhetőségi tulajdonságainak vizsgálata, valamint az ezeken a tereken definiált mérhető függvények analízise. A valószínűség számítás alkalmazási területeire való tekintettel hívhatjuk ezeket a tereket *mintavételi* tereknek, a mérhető halmazokat pedig *eseményeknek*, ez a megközelítés azonban nem mindig állja meg a helyét.

Most pedig egy következetes matematikai tárgyalásmód kerül bevezetésre. Feltesszük tehát, hogy egy Ω alaptérrel van dolgunk és hogy ezen a téren pon-

thalmazok vannak. Ezeket a halmazokat hívjuk majd *mérhető halmazoknak*. Feltesszük, hogy a mérhető halmazok osztálya Borel mezőt alkot, továbbá azt is megköveteljük, hogy legyen egy $\mathbf{P}\{\cdot\}$ függvény, amely minden mérhető halmazra definiálva van. Ezt hívjuk *valószínűségi mértéknek*, ami azt jelenti, hogy $\mathbf{P}\{\cdot\}$ értelmezhető minden mérhető halmazra, nem negatív és tökéletesen additív a teljes téren pedig 1-et vesz fel. A $\mathbf{P}\{\Lambda\}$ számot Λ *valószínűségének* vagy *mértékének* nevezzük. Egy ω függvény mérhetőségét hasonló módon kell megadni a mérhető halmazokkal valamint a valószínűségi mérték szerinti integrállal kifejezve. Ez az egy mérhető ϕ függvény mérhető Λ halmazok feletti integrálját jelenti:

$$\int_{\Lambda} \phi(\omega) d\mathbf{P} \quad \text{vagy} \quad \int_{\Lambda} \phi d\mathbf{P} \quad (1)$$

Egy tulajdonságra azt mondjuk hogy *majdnem mindenhol* igaz, ha Ω minden pontjára igaz kiveéve a 0 valószínűségű pontokat. Erre azt is mondhatjuk, hogy majdnem minden ω pontban igaz, vagy hogy 1 valószínűséggel igaz.

Nézzük például azt az esetet, amikor egyszer dobunk a dobókockával. Ebben az esetben az Ω tér az $1 \dots 6$ számokból áll. Egy j pontot azzal az eseménnyel azonosítjuk, amikor j -t dobunk a kockával. Ebben az esetben minden Ω halmaz mérhető és minden halmazhoz a benne lévő pontok számának az $\frac{1}{6}$ -át rendeljük hozzá. Így joggal beszélhetünk egy mintavételi térről, hiszen a tér pontjai és a lehetséges események között egyértelmű megfeleltetést tudunk felállítani. Most tekintsük ugyanezen kísérlet egy más matematikai modelljét. Ezen alkalommal az Ω térbe vegyük bele az összes valós számot és a kockával való dobásból származó eseményekhez rendeljük azokat a számokat a térből, mint az előbb. Más szabályt ne állítsunk fel. Ebben az esetben is minden Ω pont halmaz mérhető és minden halmaz mértéke $\frac{1}{6}$ -a annak, hogy hány egész szám kerül bele az $1 \dots 6$ intervallumból. A matematikai modell, amit így kapunk majdnem teljesen ugyan az, mint az előző esetben, azonban a mintavételi tér elnevezés kicsit kevésbé praktikus, hiszen így már nehezebb megfeleltetni minden pontot egy valós eseménnyel. Végezetül tekintsünk egy harmadik felállást, amikor is az Ω tér az $[1, 7)$ félig nyílt intervallumon található összes számból áll. Most feltessük meg a $[j, j+1)$ intervallumot annak az eseménynek, amikor j -t dobunk. Ebben az esetben a mérhető pontok halmazán intervallumokat ($[j, j+1)$), vagy ezek unióját értjük. És bármely mérhető halmaz mértéke megegyezik az intervallumok darabszámának $\frac{1}{6}$ -ával. Ezt a modellt nyilván ugyanolyan jól lehet használni mint az előző kettőt, azonban a mintatér elnevezés itt sem állja meg a helyét. Nyilván nem utasíthatjuk vissza ezt az okfejtést, mondván hogy a modell túl bonyolult. A feleslegesen bonyolult modellek még a gyakorlati alkalmazásokban sem söpörhetőek le az asztalról. A matematikai modell felállításának módja attól függ, hogy milyen paramétereket veszünk alapul. Az első modellben konkrétan a dobások kimenetelét vettük alapul. Azonban ha a dobókockára úgy tekintünk, mint egy fizikai objektum, ami az descartesi térben mozog valamilyen sebességgel, akkor a 2 oldalának hely és sebesség vektorai egyértelmű meghatározzák a dobás kimenetelét (12 paraméter). Az Ω térnek tehát természetes módon választhatnánk a 12 darab kiindulási hely és

sebesség komponenseket. Az Ω tér egy pontja tehát egyértelműen meg fogja határozni a dobás kimenetelét. Jelöljük Λ_j -vel azokat a pontokat a térben, amik végül majd j dobást eredményeznek. Az előző gondolatmenet alapján minden Λ_j halmazhoz $\frac{1}{6}$ valószínűséget kell rendelnünk. Csupán ennyi tudásra van szükségünk az Ω térről, ha csak a dobás kimenetelére vagyunk kíváncsiak. Ez a modell nagyon hasonlít az előbbiekben ismertetett harmadik modellre. Arról van szó tehát, hogy mind az elméleti, mind a gyakorlati tárgyalásmódok során, megválaszthatjuk az Ω teret olyan kritériumok alapján, ami egyszerűsíti a kezelést (hogy lehessen mondjuk mintavételi térnek nevezni), azonban ez a kritérium pont az azt követő levezetésekben fog elveszni, amikor is valószínűségi eloszlásokat hozunk ki a kezdeti felállásból. Pontosan emiatt az Ω térrel kapcsolatban nem kötünk ki semmi olyat, ami a $\mathbf{P}\{\Lambda\}$ halmazfüggvény létezéséből implicit módon ne következne. A $\mathbf{P}\{\Lambda\}$ létezéséből ugyan következik, hogy az Ω tér nem lehet üres, de ennél többet nem várunk el az Ω tértől. A továbbiakban az Ω tér alatt egy absztrakt teret értünk. Amikor konkretizáljuk a teret, akkor az Ω térnek az $0 \leq x \leq 1$ intervallumot választjuk. Ez lehet továbbá az összes véges vagy végtelen valós értékű függvénye a t változónak, ahol $-\infty < t < \infty$, vagy a véges sík is akár. Ami valamelyest a valószínűség számítással kapcsolatos, az az, hogy a $\mathbf{P}\{\Lambda\}$ függvénytől megköveteljük a normalizálási feltételt, avagy $\mathbf{P}\{\Lambda\} = 1$. Ha elrugaszkodunk a valószínűség számítástól, akkor nem is mindig kell megkövetelni, hogy a halmazfüggvények véges értékek legyenek, ha pedig végesek, akkor arra sincs feltétlenül szükség, hogy a $\mathbf{P}\{\Lambda\} = 1$ feltétel teljesüljön.

Foglalkozzunk a továbbiakban megint a dobókockával való dobással, hiszen ezen a példán keresztül érthetjük meg miért is olyan praktikus az alaptérként szorzatteret választani. Ha végtelenszer eldobjuk a dobókockát a következő módon könnyen definiálhatunk alapteret. Az Ω térben legyen minden ω pont egy sorozat: ξ_1, ξ_2, \dots , ahol minden ξ_i egy egész lehet az $1 \dots 6$ intervallumból. Az Ω tér egy pontja tehát lehet egy olyan sorozat, ami akkor valósul meg, amikor a dobókockát végtelenszer eldobjuk. Ebben az esetben azonosíthatjuk az összes olyan sor osztályát, ami mondjuk j -vel kezdődik. Ezen sorok osztályához az $\frac{1}{6}$ valószínűséget rendelhetjük hozzá és így az egyszeri dobókocka dobás egy másik matematikai modelljét kapjuk meg. Az ilyen jellegű megválasztásban az a szerencsés, hogy ezt a modellt ki lehet bővíteni tetszőleges számú dobásra ha az osztályokat a j_1, \dots, j_n -el kezdődő sorokból képezzük és ezekhez rendeljük azt az eseményt, miszerint az első n dobás során pontosan j_1, \dots, j_n -et kapunk. Ezen sorok osztályához (Ω halmaz) az $(\frac{1}{6})^n$ valószínűséget rendelhetjük hozzá. Valamennyire igazuk lehet azoknak akik azt, mondják, hogy a végtelen darabszámú kísérletet a gyakorlatban sosem végzünk el, azonban a gyakran még a legegyszerűbb esetben sem tudjuk kikerülni a végtelen sorok tárgyalását. Például ha analizálni szeretnénk azt az eseményt, amikor valami először következik be, mondjuk először hatot dobunk a kockával, akkor ezt csak végtelen sorokon értelmezett mintatérre tudjuk megtenni, hiszen azt a számot, ami ahhoz kell, hogy valami megtörténjen nem tudjuk alulról becsülni mielőtt elvégeznénk a kísérletet. Ez a szám az Ω téren egy nem korlátolt függvény. Az előző példában az Ω tér szorzattér és végtelen sok faktortere van, ezek pedig

mind hat pontot tartalmaznak.

Az Ω téren értelmezhetünk feltételeket, ezeket C -vel jelöljük és az $\{C\}$ jelölést használjuk majd azokra a pontokra, amik kielégíti az Ω téren a C feltételt. Például ha X lineáris halmaz és ha x egy ω függvény, akkor $\{x(\omega) \in X\}$ egy olyan ω halmaz, amelyen $x(\omega)$ egy szám az X halmazból.

Még nem tettük fel, hogy az alap valószínűségi mérték teljes. Arról van szó, hogy ha Λ_0 mérhető és mértéke 0 és ha $\Lambda \supset \Lambda_0$ részhalmaza, attól még nem kell, hogy Λ mérhető legyen. Ha ezt le szeretnénk fordítani a valószínűség számítás nyelvére, azt mondhatjuk, hogy ha van két eseményünk, Λ és Λ_0 és Λ bekövetkeztéből következik Λ_0 , akkor:

$$\mathbf{P}\{\Lambda_0\} = 0 \quad , \quad (2)$$

és ettől függetlenül Λ valószínűsége nem biztos, hogy értelmezett. Nyilván, ha $\mathbf{P}\{\Lambda\}$ értelmezett, akkor az 0. Egy matematikus számára ez a tulajdonság zavaró lehet ha meg szeretné magának tartani a valószínűség intuitív értelmezését. Itt tehát két választás előtt állunk, vagy módosítjuk intuíciónkat vagy megváltoztatjuk a matematikát. A választás nem túl fontos, azonban ha nagyon ragaszkodunk az intuíciónkhoz, akkor könnyen átdefiniálhatjuk fogalmainkat. Valójában a valószínűségi mértéket könnyedén teljessé lehet tenni ha az ω halmazok osztályát némileg kibővítjük. Ettől függetlenül, most azonban mégis tegyük inkább fel, hogy a valószínűségi mérték, amivel ezentúl dolgozni fogunk teljes.

2.2 Valószínűségi Változók és Eloszlások

Egy valós x függvényt egy ω ponttérén valós *valószínűségi változónak* nevezzük ha az ω halmazokon létezik valószínűségi mérték és egy tetszőleges λ valós számhoz az $x(\omega) \leq \lambda$ egyenlőtlenség bekorlátozza az ω halmazt és így is mérhető halmazt kapunk. Ebből következő módon:

$$F(\lambda) = \mathbf{P}\{x(\omega) \leq \lambda\} \quad (3)$$

létezik tetszőleges valós λ -ra. A matematika nyelvét használva ezt úgy is mondhatjuk, hogy egy (valós) valószínűségi változó tulajdonképp nem más, mint egy (valós) mérhető függvény. Egy komplex valószínűségi változó pedig egy ω függvény, aminek a valós és a képzetes részei mérhetőek. Mostantól, ha mást nem állítunk, mindig fel fogjuk tenni, hogy ha esetleg több valószínűségi változóról lenne szó, akkor a valószínűségi változókat mindig ugyanazon ω téren fogjuk értelmezni. A matematikusok egy része már a mértékelmélet megjelenése előtt is foglalkozott valószínűségi változókkal, viszont akkor még nem volt tudatosítva hogy tulajdonképp mérhető függvények tulajdonságait vizsgálják. Az akkori tisztán valószínűség számítási nyelvezetet át lehet fordítani a mai mértékelméleti nyelvezetre. Ezek mellett meg fogjuk tartani a klasszikus valószínűség számítási terminusokat is, hiszen az alkalmazások során ez segítség lehet. Az analízisben gyakran praktikus lehet ugyanazt az a jelölést használni a függvényre, valamint a felvett értékére az értelmezési tartományában és ez a kétértelműség gyakran ki is fizetődik. Sok esetben azonban ennél pontosabban kell eljárunk. A

függvényeket általában egy betűvel fogjuk jelölni és más jelölést fogunk alkalmazni a függvény értékére az adott pontban. Így tehát az $x(\omega)$ alatt az x függvény értékét értjük az ω pontban. Gyakran, amikor célszerű, akkor az x függvényt az $x(\cdot)$ jellel jelöljük majd.

Az (3)-ben definiált F függvényt az x valószínűségi változó *eloszlásfüggvényének* hívjuk. Ez a függvény monoton, nem csökkenő és jobbról folytonos és igazak rá a következők:

$$\lim_{\lambda \rightarrow -\infty} F(\lambda) = 0, \quad \lim_{\lambda \rightarrow \infty} F(\lambda) = 1. \quad (4)$$

Bármely olyan függvényt, ami eleget tesz ezen feltételeknek eloszlásfüggvényének fogunk hívni. Az eloszlásfüggvény eloszlást definiál és ez A halmazok valószínűségi mértéke lesz:

$$\int_A dF(\lambda). \quad (5)$$

Ez egy hagyományos Lebesgue-Stieltjes mérték, amit F definiál. Ha pedig F eloszlásfüggvény és valamilyen Lebesgue mérhető függvényre teljesül a következő:

$$F(\lambda) = \int_{-\infty}^{\lambda} f(\mu) d\mu, \quad -\infty < \lambda < \infty, \quad (6)$$

akkor f -et az F -hez tartozó sűrűségfüggvénynek hívjuk. Majdnem minden λ -ra a Lebesgue mértékben teljesül az is, hogy $F'(\lambda) = f(\lambda)$. Amikor tehát azt mondjuk, hogy F -hez tartozik sűrűségfüggvény, akkor ez alatt azt értjük, hogy létezik egy f függvény, ami kielégíti (6)-t. Az f függvény tehát akkor létezik, ha F abszolút folytonos.

Többváltozós eloszlásfüggvénynek hívjuk a következő x_1, \dots, x_n valószínűségi változókra értelmezett függvényt:

$$F(\lambda_1, \dots, \lambda_n) = \mathbf{P}\{x_j(\omega) \leq \lambda_j, j = 1, \dots, n\}. \quad (7)$$

A (7) által definiált függvény minden változójában monoton, nem csökkenő és jobbról folytonos. Igazak rá továbbá a következők is:

$$\lim_{\lambda_j \rightarrow -\infty} F(\lambda_1, \dots, \lambda_n) = 0, \quad j = 1, \dots, n$$

$$\lim_{\lambda_1, \dots, \lambda_n \rightarrow \infty} F(\lambda_1, \dots, \lambda_n) = 1.$$

Továbbá $\lambda_j \leq \mu_j, j = 1, \dots, n$ -re teljesül az is, hogy:

$$\begin{aligned} & F(\mu_1, \dots, \mu_n) - \sum_{j=1}^n F(\mu_1, \dots, \mu_{j-1}, \lambda_j, \mu_{j+1}, \dots, \mu_n) \\ & + \sum_{\substack{j,k=1 \\ j < k}}^n F(\mu_1, \dots, \mu_{j-1}, \lambda_j, \mu_{j+1}, \dots, \mu_{k-1}, \lambda_k, \mu_{k+1}, \dots, \mu_n) \\ & - \dots + (-1)^n F(\lambda_1, \dots, \lambda_n) \geq 0. \end{aligned}$$

Az egyenlet bal oldalán lévő mennyiség a következő valószínűséget fejezi ki F segítségével:

$$\mathbf{P}\{\mu_j \leq x_j(\omega) \leq \lambda_j, j = 1, \dots, n\}.$$

Minden olyan F függvényt, ami eleget tesz ezen feltételeknek n -változós eloszlásfüggvénynek fogunk hívni. Egy ilyen függvény n dimenzióban Lebesgue-Stieltjes mértéket definiál:

$$\int_A \dots \int_A d_{\lambda_1, \dots, \lambda_n} F(\lambda_1, \dots, \lambda_n).$$

Egy Lebesgue mérhető és integrálható f függvényt pedig hasonló módon sűrűségfüggvénynek hívunk (minden $\lambda_1, \dots, \lambda_n$ -re) ha igaz rá a következő:

$$F(\lambda_1, \dots, \lambda_n) = \int_{-\infty}^{\lambda_1} \dots \int_{-\infty}^{\lambda_n} f(\mu_1, \dots, \mu_n) d\mu_1, \dots, d\mu_n \quad (8)$$

Ha minden lineáris Borel halmazra (X_1, \dots, X_n) teljesül, hogy:

$$\mathbf{P}\{x_j(\omega) \in X_j, j = 1, \dots, n\} = \prod_{j=1}^n \mathbf{P}\{x_j(\omega) \in X_j\}, \quad (9)$$

akkor az x_1, \dots, x_n valószínűségi változókat *kölcsönösen függetlennek* nevezzük. Ezt a feltételt meg lehet adni nyílt X_j halmazokra is, intervallumokra is és akár olyan félig végtelen intervallumokra, mint például $(-\infty, b_j]$. A definíciókból következik, hogy csak azok a valószínűségi változók függetlenek, amelyek közös többváltozós eloszlása az egyes eloszlásfüggvények szorzataként áll elő.

Ez előző bekezdésben egy valószínűségi változó helyett x_j -vel akár jelölhetünk m_j darab valószínűségi változót $(x_{j1}, \dots, x_{jm_j})$ és ekkor ha X_j ennek megfelelően egy m_j dimenziós halmaz, akkor az előzőekben tárgyalt elmélet segítségével kölcsönös függetlenséget tudunk definiálni véges számú valószínűségi változó aggregátumra, amikben véges számú valószínűségi változó van. Ha valamelyik aggregátumban mégis végtelen sok valószínűségi változó lenne, akkor a kölcsönös függetlenség alatt azt fogjuk érteni, hogy ha a végtelen aggregátumokat kicserélnénk véges aggregátumokra, akkor azok kölcsönösen függetlenek lennének. Valószínűségi változók végtelen sok aggregátuma kölcsönösen független ha A_1, A_2, \dots véges darabszámú aggregátum kölcsönösen függetlenül kiválasztható ezek közül.

Az itt megjelent definíciók komplex valószínűségi változókra is igazak ha (9)-ben megengedjük, hogy X_j a két dimenziós komplex sík Borel halmaza legyen és ez a definíció később majd hasznosnak is fog bizonyulni. Az előzőeket komplex valószínűségi változókra természetesen úgy is lehet érteni, miszerint azok valószínűségi változók párok, egy valós és egy képzetes résszel.

Most vegyünk egy x valószínűségi változót, és integráljuk azon az ω téren, amit Ω -val jelöltünk. Ha ez az integrál létezik, akkor ezt $\mathbf{E}\{x\}$ -nek fogjuk jelölni (bizonyos esetekben $\mathbf{E}\{x(\omega)\}$ -nek) és ezt az x valószínűségi változó *várható értékének* hívjuk majd.

$$\mathbf{E}\{x\} = \int_{\Omega} x d\mathbf{P}. \quad (10)$$

Tudjuk, hogy ez az integrál csak akkor létezik, ha $|x|$ integrálja véges.

2.3 A konvergenciáról

Vegyünk x, x_1, x_2, \dots valószínűségi változókat. Ekkor ha teljesül az, hogy majdnem minden ω -ra:

$$\lim_{n \rightarrow \infty} x_n(\omega) = x(\omega), \quad (11)$$

akkor azt mondjuk, hogy:

$$\lim_{n \rightarrow \infty} x_n = x. \quad (12)$$

1 valószínűséggel teljesül. Ilyen jellegű konvergenciáról csak akkor beszélhetünk, ha

$$\lim_{n \rightarrow \infty} \mathbf{P}\{\text{L.F.K.} | x_m(\omega) - x_n(\omega) | \geq \epsilon\} = 0, \quad (13)$$

és ennek minden pozitív ϵ -ra teljesülni kell. Itt a legkisebb felső korlátra bevezettük az L.F.K. rövidítést. A *sztochasztikus konvergenica* egy ennél gyengébb feltétel:

$$\lim_{n \rightarrow \infty} \mathbf{P}\{|x_n(\omega) - x(\omega)| \geq \epsilon\} = 0. \quad (14)$$

Ennek minden pozitív ϵ -ra teljesülni kell. A sztochasztikus konvergenicát így jelöljük:

$$\text{plim}_{n \rightarrow \infty} x_n = x. \quad (15)$$

Erre gyakran *valószínűségben vett konvergenica*-ként és *mértékben vett konvergenica*-ként hivatkoznak. Az 1 valószínűséggel teljesülő konvergenica és a valószínűségben vett konvergenica kapcsolata jól ismert, itt most nem fogjuk bebizonyítani a következőket:

- (a) az 1 valószínűséggel teljesülő konvergenica magával vonja a valószínűségben vett konvergenciát
- (b) $\text{plim}_{n \rightarrow \infty} x_n = x$ csak akkor teljesül ha az x_j sor minden x_{a_n} részsora tartalmaz olyan részsort, amely x -hez 1 valószínűséggel konvergál.

Azt mondjuk továbbá, hogy egy $\{x_n\}$ sor x -hez *középértékben* konvergál ha $\mathbf{E}\{|x_n|^2\} < \infty$ minden n -re és ha $\mathbf{E}\{|x|^2\} < \infty$ és ha

$$\lim_{n \rightarrow \infty} \mathbf{E}\{|x - x_n|^2\} = 0. \quad (16)$$

Erre a következő jelölést használjuk:

$$\text{l. i. m.}_{n \rightarrow \infty} x_n = x. \quad (17)$$

A középértékben vett konvergenica magával vonja a valószínűségben vett konvergeniciát. Valójában ha (17) teljesül, akkor minden $\epsilon > 0$ -ra teljesülni fog, hogy

$$\lim_{n \rightarrow \infty} \mathbf{P}\{|x_n(\omega) - x(\omega)| \geq \epsilon\} \leq \lim_{n \rightarrow \infty} \frac{\mathbf{E}\{|x_n - x|^2\}}{\epsilon^2} = 0. \quad (18)$$

Most vegyünk egy egydimenziós eloszlásfüggvény sort és ezt jelöljük $\{F_n\}$ -el. A legcélszerűbb ha az eloszlásfüggvényhez való konvergenciát úgy definiáljuk,

hogy $\lim_{n \rightarrow \infty} F_n(\lambda) = F(\lambda)$ teljesül F minden folytonossági pontjában. Ha ez a feltétel teljesül, akkor akkor a konvergencia egyenletes lesz F minden zárt folytonos intervallumán mindegy hogy az véges-e vagy végtelen. Hogyan tudunk most két eloszlásfüggvény G_1 és G_2 között – olyan távolság-definíciót bevezetni, ami ilyen jellegű konvergenciát eredményez? Ehhez rajzoljuk fel G_1 -et és G_2 -t, majd a szakadási pontokba húzzunk be függőleges vonalakat és ekkor a két függvény között a távolságot definiáljuk úgy, hogy maximálisan mennyit kell megtenni a két függvény között egy -1 meredekségű egyenes mentén. A távolság ilyen definíciója mellett az eloszlásfüggvények tere teljes metrikus tér lesz és $\lim_{n \rightarrow \infty} F_n(\lambda) = F(\lambda)$ teljesül F minden folytonossági pontjában, de csak akkor ha F_n és F között a távolság 0-hoz tart $1/n$ -el.

Azt mondjuk továbbá, hogy az x valószínűségi változó *eloszlásban konvergál* ha az $\{F_n\}$ függvény-sora az $\{x_n\}$ valószínűségi változóknak az előbb ismertetett módon konvergál egy eloszlásfüggvényhez. Ez a meghatározás valamelyest szerencsétlen, hiszen lesznek olyan esetek, amikor $\{x_n\}$ sehogy sem konvergál. Ha például az eloszlásfüggvények megegyeznek, de ettől függetlenül a valószínűségi változókat tetszőlegesen megválaszthatjuk (mondjuk kölcsönösen függetlenre választjuk őket) és így eloszlásban konvergáló valószínűségi változó sort kapunk.

2.4 A valószínűségi változók családjai

Vannak olyan esetek, amikor a valószínűségi változókat közvetlen módon adjuk meg. Ha mondjuk vesszük a számegyeneset, mint az ω teret, és az ω halmazoknak a Lebesgue mérhető halmazokat választjuk meg és ha a valószínűségi mértéket így definiáljuk:

$$\mathbf{P}\{A\} = \frac{1}{\sqrt{2\pi}} \int_A e^{-\lambda^2/2} d\lambda, \quad (19)$$

akkor az $1, \omega, \omega^2, \dots$ függvény-sor valószínűségi változók sora lesz. Sokszor azonban teljesen mindegy, hogy milyen ω téren dolgozunk és csak azt követeljük meg, hogy létezzen a valószínűségi változók egy olyan osztálya, ami kielégít bizonyos feltételeket. Ebben az esetben az történik, hogy véges aggregátumok többváltozós eloszlásai vannak megadva és feltesszük, hogy létezik a valószínűségi változók olyan családja, amire az adott eloszlás lesz igaz. Most ezt a gondolatot fogjuk kicsit részletesebben kifejteni. 3§2-ben lesz még szó arról hogyan kell más módokon megadni eloszlásfüggvényeket.

Például, amikor majd így kezdünk egy tételt: "vegyünk x_1, \dots, x_n kölcsönösen független valószínűségi változókat, amelyek eloszlásfüggvényei F_1, \dots, F_n " a tétel triviális lesz anélkül, hogy kimondanánk egy tételt az n valószínűségi változóról valamilyen téren. Nézzünk most konkrétan egy ilyen létezésről szóló tételt. Ez remélhetőleg tisztává fogja varázsolni a képet. Vegyük az ω teret, amiben n dimenzióban ξ_1, \dots, ξ_n pontok vannak. A Borel halmazokon az ω téren a következő módon lehet valószínűségi mértéket definiálni:

$$\mathbf{P}\{A\} = \int_A \dots \int_A dF_1(\xi_1) \dots dF_n(\xi_n), \quad (20)$$

ahogyan erről már §2.2-ben volt szó. Most pedig jelöljük x_j -vel a j . koordináta változót, avagy $x_j(\omega) = \xi_j$ ha ω a (ξ_1, \dots, ξ_n) pont. Ebben az esetben az x_j valószínűségi változók kölcsönösen függetlenek, eloszlásfüggvényük pedig F_1, \dots, F_n . Ezáltal tehát bebizonyítottuk, hogy találhatunk olyan ξ_1, \dots, ξ_n valószínűségi változókat, amelyek kielégítik a kívánt feltételeket és hogy pontosan ezeket a valószínűségi változókat lehet úgy értelmezni, mint koordináta-függvényeket az n dimenziós térben. Most pedig nézzünk egy általánosított eljárást, ami valószínűségi változók teljesen általános aggregátumaira is érvényes lesz.

Általánosságban meg fogunk adni egy index osztályt és ezt T -vel fogjuk jelölni, majd ehhez valószínűségi változók osztályát definiáljuk. Ezt jelöljük $\{x_t, t \in T\}$ -vel. Ha veszünk egy tetszőleges véges t halmazt (t_1, \dots, t_n) akkor $\xi_{t_1}, \dots, \xi_{t_n}$ többváltozós eloszlása (F_{t_1, \dots, t_n}) meghatározott. Az így meghatározott többváltozós eloszlásfüggvénynek kölcsönösen konzisztensnek kell lennie. Ez alatt azt értjük, hogy ha permutáljuk az indexeket, $\alpha_1, \dots, \alpha_n$ szerint, akkor teljesülnie kell annak, hogy:

$$F_{t_1, \dots, t_n}(\lambda_1, \dots, \lambda_n) \equiv F_{\alpha_1, \dots, \alpha_n}(\lambda_{\alpha_1}, \dots, \lambda_{\alpha_n}),$$

és ha $m < n$:

$$F_{t_1, \dots, t_m}(\lambda_1, \dots, \lambda_m) \equiv \lim_{\substack{\lambda_j \rightarrow \infty \\ j=m+1, \dots, n}} F_{t_1, \dots, t_n}(\lambda_1, \dots, \lambda_n).$$

Kolmogorov látta be, hogy csak ezeket a konzisztencia kritériumokat kell megkövetelni. Az x_t valószínűségi változókat a következő módon lehet megadni: Az Ω térnek pontokból álló teret választunk. $\omega : \xi_t, t \in T$, ahol ξ_t egy tetszőleges valós szám lehet. Az Ω tér tehát $t \in T$ -k függvényéből álló tér. Más megközelítésben, Ω koordináta tér, dimenziója pedig T kardinalitása. $t = s$ -ben a t függvény értéke egy ω függvényt definiál, ezt x_s -el jelöljük ha $x_s(\omega) = \xi_s$, ahol az ω függvény $\xi_{(\cdot)}$. A következő ω halmazhoz:

$$\{x_{t_j}(\omega) \leq \lambda_j, j = 1, \dots, n\}$$

a következő számot rendeljük mint mérték:

$$F_{t_1, \dots, t_n}(\lambda_1, \dots, \lambda_n).$$

Vagy még általánosabban, a következő ω halmazhoz:

$$\{[x_{t_1}(\omega), \dots, x_{t_n}(\omega)] \in A\},$$

ha A n dimenziós Borel halmaz, akkor a következő számot rendeljük mint mérték:

$$\int_A \dots \int_A d_{\xi_1, \dots, \xi_n} F_{t_1, \dots, t_n}(\xi_1, \dots, \xi_n).$$

Ha ilyen módon rendelünk az ω halmazokhoz mértéket akkor az adott osztályú ω halmazok által generált Borel mezőn valószínűségi mértéket kapunk. Az

ω függvények családja $\{\xi_t, t \in T\}$ az így kialakított eloszlásoknak megfelelő valószínűségi változók családja lesz.

Az alap ω tér (Ω) most Ω_T volt az előzőekben, ami minden $t \in T$ függvényt magában foglalja és így a valószínűségi változó család az descartesi térben lévő koordináta-függvények családja lett. Azonban még akkor is ha az alap tér nem a descartesi térben lévő valószínűségi változó család ($\{\xi_t, t \in T\}$), akkor is sok esetben lehet ilyenrel helyettesíteni és ezt majd lentebb látni is fogjuk.

Vegyünk most egy $\{\xi_t, t \in T\}$ valószínűségi változó családot. Ebben az esetben rögzített ω -ra az $x_t(\omega)$ függvényérték t egy függvényét adja meg. Az így kapott t függvényeket a család *mintavételi* függvényének fogjuk hívni. Abban a konkrét esetben, amikor Ω koordináta tér (Ω_T), és x_t a t . koordináta függvény, az ω bázispont és a mintavételi függvény koncepciója egybeesik. Mindenesetre gyakran praktikus lesz olyan kifejezéseket használni, mint mondjuk "majdnem minden mintavételi függvény" és emögött mindig azt a mértékelméleti koncepciót fogjuk érteni ami visszautal az Ω térre. Például a *majdnem minden mintavételi függvény* kifejezés egyenértékű azzal, hogy "majdnem minden ω ". Ha a T paraméterhalmaz véges vagy megszámlálhatóan végtelen, akkor helyesebb *mintavételi sorról* beszélni, mint mintavételi függvényekről és általában így is fogunk eljárni.

Végül még említsük meg azt is, hogy amikor eloszlásfüggvények osztályáról beszélünk (mint ahogyan a bevezetésben ezt meg is tettük), akkor és ennek segítségével Ω_T halmazokhoz mértéket definiálunk, gyakran célszerűbb nem az összes t ($t \in T$) függvényosztályt venni, hanem jobban járhatunk ha kikötjük, hogy a t paraméterek csak valamilyen X halmazból kerülhetnek ki. Nincs azonban feltétlenül szükség arra, hogy ezt az X halmazt t -től függően állapítsuk meg. Minden eddig fentebb elmondott teljesül akkor is ha X a végtelen számegyenes ($-\infty \leq x \leq \infty$) Borel halmaz és a megfelelően megválasztott eloszlások 1 valószínűséggel korlátozzák a valószínűségi változókat X -re.

Nézzük most azt az esetet, amikor többször egymás után dobunk egy dobókockával. Ezen példa segítségével próbálom illusztrálni a fentieket. Tegyük fel, hogy n -szer dobjuk el a kockát. A T paraméterhalmaznak akkor választhatjuk az $1, \dots, n$ egészek halmazát. Az Ω_T terünk ekkor egy n dimenziós tér, és ebben (ξ_1, \dots, ξ_n) pontok vannak. Bármely ponthoz, amely koordinátái egészek az $(1/6)^n$ mértéket tudjuk hozzárendelni. Ha pedig ilyen pontokból halmazokat készítünk, akkor minden ilyen halmaz mértéke a halmazban található pontok száma osztva 6^n -el lesz. Ezen módszer segítségével egy szabályos dobókocka n -szer való eldobásának matematikai modelljét kaptuk (lásd még §2.1). Az Ω_T tér egy j pontja pedig egy valószínűségi változó, ami a j . dobásban kapott értéket veszi fel. Nyilván így az Ω_T térben feleslegesen sok pontunk lesz. A $(0, \dots, 0)$ pont például fizikailag értelmezhetetlen és feleslegesen elbonyolítja a matematikai modellt. Pontosan ezen okból vezettük be az X halmazt az előző bekezdésben. Válasszuk most az X halmaznak az $1, 2, 3, 4, 5, 6$ egészeket, és így az ω tér 6^n pontból álló tér lesz n dimenzióban, a pontok koordinátái pedig az $1, 2, 3, 4, 5, 6$ egész értékeket vehetik majd fel. Ha pedig a pontokból halmazokat készítünk, akkor minden ilyen halmaz mértéke a halmazban található pontok száma osztva 6^n -el lesz.

2.5 Szorzattér reprezentációk

Vegyünk most egy x valószínűségi változót, az eloszlásfüggvényét pedig jelöljük F -el. Ha így teszünk, akkor F -ből Lebesgue-Stieltjes mértéket kapunk lineáris halmazokra:

$$\tilde{\mathbf{P}}\{A\} = \int_A dF(\xi).$$

A mérhető lineáris halmazok azok lesznek, amik mérhetőek ezen F mérték szerint. Ezen halmazokba beleértjük az összes Borel halmazt. Ha pedig valószínűségi mértékkel van dolgunk a számegyenesen és egy x -re adott feltételből következő eseményt vizsgálunk, akkor tekinthetjük azt mondjuk, hogy $x(\omega) \in A$ és itt A lineáris ponthalmaz és nem ω halmaz ($x(\omega) \in A$). Az előbb definiált lineáris mérték pedig biztosítja, hogy a valószínűség mind a két interpretációban meg fog egyezni. Formálisan azt mondhatjuk, hogy szükségünk van egy mértéktartó transzformációra, ami az ω tér és a számegyenes között működik. Ezt túlságosan részletes magyarázatok nélkül is megérthetjük. Ehhez vegyünk egy \tilde{x} koordináta (valószínűségi) változót a számegyenesen. Ez egy függvény, ami ξ értéket vesz fel a ξ koordinátájú pontban. Most tegyük fel, hogy csak olyan eseteket vizsgálunk, amikor az ω valószínűségi változók csak $\Phi(x)$ alakot öltenek, ahol Φ egy valós Baire függvény. Még általánosabban azt is mondhatnánk, hogy mérhető az F mérték szerint. Ebben az esetben gyakran praktikus ha az eredeti ω tér helyett a számegyeneset használjuk és ehhez alkalmazzuk még az F mértéket, ahogyan azt fentebb definiáltuk. Például ha van egy ω valószínűségi változónk $\Phi(x)$, akkor abból csinálhatunk egy $\Phi(\tilde{x})$ valószínűségi változót és ez már a számegyenesen lesz definiálva. \tilde{x} és x ugyan különböző tereken definiált valószínűségi változó, eloszlásuk azonban megegyezik. Még általánosabban azt is mondhatjuk, hogy ha $\Phi_1(x), \dots, \Phi_n(x)$ ilyen jellegű ω valószínűségi változók, akkor lehet őket a számegyenesen értelmezni és a többváltozós eloszlásuk meg fog egyezni, attól függetlenül, hogy különböző tereken vannak definiálva. Az alkalmazhatósági szempontokat figyelembe véve, nézzük most például $\Phi(x)$ várható értékét ha $\Phi(x)$ olyan tulajdonságú, mint amiről eddig szó volt. Ha az ω téren dolgozunk, akkor $\mathbf{E}\{\Phi(x)\}$:

$$\int_{\Omega} \Phi[x(\omega)] d\mathbf{P}.$$

Ha pedig most az új bázisteret használjuk ($\tilde{\Omega}$), ami a ξ tengely, akkor a $\Phi(\tilde{x})$ valószínűségi változóval van dolgunk, ennek várható értéke pedig így számolható:

$$\int_{-\infty}^{\infty} \Phi(\xi) dF(\xi).$$

A két mennyiség egyenlő (belátható) és ezáltal, amikor várható értéket szeretnénk számolni, akkor nincs feltétlenül szükség arra, hogy visszaközlekedjünk az eredeti bázistérbe.

Most vegyünk két valószínűségi változót, x -et és y -t. Közös kétváltozós eloszlásukat jelöljük F -el. Most is, hasonlóan az egydimenziós esethez, F a

síkon vett halmazokhoz valószínűségi mértéket fog definiálni:

$$\mathbf{P}\{A\} = \iint_A d_{\xi,\eta} F(\xi, \eta),$$

ha A mérhető F szerint. Tegyük fel továbbá, hogy minket csak olyan mennyiségek érdekelnek, amik kifejezhetőek $\Phi(x, y)$ alakban. Itt Φ két valós változó függvénye és F szerint mérhető. Ebben az esetben hasonló módon járhatunk el, mint az egydimenziós esetben és az eredeti Ω tér helyett vehetünk egy $\tilde{\Omega}$ teret, amit a ξ, η síkkal azonosítunk. Itt az a fontos, hogy az eredeti ω téren értelmezett mérték egy pont mértéktartó transzformáció segítségével az $\tilde{\Omega}$ téren is mértéket indukál. Ez az új tér gyakran jobban kezelhető mint az eredeti. Nagyon hasonlóan az előzőekhez, $\mathbf{E}\{\Phi(x, y)\}$, amit az ω téren ω mérték mellett értelmezzünk:

$$\int_{\tilde{\Omega}} \Phi[x(\omega), y(\omega)] d\mathbf{P}.$$

Ez pedig az új téren az új mérték mellett így értékelhető ki:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(\xi, \eta) d_{\xi,\eta} F(\xi, \eta).$$

A két megközelítés ugyanazon várható értéket adja.

Általánosságban vehetünk egy tetszőleges valószínűségi változó családot $(x_t, t \in T)$. Az ω halmazok legkisebb olyan Borel mezejét, ami szerint az x_t -k még mérhetők jelöljük $\mathcal{B}(x_t, t \in T)$ -vel. Ezt a halmazokból álló Borel mezőt ω halmazok osztálya generálja, amik $\{x_t(\omega) \in X\}$ alakúak. Itt $t \in T$, X pedig egy intervallum. Ebben az esetben gyakran célszerű az eredeti Ω tér helyett egy $\tilde{\Omega} = \Omega_T$ térben dolgozni. Ez a tér a t ($t \in T$) függvények tere. Belátható az, hogy ha valószínűségi mértéket tudunk definiálni ezen a függvénytéren olyan módon, hogy $\{\tilde{x}_t(\omega) \in X\}$ a koordináta változók osztálya, akkor minden véges \tilde{x}_t -ből képzett halmaz ugyanazon többváltozós eloszlással fog rendelkezni, mint a nekik megfelelő x_t változók. Kiderül, hogy lesz egy egyértelmű transzformáció, ami a $\mathcal{B}(x_t, t \in T)$ szerint mérhető ω valószínűségi változókat $\tilde{\omega}$ valószínűségi változókba viszi. Itt megegyező valószínűségi változók alatt azokat értjük, amik 1 valószínűséggel megegyeznek. A transzformáció kielégíti a következő feltételeket:

- (i) a transzformáció x_t -t \tilde{x}_t -be viszi, véges számú x_t -k Baire függvényei ugyanazon Baire függvényekbe mennek át, \tilde{x}_t -kre értelmezve,
- (ii) ha x egy ω valószínűségi változó és az egy $\tilde{\omega}$ \tilde{x} valószínűségi változóba megy át és ha valamelyik valószínűségi változó várható értéke létezik, akkor létezik a másiké is és ezek a mennyiségek megegyeznek,
- (iii) ha x egy ω valószínűségi változó és az egy Λ halmazon ($\Lambda \in \mathcal{B}(x_t, t \in T)$) 1-et, annak komplementén 0-t vesz fel, akkor a transzformáció azt egy $\tilde{\omega}$ \tilde{x} valószínűségi változóba viszi át és \tilde{x} egy $\tilde{\Lambda}$ halmazon 1-et, annak

komplementesen 0-t vesz fel és az így definiált halmaztranszformáció 1–1 megfeleltetés (azonos halmazok alatt a 0 mértékkel eltérő halmazokat értjük), valamint mértéktartó.

Így tehát azt kapjuk, hogy ha ω valószínűségi változókkal van dolgunk, amik $\mathcal{B}(x_t, t \in T)$ mérhető (vagy ezen Borel mező halmazaival), akkor a probléma átfogalmazható az $\tilde{\omega}$ téren értelmezett valószínűségi változókra. Az így kezelhető problémák osztálya némileg kibővíthető ha teljessé tesszük a $\mathcal{B}(x_t, t \in T)$ halmazokon definiált valószínűségi mértéket. Ezen szakasz elején már volt szó olyan esetekről, amikor T pontosan egy vagy két pontból állt. Az olyan problémák, amikben n darab valószínűségi változókkal van dolgunk redukálhatók n dimenziós koordinátatérre, az olyan esetek pedig, amikor végtelen sok valószínűségi változókkal kell számolnunk végtelen dimenziós koordinátatéren kezelendők. Minden egyes esetben az eredeti valószínűségi változókból az új térben koordináta-függvényeket kapunk.

Az előbbi gondolatmenet alkalmazása céljából nézzünk egy olyan esetet, amikor T két pontból áll. Vegyük azt a tételt, amikor x és y kölcsönösen független valószínűségi változó, várható értékeik léteznek és $\mathbf{E}\{xy\}$ is létezik, valamint teljesül, hogy:

$$\mathbf{E}\{xy\} = \mathbf{E}\{x\}\mathbf{E}\{y\}. \quad (21)$$

Első ránézésre ez nem úgy tűnik mintha egy hagyományos integrálokról kimondott tétel lenne és gyakran valóban úgy hivatkoznak rá mint a valószínűség számítás egy speciális tételére. Vegyük észre azonban, hogy ebben a tételben csupán két változó szerepel x és y , így tehát reprezentálhatjuk őket a síkon. Ebben a reprezentációban valóban csupán egy integrálokról kimondott tétellel lesz dolgunk. Ha x és y eloszlásfüggvényei G és H , F pedig a közös eloszlásfüggvény, akkor:

$$F(\xi, \eta) \equiv G(\xi)H(\eta),$$

mivel a valószínűségi változók függetlenek. A síkon vett reprezentációban (21) ekkor a következő alakot ölti:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi \eta dG(\xi) dH(\eta) = \int_{-\infty}^{\infty} \xi dG(\xi) \int_{-\infty}^{\infty} \eta dH(\eta) \quad (22)$$

A jobb oldalon álló kettős integrál egyszeres integrálra redukálható és így a következőt kapjuk:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi \eta dG(\xi) dH(\eta) = \int_{-\infty}^{\infty} \xi dG(\xi) \int_{-\infty}^{\infty} \eta dH(\eta). \quad (23)$$

A (21)-ben szereplő kettős integrált tehát iterált integrálként tudjuk kiértékelni. Minden (21)-re adott bizonyítás erre a szabványos bizonyításra vezethető vissza.

Eddig mindig feltettük, hogy a valószínűségi változók, amikkel dolgozunk valósak. A reprezentáció-elmélet komplex valószínűségi változókra való kiterjesztése triviális, ezért most ezt nem részletezzük.

2.6 Feltételes valószínűségek és várható értékek

Vegyünk most egy y valószínűségi változót és egy mérhető ω halmazt, amit jelöljünk M -el. Szeretnénk most feltételes valószínűséget definiálni az M halmazon. Szeretnénk bizonyos feltételek mellett y várható értékét is definiálni. Mielőtt ezt megtennénk nézzünk két speciális esetet.

Eset-1

Tegyük fel, hogy az x valószínűségi változó csak véges, vagy megszámlálható számosságú sor-értékeket (a_1, a_2, \dots) vehet fel. Ha $x(\omega) = a_j$, akkor M feltételes valószínűségét $\mathbf{P}\{M|x(\omega) = a_j\}$ -vel jelöljük és mindig meghatározott a következő módon ha $\mathbf{P}\{x(\omega) = a_j\} > 0$:

$$\mathbf{P}\{M|x(\omega) = a_j\} = \frac{\mathbf{P}\{\omega \in M, x(\omega) = a_j\}}{\mathbf{P}\{x(\omega) = a_j\}}. \quad (24)$$

Abban a speciális esetben, amikor y csak b_1, b_2, \dots értékeket vehet fel az $x(\omega) = a_j$ feltételre y feltételes eloszlását a következő módon kapjuk:

$$\mathbf{P}\{y(\omega) = b_k|x(\omega) = a_j\} = \frac{\mathbf{P}\{y(\omega) = b_k, x(\omega) = a_j\}}{\mathbf{P}\{x(\omega) = a_j\}} \quad (25)$$

A $\mathbf{P}\{M|x(\omega) = a_j\}$ feltételes valószínűség függ a_j -től. Így az x valószínűségi változó által felvett értékeinek függvényét kapjuk. Ha konkrétan $x(\omega)$ helyett a valószínűségi változó értékét írjuk be a feltételes valószínűség definíciójába, akkor egy új z valószínűségi változót kapunk:

$$z(\omega) = \mathbf{P}\{M|x(\omega) = a_j\}, \quad \text{ahol } x(\omega) = a_j,$$

ha $\mathbf{P}\{x(\omega) = a_j\} > 0$. $z(\omega)$ -t tetszőlegesen választjuk ha ezen az ω halmazon $(\{x(\omega) = a_j\})$ 0 valószínűséget venne fel. Egyértelműen megadtuk tehát a z valószínűségi változót, ha nem foglalkozunk a 0 valószínűségű ω halmazokkal. M feltételes valószínűsége x szerint $\mathbf{P}\{M|x\}$ -el jelölendő és ez alatt a z valószínűségi változót értjük (vagy egy ezzel ekvivalens meghatározást).

$$\mathbf{P}\{M|x\}_{x(\omega)=a_j} = \mathbf{P}\{M|x(\omega) = a_j\} \quad (26)$$

ha $\mathbf{P}\{x(\omega) = a_j\} > 0$. Most vegyünk a_j -k egy tetszőleges halmazát és jelöljük A -val, majd vezessük be a következő definíciót:

$$\Lambda = \{x(\omega) \in A\} = \bigcup_{a_j \in A} \{x(\omega) = a_j\}.$$

Ebben az esetben észre lehet venni a következőt:

$$\mathbf{P}\{\Lambda M\} = \int_{\Lambda} \mathbf{P}\{M|x\} d\mathbf{P}. \quad (27)$$

Hasonló módon az y valószínűségi változó x szerinti feltételes várható értékét $\mathbf{E}\{y|x\}$ -vel jelöljük és olyan valószínűségi változókra értelmezzük, amik eleget tesznek az alábbi egyenleteknek:

$$\mathbf{E}\{y|x\}_{x(\omega)=a_j} = \mathbf{E}\{y|x(\omega) = a_j\} = \sum_k b_k \mathbf{P}\{y(\omega) = b_k | x(\omega) = a_j\}$$

és $\mathbf{P}\{x(\omega) = a_j\} > 0$. A következő egyenlőség azonnal következik a definícióból:

$$\sum_k b_k \mathbf{P}\{y(\omega) = b_k, x(\omega) = a_j\} = \sum_{a_j \in A} \mathbf{E}\{y|x(\omega) = a_j\} \mathbf{P}\{x(\omega) = a_j\}.$$

Ezt a következő formában is fel lehet írni:

$$\int_{\Lambda} y d\mathbf{P} = \int_{\Lambda} \mathbf{E}\{y|x\} d\mathbf{P}. \quad (28)$$

(27) és (28)-ból a feltételes valószínűség és várható érték általánosabb definícióját is bevezethetjük és ezt most meg is tesszük.

Eset-2

Tegyük Ω az ξ, η sík és hogy a mérhető ω halmazok Lebesgue mérhető halmazok, a valószínűségi mértéket pedig a sűrűségfüggvény segítségével a következő módon lehet megadni:

$$\mathbf{P}\{\Lambda\} = \iint f(\xi, \eta) d\xi d\eta.$$

Az első és második koordináta x és a y függvényeket definiálnak, amelyek a (ξ, η) pontokban ξ és η értékeket vesznek fel. Ezekre a függvényekre mint valószínűségi változókra tekinthetünk a közös sűrűségfüggvényüket pedig f -el jelölhetjük. Bevezethetünk egy új sűrűségfüggvényt (η változó szerint):

$$\frac{f(\xi, \eta)}{\int_{-\infty}^{\infty} f(\xi, \zeta) d\zeta},$$

figyelve arra, hogy a nevező sose legyen 0. A fenti esettel analóg módon természetesen bevezethetünk egy olyan definíciót y valószínűségi változó feltételes eloszlására, amikor $x(\omega) = \xi$ és az y valószínűségi változó feltételes ($x(\omega) = \xi$) várható értékét megadhatjuk a következő módon is:

$$\frac{\int_{-\infty}^{\infty} \eta f(\xi, \eta) d\eta}{\int_{-\infty}^{\infty} f(\xi, \eta) d\eta}.$$

Itt most feltettük, hogy $\mathbf{E}\{|y|\} < \infty$. Az így megadott definíciók konzisztensek lesznek a következő általános értékű tárgyalásmóddal.

Általános eset

Itt most nem egy valószínűségi változó szerinti feltételes valószínűséget és várható értéket fogunk megadni hanem általános koncepciókat ismertetünk és ezeket fogjuk majd speciális esetekre is kimondani. Látni fogjuk, hogy a feltételes valószínűség a feltételes várható érték egy speciális esete, így célszerű először ezzel foglalkozni.

Vegyünk egy tetszőleges y valószínűségi változót, amelynek létezik a várható értéke. Jelöljük \mathcal{F} -vel mérhető ω halmazok Borel mezejét. Az ω halmazok azon Borel mezejét, ami \mathcal{F} halmaz vagy az \mathcal{F} halmazoktól 0 valószínűségi halmazokkal tér el $\mathcal{F}' \subset \mathcal{F}$ -el fogjuk jelölni. Ismeretes az a tény, miszerint ha van egy valószínűségi változónk, ami \mathcal{F}' mérhető, akkor az 1 valószínűséggel egyenlő egy olyan valószínűségi változóval, ami \mathcal{F} mérhető. Az y valószínűségi változó \mathcal{F} szerinti feltételes várható értékét $\mathbf{E}\{y|\mathcal{F}\}$ -el jelöljük és egy tetszőleges ω függvényként értelmezzük, ami mérhető az \mathcal{F}' Borel mezőn. A függvénytől megköveteljük továbbá, hogy integrálható legyen és hogy igaz legyen rá a következő egyenlet:

$$\int_{\Lambda} \mathbf{E}\{y|\mathcal{F}\} d\mathbf{P} = \int_{\Lambda} y d\mathbf{P}, \quad \Lambda \in \mathcal{F}. \quad (29)$$

Ugyan ez igaz lesz az \mathcal{F}' Borel mezőre és ha figyelembe vesszük \mathcal{F} és \mathcal{F}' kapcsolatát, így tehát a feltételes várható érték definíciója ekvivalens lesz \mathcal{F} -re és \mathcal{F}' -re. Vegyük észre, hogy (29) jobb oldala $\Lambda \in \mathcal{F}$ egy függvényét definiálja. Ez a függvény tökéletesen additív és eltűnik, amikor $\mathbf{P}\{\Lambda\} = 0$. A Radon-Nikodym tétel értelmében a Λ változó függvénye kifejezhető egy Λ feletti integrállal, az integrandus pedig egy \mathcal{F} Borel mezőn mérhető ω függvény. Az így megadott ω függvény tehát az $\mathbf{E}\{y|\mathcal{F}\}$ várható érték egy jó definíciója lesz, de elképzelhető más ω függvény is, ami ezzel 1 valószínűséggel egyenlő. Megfordítva azt is mondhatjuk, hogy a Radon-Nikodym tétel értelmében a feltételes várható érték két különböző definíciója majdnem mindenhol egyenlő. Így tehát az $\mathbf{E}\{y|\mathcal{F}\}$ várható értéket úgy adtuk meg, mint egy jól megválasztott valószínűségi változó osztály tetszőleges eleme. Ezen osztályból bárhogyan is választunk ki két tetszőleges valószínűségi változót, azok majdnem mindenhol egyenlőek lesznek és bármely olyan valószínűségi változó, ami majdnem mindenhol egyenlő egy olyan valószínűségi változóval, ami az osztályban van maga is az osztályban lesz. Hacsak nem jelezzük, hogy egy speciális esettel van dolgunk, bármikor amikor megjelenik a feltételes várható érték fogalma, akkor azalatt azt értjük, hogy azt bármely ekvivalens definícióra fel lehet cserélni.

Most vegyünk egy M tetszőleges ω halmazt. Jelöljük \mathcal{F} -vel mérhető ω halmazok Borel mezejét. Az y valószínűségi változót adjuk meg a következő módon:

$$y(\omega) = \begin{cases} 1, & \text{ha } \omega \in M \\ 0, & \text{ha } \omega \in \Omega - M. \end{cases}$$

Ebben az esetben M \mathcal{F} szerinti feltételes várható értékét $\mathbf{E}\{M|\mathcal{F}\}$ -el jelöljük és $\mathbf{E}\{y|\mathcal{F}\}$ -ként értelmezzük és ez alatt a feltételes várható érték egy tetszőleges definícióját értjük. A feltételes várható érték tehát egy tetszőleges ω függvény, ami \mathcal{F} -vel mérhető, ha pedig nem, akkor megköveteljük, hogy majdnem mindenhol egyenlő legyen egy olyan függvénnyel, ami mérhető \mathcal{F} -en. A függvény

legyen továbbá integrálható és teljesítse a következő egyenletet:

$$\int_{\Lambda} \mathbf{P}\{M|\mathcal{F}\}d\mathbf{P} = \mathbf{P}\{\Lambda M\}. \quad (30)$$

Ez a leírás sokat egyszerűsödik, ha fel tudjuk tenni, hogy \mathcal{F} -ben benne vannak a 0 mértékű halmazok is, mert ebben az esetben $\mathcal{F}' = \mathcal{F}$, továbbá a feltételes várható értékek és feltételes valószínűség mérhető lesz \mathcal{F} szerint. Láttuk azonban azt is, hogy minden esetben $\mathbf{E}\{y|\mathcal{F}\}$ -nak mindig lesz egy olyan változata, ami mérhető \mathcal{F} -en.

Most vegyünk egy tetszőleges $\{x_t, t \in T\}$ valószínűségi változó családot. Az ω halmazok legkisebb olyan Borel mezejét, ami szerint az x_t -k még mérhetőek jelöljük $\mathcal{F} = \mathcal{B}(x_t, t \in T)$ -vel. Ez alatt azt értjük, hogy A egy olyan Borel halmaz, amire $x_t(\omega) \in A$ és az \mathcal{F} egy olyan Borel mező, amit ilyen valószínűségi változók osztálya generál. Jelöljük továbbá \mathcal{F}' -vel ω halmazok Borel mezejét, ahol az ω halmazok vagy az \mathcal{F} Borel mezőn vannak, vagy azoktól 0 mértékű halmazokkal térnek el. Mostantól az \mathcal{F}' Borel mezőn lévő ω halmazok alatt azokat értjük, amelyek mérhetőek az x_t valószínűségi változók mintavételi terén és olyan ω függvényekről lesz majd szó, amelyek mérhetőek az \mathcal{F}' Borel mezőn. Ezek olyan függvények, amelyek majdnem mindenhol megegyeznek egy \mathcal{F} Borel mezőn mérhető függvénnyel (mint egy valószínűségi változó az x_t mintavételi téren). Abban a speciális esetben, amikor T az $1, \dots, n$ egészekből áll az ω halmaz (Λ) csak akkor mérhető az x_t mintavételi téren ha az

$$\{[x_1(\omega), \dots, x_n(\omega)] \in A\},$$

jellegű halmazoktól csak a 0 mértékű halmazokkal tér el. Itt A Borel halmaz (a valós esetben n dimenziós, komplex esetben $2n$ dimenziós). Az x ω függvény az x_t valószínűségi változók mintavételi terén valószínűségi változó, de csak akkor ha x majdnem mindenhol megegyezik x_1, \dots, x_n egy Baire függvényével.

Most vegyünk egy y valószínűségi változót és tegyük fel hogy létezik várható értéke, M pedig legyen egy mérhető ω halmaz. Az y valószínűségi változó $[M]$ feltételes várható értéke [valószínűsége] x_t szerint a következő módon jelölendő:

$$\mathbf{E}\{y|x_t, t \in T\}, \quad [\mathbf{P}\{M|x_t, t \in T\}],$$

és a következő módon definiálhatjuk ezeket:

$$\mathbf{E}\{y|\mathcal{F}\}, \quad [\mathbf{P}\{M|\mathcal{F}\}],$$

és ezen ekvivalens definíciók közül használhatjuk bármelyiket. Az itt megjelenő \mathcal{F} -et úgy kell érteni, mint ahogyan azt az előző bekezdésben bemutattuk. Most is, mint mindig \mathcal{F} helyett vehetünk \mathcal{F}' -t. A vizsgálatunk tárgyául szolgáló feltételes várható érték tehát olyan valószínűségi változóként van definiálva, ami x_t mintavételi terén mérhető, integrálja pedig megegyezik y integráljával minden olyan halmaz felett, ami x_t mintavételi terén mérhető. Ha ilyen módon tudjuk \mathcal{F} -et definiálni, akkor (29) és (30) egy kicsit kényelmesebb alakra hozható,

hiszen a definíciót az ω halmazok osztályára értelmezhetjük, amire az egyenleteknek van értelme. Ezen egyenletek jobb és bal oldala Λ egy tökéletesen additív függvényét adja meg és ezen függvényeket értékeik egyértelműen meghatározzák bármely \mathcal{F}_0 részmezőn (\mathcal{F} -en), ami \mathcal{F} -et generálja. A jelen esetben \mathcal{F}_0 -t választhatjuk ω halmazok osztályának, amely a következő alakú halmazok véges unióképzéséből jön létre:

$$\{x_{t_j} \in X_j, j = 1, \dots, n\},$$

ahol (t_1, \dots, t_n) T egy tetszőleges véges részhalmaza, az X_j -k pedig Borel halmazok. Elegendő tehát ha (29) vagy (30) (attól függ melyikről van szó) egy ilyen Λ -ra teljesül. Mivel (29) és (30) mindkét oldalán az integrálandó halmazokban additív elegendő csak külön az összegzendő tagokat vizsgálni. Ha nagyon szükséges, még akár azt is feltehetjük, hogy az X_j -k jobbról félig zárt intervallumok (vagy nyíltak vagy zártak).

Nézzünk egy speciális esetet, amikor T az $1, \dots, k$ egészekből áll. Ebben az esetben ez a két mennyiség:

$$\mathbf{E}\{y|x_1, \dots, x_k\}, \quad \mathbf{P}\{M|x_1, \dots, x_k\}$$

az előző két esetben adott diszkusszió alapján definiált. Valójában, az "Eset-1"-ben megadott $\int_{\Lambda} y d\mathbf{P} = \int_{\Lambda} \mathbf{E}\{y|x\} d\mathbf{P}$ általános esetben feltétel lesz a $\int_{\Lambda} \mathbf{E}\{y|\mathcal{F}\} d\mathbf{P} = \int_{\Lambda} y d\mathbf{P}$, $\Lambda \in \mathcal{F}$ definícióra. Most nézzük a feltételes valószínűség olyan definícióját, ami mérhető $\mathcal{F} = \mathcal{B}(x_1, \dots, x_k)$ szerint. Már láttuk, hogy ezt fel lehet írni a következő módon:

$$\mathbf{E}\{y|x_1, \dots, x_k\} = \Phi(x_1, \dots, x_k),$$

és itt $\Phi(x_1, \dots, x_k)$ egy k változós Baire függvény. Ebben az esetben ezt gyakran így írjuk:

$$\mathbf{E}\{x_j(\omega) = \xi_j, j = 1, \dots, k\}$$

ehelyett:

$$\mathbf{E}\{y|x_1, \dots, x_k\}|_{x_j(\omega)=\xi_j, j=1, \dots, k} = \Phi(\xi_1, \dots, \xi_k).$$

Láthatjuk tehát, hogy miért lehet közös alapokon megadni

$$\mathbf{E}\{y|x_t, t \in T\}, \quad \mathbf{P}\{M|x_t, t \in T\},$$

úgy mint egy y valószínűségi változó szerinti feltételes várható értéket és M valószínűséget x_t vagy $x_t(\omega), t \in T$ adott értékeire.

2.7 A feltételes valószínűség és várható érték általános tulajdonságai

Vegyünk egy várható értékkel rendelkező tetszőleges valószínűségi változót (y). Legyenek \mathcal{F} és \mathcal{G} mérhető ω halmazok Borel mezejei. Jelöljük továbbá \mathcal{F}' -vel ω halmazok Borel mezejét, ahol az ω halmazok vagy az \mathcal{F} Borel mezőn vannak, vagy azoktól 0 mértékű halmazokkal térnek el. Hasonlóan \mathcal{G} -re is. Most tegyük fel hogy $\mathcal{G}' \subset \mathcal{F}'$. Ebben az esetben $\mathbf{E}\{y|\mathcal{F}\}$ és $\mathbf{E}\{y|\mathcal{G}\}$ nem biztos hogy 1

valószínűséggel megegyeznek. A második átlagolás egy ritkább átlagolás mint az első. Még pontosabban fogalmazva azt mondhatjuk, hogy mindkét esetben \mathcal{G}' halmazok felett integráljuk az y valószínűségi változót, de az első nem biztos, hogy \mathcal{G}' mérhető. Abban az esetben viszont ha az első esetben beszélhetünk \mathcal{G}' mérhetőségről, akkor a két várható érték 1 valószínűséggel megegyeznek. Erről szól a következő tétel:

Tétel 1 *Ha veszünk $\mathcal{G}' \subset \mathcal{F}'$ Borel mezőket és feltesszük, hogy a várható érték valamelyik meghatározását használva (így az összesre igaz lesz) $\mathbf{E}\{y|\mathcal{F}\}$ mérhető \mathcal{G}' szerint, akkor 1 valószínűséggel igaz a következő:*

$$\mathbf{E}\{y|\mathcal{F}\} = \mathbf{E}\{y|\mathcal{G}'\}. \quad (31)$$

A bizonyításhoz elegendő azt megjegyezni, hogy $\mathbf{E}\{y|\mathcal{F}\}$ \mathcal{G}' mérhető a hipotézisünk miatt és hogy \mathcal{G}' felett az integráljai megegyeznek y integráljaival (még \mathcal{F}' feletti integráljai is megegyeznek y integráljaival). $\mathbf{E}\{y|\mathcal{F}\}$ így tehát kielégíti azokat a feltételeket, amik $\mathbf{E}\{y|\mathcal{G}'\}$ -t definiálják.

Tétel 2 *Vegyünk egy valószínűségi változó családot és jelöljük ezt így: $\{x_t, t \in T\}$. Tegyük fel, hogy T nem megszámlálhatóan végtelen. Ekkor ha y olyan hogy létezik várható értéke, akkor lesz egy olyan megszámlálhatóan végtelen T rész-aggregátum, $\{t_n, n \geq 1\}$ (ez y -től függ), amire:*

$$\mathbf{E}\{y|x_t, t \in T\} = \mathbf{E}\{y|x_{t_1}, x_{t_2}, \dots\}. \quad (32)$$

Jelöljük most T tetszőleges részhalmazát S -el, azt a legkisebb olyan Borel mezőt, ami szerint az x_t -k, $t \in S$ még mérhetőek jelöljük $\mathcal{F}_S = \mathcal{B}(x_t, t \in S)$ -vel. $\mathbf{E}\{y|x_t, t \in T\}$ -re tekinthetünk úgy mint $\mathbf{E}\{y|\mathcal{F}_T\}$ egy verziója. Mostantól az \mathcal{F}_T szerint mérhető feltételes várható értéket jelöljük $\mathbf{E}\{y|\mathcal{F}_T\}$ -vel. Mivel ez a várható érték mérhető \mathcal{F}_T szerint, T -nek mindig lesz egy olyan megszámlálhatóan végtelen S részhalmaza úgy hogy $\mathbf{E}\{y|\mathcal{F}_T\} \mathcal{F}_S \subset \mathcal{F}_T$ mérhető. Ekkor pedig (Tétel-1) miatt 1 valószínűséggel teljesül, hogy

$$\mathbf{E}\{y|\mathcal{F}_T\} = \mathbf{E}\{y|\mathcal{F}_S\}, \quad \text{Q.E.D.}$$

Most ha \mathcal{F} mérhető ω halmazok Borel mezeje, és ha z egy olyan ω függvény, ami \mathcal{F} szerint mérhető és ha $\mathbf{E}\{|z|\} < \infty$, akkor:

$$\mathbf{E}\{z|\mathcal{F}\} = z$$

1 valószínűséggel teljesül. z -t tehát definiálja az adott feltételes várható érték. Ezt a gondolatot még jobban általánosítva jutunk el a következő tételhez:

- 3 A Véletlen Folyamatok Definíciója
- 4 Folyamatok Kölcsönösen Független Valószínűségi Változókkal
- 5 Folyamatok Kölcsönösen Korrelálatlan vagy Ortogonális Valószínűségi Változókkal
- 6 Diszkrét Paraméterű Markov Folyamatok
- 7 Folytonos Paraméterű Markov Folyamatok
- 8 Martingálok
- 9 Független Növekményű Folyamatok
- 10 Ortogonális Növekményű Folyamatok
- 11 Diszkrét Paraméterű Stacionárius Folyamatok
- 12 Folytonos Paraméterű Stacionárius Folyamatok
- 13 Lineáris Legkisebb Négyzetes Predikció — Stacionárius Tág Értelmű Folyamatok