A Valószínűségszámítás I. előadássorozat hetedik előadása.

2003. március 18.

Összefoglaló:

A Poisson eloszlás (folytatás)

Láttuk az előző előadáson, hogy amennyiben az S_n valószínűségi változók, $n = 1, 2, \dots$, binomiális eloszlásúak n és p_n paraméterekkel, azaz S_n eloszlása megegyezik egy $\sum\limits_{j=1}^{n} \xi_j^{(n)}$ összeg eloszlásával, ahol a $\xi_j^{(n)}$ valószínűségi változók függetlenek, és

$$P\left(\xi_j^{(n)} = 1\right) = 1 - P\left(\xi_j^{(n)} = 0\right) = p_n, \quad 1 \le j \le n,$$

továbbá $\lim\limits_{n\to\infty} np_n = \lambda$ valamely $\lambda > 0$ számmal, akkor $\lim\limits_{n\to\infty} P(S_n = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$ minden k egész számra.

Jegyezzük meg továbbá, hogy szintén következik az előző előadás eredményeiből, hogy amennyiben olyan T_n, $n = 1, 2, \dots$, véletlen összegeket tekintünk, amelyek $T_n = \sum\limits_{j=1}^{n} \eta_j^{(n)}$ alakban írhatók, ahol rögzített n számra az $\eta_j^{(n)}$ valószínűségi változók függetlenek, és Poisson eloszlásúak $\dfrac{\lambda_n}{n}$ paraméterrel, és $\lim\limits_{n\to\infty} \lambda_n = \lambda > 0$, akkor $\lim\limits_{n\to\infty} P(T_n = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$ minden k egész számra. Ekkor ugyanis T_n Poisson eloszlású λ_n paraméterrel.

Belátunk egy eredményt, amelynek a fent tekintett két példa speciális esete. Azután megtárgyaljuk egy példán keresztül, hogy ez az eredmény magyarázatot ad arra, hogy miért fontos a Poisson eloszlás, miért tételezhetjük fel, hogy bizonyos véletlen jelenségeket Poisson eloszlású valószínűségi változók írnak le.

Tétel. *Legyenek minden rögzített $n = 1, 2, \dots$ számra $\xi_1^{(n)}, \dots, \xi_n^{(n)}$ független egyforma eloszlású, nem negatív egész értékeket felvevő valószínűségi változók, amelyek teljesítik a következő feltételeket:*

a) $\lim\limits_{n\to\infty} nP\left(\xi_1^{(n)} = 1\right) = \lambda > 0.$

b) $\lim\limits_{n\to\infty} nP\left(\xi_1^{(n)} \ge 2\right) = 0.$

Ekkor az $S_n = \sum\limits_{j=1}^{n} \xi_j^{(n)}$, $n = 1, 2, \dots$, *véletlen összegekre teljesül a*

$$\lim_{n\to\infty} P(S_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

reláció minden k nem negatív egész számra.

Megjegyzés: Érvényes e tétel állításának általánosítása megfelelő feltételek mellett független nem negatív egész értékű, de nem feltétlenül azonos eloszlású valószínűségi változók összegére is. Egy ilyen állítást, amelyet a fent megfogalmazott tételhez hasonlóan lehet bizonyítani, nem kötelező házi feladat formájában megfogalmazok.

A tétel bizonyításának vázlata: Tekintsük az S_n valószínűségi változók

$$G_n(x) = \sum_{j=0}^{\infty} P(S_n = j)x^j, \quad n = 1, 2, \dots,$$

generátorfüggvényeit. Elég belátni, hogy $\lim_{n\to\infty} G_n(x) = G(x)$ valamilyen $-A < x < A$, $A > 0$, intervallumban, ahol $G(x) = e^{\lambda(x-1)}$ a λ paraméterű Poisson eloszlás generátorfüggvénye. Ugyanis, mivel hatványsorok konvergenciája esetében szabad tagonként deriválni, ezért teljesül a $\lim_{n\to\infty} \left.\frac{d^k G_n(x)}{dx^k}\right|_{x=0} = \left.\frac{dG^k(x)}{dx^k}\right|_{x=0}$ reláció minden $k = 0, 1, 2, \dots$ számra. Azaz $k! \lim_{n\to\infty} P(S_n = k) = k! \frac{\lambda^k}{k!} e^{-\lambda}$ minden $k = 0, 1, \dots$ számra. Ennek a bizonyítandó relációnak az érvényességét a következő észrevétel segítségével bizonyíthatjuk:

Legyen $g_n(x) = \sum\limits_{j=0}^{\infty} P(\xi_1^{(n)} = j)x^j$, a $\xi_1^{(n)}$ valószínűségi változó generátorfüggvénye. Ekkor a Tétel feltételei miatt $G_n(x) = g_n^n(x)$ minden x számra. Ezért logaritmust véve a bizonyítandó relációban elég megmutatni azt, hogy $\lim_{n\to\infty} n \log g_n(x) = \lambda(x-1)$. Továbbá, mivel $P\left(\xi_1^{(n)} = 0\right) = 1 - P\left(\xi_1^{(n)} = 1\right) - \sum\limits_{j=2}^{\infty} P\left(\xi_1^{(n)} = j\right)$, ezért $g_n(x) = 1 + P\left(\xi_1^{(n)} = 1\right)(x-1) + \sum\limits_{j=2}^{\infty} P\left(\xi_1^{(n)} = j\right)\left(x^j - 1\right)$. Ezért az a) és b) feltételek miatt azt várjuk, hogy a $g_n(x) \sim 1 + \frac{\lambda}{n}(x-1)$ formula jó közelítés. Mivel $\log(1+u) \sim u$ kis u számokra ezért természetes azt várni, hogy $\log g_n(x) \sim \frac{\lambda}{n}(x-1)$ jó közelítés, és $\lim_{n\to\infty} \log g_n(x) = \lambda(x-1)$, ahonnan következik a Tétel állítása. A Tétel bizonyítása a fenti közelítések jogosságának ellenörzéséből áll.

A Tétel bizonyításának befejezése. Vegyük észre, hogy minden $\varepsilon > 0$ számra létezik olyan $n_0 = n_0(\varepsilon)$ küszöbindex, amelyre $\left| g_n(x) - \left(1 + \frac{\lambda}{n}(x-1)\right)\right| < \frac{\varepsilon}{n}$, ha $n \ge n_0$ és $|x| < 1$. Valóban a $g_n(x) = 1 + P\left(\xi_1^{(n)} = 1\right)(x-1) + \sum\limits_{j=2}^{\infty} P\left(\xi_1^{(n)} = j\right)\left(x^j - 1\right)$ azonosság teljesül. Ezenkívül a b) relációból következik, hogy

$$\left|\sum_{j=2}^{\infty} P\left(\xi_1^{(n)} = j\right)\left(x^j - 1\right)\right| \le 2\sum_{j=2}^{\infty} P\left(\xi_1^{(n)} = j\right) = 2P\left(\xi_1^{(n)} \ge 2\right) \le \frac{\varepsilon}{2n},$$

és az a) relációból pedig az, hogy $\left|P\left(\xi_1^{(n)}=1\right)(x-1)-\frac{\lambda}{n}(x-1)\right|<\frac{\varepsilon}{2n}$, ha $n\ge n_0$ és $|x|<1$. Ezért igaz a fenti azonosság.

Továbbá, mivel mint azt például a $\log(1+x)$ függvény Taylor sorfejtéséből lehet látni, $|\log(1+u)-u|<u^2$, ha $|u|<\frac{1}{2}$, ezért a fenti egyenlőtlenségből $u=g_n(x)-1$ választással kapjuk, hogy $|\log g_n(x)-(g_n(x)-1)|<\frac{2\varepsilon}{n}$, és $\left|\log g_n(x)-\frac{\lambda(x-1)}{n}\right|<\frac{3\varepsilon}{n}$, ha $n\ge n_1$, és $|x|\le 1$ alkalmas $n_1=n_1(\varepsilon)$ küszöbindexre. Mivel ez az állítás igaz minden $\varepsilon>0$ számra, ezért $\lim\limits_{n\to\infty} n\log g_n(x)=\lambda(x-1)$, $\lim\limits_{n\to\infty} g_n(x)^n=e^{\lambda(x-1)}$, ha $x<1$. Viszont láttuk, hogy innen következik a Tétel állítása.

Nem kötelező házi feladat.

Legyen

$$\begin{matrix} \xi_{1,1}\dots,\xi_{1,n_1} \\ \vdots \qquad \vdots \\ \xi_{k,1}\dots,\xi_{k,n_k} \\ \vdots \qquad \vdots \end{matrix}$$

szériasorozat, azaz tegyük fel, hogy az egy sorban álló valószínűségi változók függetlenek. Tegyük fel továbbá, hogy e valószínűségi változók teljesítik a következő feltételeket:

1.) A $\xi_{k,j}$ valószínűségi változók nem negatív egész értékeket vesznek fel.

2.) $P(\xi_{k,j}=1)=\lambda_{k,j}$, $\lim\limits_{k\to\infty}\sum\limits_{j=1}^{n_k}\lambda_{k,j}=\lambda>0$.

3.) $\sup\limits_{1\le j\le n_k}\lambda_{k,j}\to 0$, ha $k\to\infty$, és $\sum\limits_{j=1}^{n_k}P\left(\xi_{k,j}\ge 2\right)\to 0$, ha $k\to\infty$.

Ekkor az $S_k=\sum\limits_{j=1}^{n_k}\xi_{k,j}$ valószínűségi változók eloszlásban konvergálnak a λ paraméterű Poisson eloszláshoz, ha $k\to\infty$, azaz $\lim\limits_{k\to\infty}P(S_k=l)=\frac{\lambda^l}{l!}e^{-\lambda}$ minden $l=0,1,2,\dots$ számra.

A Tétel szemléletes tartalma: Tekintsük például a csillaghullást. Arra vagyunk kíváncsiak, hogy amennyiben egy éjszaka azt figyeljük, hány hullócsillagot látunk egy adott időintervallumban, (mondjuk egy óra alatt), akkor a lehullott csillagok (véletlen száma) milyen valószínűségi törvényeknek tesz eleget. Osszuk fel az egy óra időintervallumot rövid ΔT hosszúságú időintervallumokra. Akkor a lehullott csillagok száma e rövid ΔT időintervallumokban lehullott csillagok összege. Ezenkívül feltehetjük, hogy diszjunkt időintervallumokban lehullott csillagok száma egymástól független, és a különböző rövid intervallumokban lehulló csillagok száma hasonló valószínűségi törvényeknek tesz eleget.

Annak valószínűsége, hogy egy csillag egy rövid időintervallumban lehullik nagyon kicsi, és arányos az időintervallum hosszával. Annak valószínűsége, hogy egy kis időintervallumban kettő vagy még több csillag hullik, még ehhez képest is elhanyagolható. Ez azt jelenti, hogy természetes feltenni, hogy teljesülnek az előbb megfogalmazott tétel feltételei. Ezért az alkalmazható, és egy adott időintervallumban lehullott csillagok száma Poisson eloszlású. Hasonló érvelés alkalmazható sok más hasonló esetben, és ez magyarázza meg, miért különösen fontos a Poisson eloszlás.

Egy a hipergeometrikus eloszlással kapcsolatos statisztikai problémáról.

Tekintsük először a következő problémát:

Feladat:

Egy tóban 3000 hal van. Véletlenül kihalásznak belőle 1000 darabot, és ezekre piros pöttyöt festenek és visszaengedik őket. Ezután ismét kifognak véletlenül 1000 halat. Mi annak a valószínűsége, hogy a kifogott halak között 100 megfestett van?

Megjegyzés: A gyakorlatban előforduló kérdés ennek fordítottja. Elvégezzük a fenti kisérletet, összeszámláljuk a megfestett halakat, és ebből próbálunk következtetni a tóban levő halak számára. Hogyan érdemes ezt csinálni?

Ezt a problémát érdemes részletesebben megtárgyalni. Valójában, nem tudjuk, hogy hány hal van a tóban. De ennek megbecslése érdekében a következő eljárást alkalmazhatjuk:

Végezzünk két fogást, az első fogásban kifogott halakat jelöljuk meg. Ezután meg akarjuk állapítani mennyi hal lehet összesen a tóban. Ezt természetesen csak bizonyos (véletlentől függő) pontossággal tudjuk meghatározni. Az ilyen tipusú feladatok tipikusak a matematikai statisztikában, az ilyen problémák vizsgálatát nevezik becsléseleméletnek. Világos, hogy az, hogy 1000-nél alig több hal van, nem túl valószínű, mert akkor sokkal több megjelölt hal lenne a második fogásban. Az hogy rengeteg, mondjuk 1 000 000 hal lenne a tóban szintén nem túl valószínű, mert akkor sokkal kevesebb megjelölt hal lenne a második fogásban. A matematikai statisztikában kidolgoztak egy általános elvet, a maximum likelihood módszernek nevezett eljárást, amely nagyon általános feltételek mellett nagyon jó módszert ad, és ez a jelen esetben is alkalmazható. Tárgyaljuk meg ezt a módszert a jelen esetben. Tekintsünk kissé általánosabb esetet. Vezessük be a következő jelöléseket:

x a tóban lévő halak (ismeretlen) száma,

n az első fogásban kifogott (és megjelölt) halak száma,

r a második fogásban kifogott halak száma,

k a második fogásban kifogott előzöleg megjelölt halak száma. Annak valószínűsége, hogy adott (ismeretlen) x és n, r számok esetén pontosan k megjelölt halat fogunk ki

$$q_k(x,n,r) = \frac{\binom{n}{k}\binom{x-n}{r-k}}{\binom{x}{r}}.$$

Tekintsük az ismeretlen x szám (maximum likelihood) becslésének azt az x számot, amelyre a $q_k(x,n,r)$ mennyiség (rögzített n, k és r számok mellett) maximális. Határozzuk meg a fenti feladatban a maximum likelihood becslést.
Némi számolás mutatja, hogy

$$\frac{q_k(x,n,r)}{q_k(x-1,n,r)}=\frac{x-n}{x-n-r+k}\cdot\frac{x-r}{x}=\frac{x^2-rx-nx+rn}{x^2-rx-nx+kx}.$$

Ez a tört kisebb mint egy, ha $rn<kx$, nagyobb mint egy, ha $rn>kx$. Ezért a becslés $rn=kx$, azaz $x=\dfrac{rn}{k}$, pontosabban az e számot közrefogó egész számok valamelyike. Valóban $x<\dfrac{rn}{k}$ esetében a $q_k(x,n,r)$ függvény (mint az x változó függvénye rögzített k, n és r paraméterekkel) monoton nő, $x>\dfrac{rn}{k}$ esetében pedig a $q_k(x,n,r)$ függvény monoton csökken.

Természetes kérdés az, hogy az így kapott becslés valóban jó-e. Azt nem várhatjuk, hogy az adott becslés teljesen pontos. A természetes elvárás az, hogy meg tudunk adni az x pontnak viszonylag kis környezetét, egy olyan $[x-a,x+a]$ intervallumot, amelyre igaz, hogy annak valószínűsége, hogy a halak valódi száma ebbe az intervallumba esik nagyobb mint egy előírt egyhez közeli szám. Az ilyen intervallumot a matematikai statisztikában kondfidencia intervallumnak szokták nevezni.

Ahhoz, hogy ilyen konfidenciaintervallumot tudjunk szerkeszteni szükség van bizonyos valószínűségi változók eloszlásának jobb ismeretére, és ez a valószínűségszámítás egyik alapvető feladata. Jegyezzük meg, hogy a most vizsgált feladatban, amennyiben a tóban x számú hal van, az első fogásban n, a második fogásban r pedig r halat fogunk ki, akkor a második fogásban kifogott véletlen k számú megjelölt halak számának a várható értéke $Ek=\dfrac{rn}{x}$. Ahhoz, hogy vizsgálni tudjuk mennyire jó a becslés, hogyan lehet jó konfidenciaintervallumot konstruálni, azt kell megértenünk, hogy mekkora az ingadozása a k valószínűségi változónak a várható értéke körül. Ezért érdemes kiszámolni egy hipergeometrikus eloszlás szórásnégyzetét kiszámolni, amit gyakorlaton megtettünk. További értékes információkat nyerhetünk, ha tételeket bizonyítunk hipergeometrikus eloszlások aszimptotikus eloszlására, ha a benne szereplő paraméterek nagyok. Bár ezzel a kérdéssel nem fogunk foglalkozni, hasonló problémákat fogunk tárgyalni, amelyek vizsgálatában ilyen jellegű kérdések megoldása hasznos.

Általános valószínűségi változók, velük kapcsolatos alapvető fogalmak.

Megbeszéltük, hogy egy ξ (valós értékű) valószínűségi változó egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn értelmezett (mérhető) függvény. Általában, nem a valószínűségi változókat adjuk meg, hanem azoknak alább definiált eloszlásfüggvényét.

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.*

Tekintsük például egy szabályos dobókocka feldobását, illetve azt a valószínűségi változót, amely megmondja mi a dobás eredménye. Hogy néz ki ennek a ξ valószínűségi változónak az $F(x)$ eloszlásfüggvénye?

Ez 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 $P(\xi < x)$ 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 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 $6 < x < \infty$,

Feladat:

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.

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 történik az, hogy ez a valószínűségi változó bizonyos értékeket vesz fel. Ezért természetes, hogy csak ezeket a valószínűségeket adjuk meg. 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. Viszont bizonyos mértékelméleti eredményekből következik, hogy az $F(x)$ eloszlásfüggvény meghatározza az összes ilyen esemény valószínűségét. Az, hogy egy halmaz ,,szép" pontosan azt jelenti, hogy ez a halmaz Borel mérhető. Lássuk, hogyan lehet néhány ilyen esemény valószínűségét meghatározni.

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

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

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

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

Hasonlóan, ha adva vam 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.

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

Lemma. *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 A. *Ha egy $F(x)$ függvény teljesíti az előző 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 A felhasznál egy fontos mértékelméleti eredményt, amelyet megfogalmazunk, de nem bizonyítunk.

Tétel B. *Ha egy $F(x)$ függvény teljesíti az előző lemmában megfogalmazott a)—d) tulajdonságokat, akkor létezik egy olyan úgynevezett $\mu_F(\cdot)$ az F függvényhez kapcsolódó 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 $-\infty \le a < b \le \infty$ számra. Speciálisan, $\mu_F((-\infty, b)) = F(b)$ minden $-\infty < b < \infty$ számra.

Egy fenti tulajdonságú $F(x)$ függvény, egyértelműen meghatározza azokat az $\mu_F(B)$ számokat minden Borel mérhető B halmazra, amelyekre teljesül, hogy $\mu_F((-\infty, b)) =$

$F(b)$ minden $-\infty < b < \infty$ számra, és $\mu_F(\cdot)$ σ-additív halmazfüggvény a számegyenes Borel σ-algebráján.

Jegyezzük meg, hogy a $\mu_F(B) = P(\{\omega\colon \xi(\omega) \in B\})$ halmazfüggvény σ-additív a számegyenes Borel σ-algebráján, és $P(\{\omega\colon \xi(\omega) < b\}) = F(b)$. Ezért az előbb megfogalmazott Tétel B azt is állítja, hogy az $F(x)$ eloszlásfüggvény meghatározza a $P(\{\omega\colon \xi(\omega) \in B\})$ valószínűségeket minden Borel–mérhető B halmazra.

A 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 Tétel A bizonyítása: 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 a Tétel B állítása fogalmazza meg. Legyen $\xi(x) = x$, azaz jelen példában az ω elemi események az x valós számok, és a ξ valószínűségi változók 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: A Tétel A bizonyításának egyetlen nehezebb lépése annak bizonyítása, hogy a konstruált μ_F valószínűségi mérték valóban σ-additív, azaz a Tétel B állításának bizonyítása. Megjegyezzük, hogy ennek az állításnak a bizonyítása elkerülhetetlen. Ugyanis a Tétel B állítása könnyen levezethető a Tétel A állításából.

2. megjegyzés: Felmerülhet a kérdés, hogy miért tárgyaltuk külön a Tétel A-t, miért érdekel minket annak az eredménye. Ennek 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 a feladat értelmes-e, azaz valóban létezik-e olyan valószínűségi változó, amelyiknek az eloszlásfüggvényét megadtuk. A Tétel A azt mondja ki, hogy amennyiben az általunk megadott képlet teljesít néhány olyan feltételt, amelyiket minden eloszlásfüggvénynek teljesíteni 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.

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 választjuk. Elmagyarázzuk azt a képet, amely természetessé teszi a használt Lebesgue mérték definicióját, és szemléletesen megmutatjuk, 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 *Fontos Tétel*-nek nevezett eredmény lehetővé teszi egy valószínűségi változó várható értékének kiszámítását csak a valószínűségi változó eloszlásfüggvényének az ismeretében.

Fontos Tétel. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelynek $F(x)$ az $F(x) = P(\xi < x)$, $-\infty < x < \infty$, eloszlásfüggvénye. Ekkor*

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

ahol a fenti integrál Lebesgue–Stieltjes integrált jelöl az F 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 Fontos Tétel következő általánosítása.

A Fontos Tétel általánosítása. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelynek $F(x)$ az $F(x) = P(\xi < x)$, $-\infty < x < \infty$, 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 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.

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 szerep;ő feltételnek, hogy egy diszkrét valószínűségi változó várható értékét definiáló összeg legyen abszolut konvergens.

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. Ezt tesszük meg az alábbiakban egy rövid kiegészítésben, amelyben nem törekszünk 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 maga az integrálás is az ezen a mezőn értelmezett téren létezik. Mégis a Fontos tételben, illetve annak általánosításában olyan eredményt fogalmaztunk meg, amelyik lehetővé teszi egy ilyen integrálnak a kiszámolását anélkül, hogy definiálnánk a valószínűségi mezőt, ahol dolgozunk. A tétel talán legfontosabb állítása az, hogy elég tudnunk a valószínűségi változó eloszlásfüggvényét, és akkor ki tudjuk számolni mind a valószínűségi változó, illetve a valószínűségi változó 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 jegyezzük meg, hogy tanultuk korábban, hogyan lehet kiszámolni diszkrét eloszlású valószínűségi változók, illetve azon függvényeinek várható értékét. Ezt érdemes ideírni, hogy lássuk hasonlóságukat a Fontos Tétellel, illetve annak általánosításával. Valójában a korábban megfogalmazott állítások ezeknek az eredményeknek az általánosításai.

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.

Kiegészítés: A Lebesgue integrál fogalama

A Lebesgye 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 σ-additív halmazfüggvény akkor σ-véges, ha létezik az X alaphalmaz olyan felbontása megszámlálhatóan sok B_j, halmazra, $X = \bigcup\limits_{j=1}^{\infty} B_j$, melyekre $\mu(B_j) < \infty$. Az $(X, \mathcal{A}, \mu)$ téren értelmezett (mérhető) $f(\cdot)$ függvényeknek 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, hogy 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 analog ahhoz, ahogy diszkrét eloszlású függvények integrálját definiáltuk. Az egyetlen különbség az, hogy itt csak véges sok értékeket felvevő függvényeket tekintünk, míg ott megengedtük azt is, hogy megszámlálhatóan sok értéket vegyenek fel. Definiálhattuk volna az elemi függvényeket is általánosbban, megengedhettük volna, hogy megszámlálható sok értéket vegyenek fel, de az itt tekintett definició technikailag kényelmesebb.) 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 $\int f(x)\mu(\,dx) = \int f_+(x)\mu(\,dx) - \int(-f_-(x)\nu(\,dx)$. 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, mi a kapcsolat az így definiált integrál és az analízisben tanult Riemann és Riemann–Stieltjes integrál között, 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 definiáljuk, hogy először alkalmas közelítő összegeket vezetünk be a következő módon. Felosztjuk 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, és mindegyik $[x_{k-1},x_k]$ intervallumban kijelölünk egy $\xi_k\in[x_{k-1},x_k]$ pontot, és tekintjük a $\sum\limits_{k=1}^n f(\xi_k)(x_k-x_{k-1})$ közelítő összegeket. Az $\int_a^b f(u)\,du$ integrált úgy definiáljuk mint ezen integrálközelítő összegek limeszét, ha $\sup\limits_{1\le k\le n}(x_k-x_{k-1})$ nullához tart.

Adva egy $F(u)$ függvény is az $[a,b]$ intervallumon az $\int_a^b f(u)F(\,du)$ Riemann–Stieltjes intergrált hasonlóan definiáljuk. 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 helyettesítjük, azaz a $\sum\limits_{k=1}^n f(\xi_k)\,(F(x_k)-F(x_{k-1}))$ közelítőösszegeket, illetve azok határértékét tekintjük.

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. Ezt is szokás Lebesgue mértéknek nevezni. Vegyük észre, hogy mind a Riemann mind az itt tekintett Lebesue integrál esetében először bizonyos elemi függvényeknek definiáljuk az integrálját alkalmas közelítő összegek segítségével, majd a definiciókat 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. 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 $a\le u<v\le b$. 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. Ebben az esetben is elmondhatjuk, hogy a Riemann-Stieltjes és Lebesgue–Stieltjes integrálok megegyznek, ha mind a két integrál létezik.

Láttuk, hogy a Lebesgue és Riemann integrálok definiciójának a szelleme 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 dediniálni tudjuk a Lebesgue integrálját. Ugyanis az integrál definiciójában szükséges határátmeneteketet a Lebesgue integrál definiciójában könnyebb végrehajtani mint a Riemann integráléban. 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, azaz tetszőleges f és g 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 milyen feltételek mellett cserélhető fel a limesz és integrálás sorrendje, azaz a $\lim \int f_n(x)\mu(\,dx) = \int \lim f_n(x)\mu(\,dx)$ azonosság milyen feltételek mellett érvényes. A következő előadásban és általában független valószínűségi változók vizsgálatában fontos szerepet játszik a következő alkalommal ismertetendő Fubini tétel, amely szerint egy szorzatmérték szerinti kétváltozós függvény integrálja szukcesszív integrálással is kiszámolható.

Végül megjegyezzük, hogy a Fontos Tételnek illetve annak általánosításának a bizonyítása viszonylag egyszerű, ha jól megértjük a Lebesgue integrál definicióját. Az állítás könnyen látható 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 megkapjuk a Fontos Tétel általánosításának a bizonyítását.