Valószínűségszámítás előadás informatika BSC/A szakosoknak és matematikai elemző BSC-seknek

2017/2018 1. félév

9. előadás

november 20.

Valószínűségi változók függetlensége (ismétlés)

- n eseményrendszer független, ha

$$P(A_{i_1} \cap A_{i_2} \cap \dots A_{i_k}) = P(A_{i_1})P(A_{i_2})\dots P(A_{i_k})$$

teljesül tetszőleges

$$A_{i_1} \in \mathcal{A}_{i_1}, A_{i_2} \in \mathcal{A}_{i_2}, \dots, A_{i_k} \in \mathcal{A}_{i_k}$$

eseményekre, $1 \le i_1 < i_2 < \dots < i_k \le n$ indexsorozatra és minden $2 \le k \le n$ számra.

Def. Az $X_1,\dots,X_n$ valószínűségi változók függetlenek, ha az előző tulajdonság teljesül

$$A_{i_j} = \{\omega : X_{i_j}(\omega) \in B_{i_j}\}$$

választással (ez matematikailag azt jelenti, hogy az $\mathcal{F}_{X1}$, $\mathcal{F}_{X2}$, …, $\mathcal{F}_{Xn}$ generált σ-algebrák függetlenek).

A függetlenség karakterizációi

- Ha $\underline{X}$ koordinátái függetlenek, akkor definíció szerint
- $F_{\underline{X}}(\underline{z})=P(X_1<z_1, X_2< z_2,\dots, X_d<z_d)=F_1(z_1)F_2(z_2)\dots F_d(z_d)$ (minden $\underline{z} \in \boldsymbol{R}^d$ –re). Meg is fordítható: F szorzatelőállításából következik a függetlenség.
- Deriválva: a függetlenség abszolút folytonos változókra ekvivalens a sűrűségfüggvény $f_{\underline{X}}(\underline{z})=f_1(z_1)f_2(z_2)\dots f_d(z_d)$ alakú előállításával is.
- Példa: az egységnégyzeten egyenletes eloszlás sűrűségfüggvénye ($f(\underline{z})=1$ ha $0<\underline{z}<1$) előáll $f_1(z_1)f_2(z_2)$ alakban, ahol $f_i(z_i)=1$, ha $0<z_i<1$ $(i=1,2)$, ez éppen a [0,1] intervallumon egyenletes eloszlás.

Tulajdonságok

- 1. Az $X_1,\dots,X_n$ diszkrét valószínűségi változók függetlenek, ha $P(X_1 = x_1,\dots, X_n = x_n)=P(X_1 = x_1)\cdot\dots\cdot P(X_n = x_n)$ teljesül minden $x_1,\dots,x_n$ értékre.
- 2. Ha az $X_1,\dots,X_n$ valószínűségi változók függetlenek, a $g_1,\dots,g_n$ függvények Borel-mérhetőek, akkor $g_1(X_1),\dots, g_n(X_n)$ is függetlenek.
- 3. Ha az $X_1,\dots,X_n$ valószínűségi változók függetlenek, a h k-változós Borel-mérhető fv., akkor $h(X_1,\dots, X_k)$, X_{k+1} ,…,X_n is függetlenek.

Függvény eloszlása

- Tfh. az A tartományon adott egy g, folytonosan deriválható függvény, melynek létezik inverze. Ha az $\underline{X}$ abszolút folytonos változó értékei A-beliek, akkor g($\underline{X}$) is abszolút folytonos, sűrűségfüggvénye

$$f_{g(X)}(\underline{y}) = f_X(g^{-1}(\underline{y}))\,|J|$$

ahol J a g^{-1} függvény Jacobi determinánsának abszolút értéke.

Példák

- Számoljuk ki (X+Y,X-Y) együttes sűrűségfüggvényét, ha X,Y független, standard normálisok!
- Adódik, hogy az összeg (és a különbség is) normális eloszlású!
- X+Y és X-Y független is (ez csak a normális eloszlásra igaz).

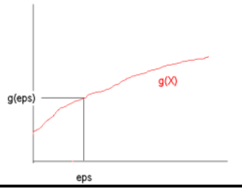
Markov-típusú egyenlőtlenségek

- Legyen $X \geq 0$ valószínűségi változó, g: $\boldsymbol{R}_+ \to \boldsymbol{R}_+$ monoton növő. Ekkor

$$P(X \geq \varepsilon) \leq E(g(X))/g(\varepsilon).$$

- Bizonyítás.

$E(g(X)) \geq g(\varepsilon)P(X \geq \varepsilon)$
mert $X \geq \varepsilon$ eseményen
$g(X)) \geq g(\varepsilon)$



Alkalmazások

- $g(x)=x$: Ha $X \geq 0$ valószínűségi változó, akkor $P(X \geq \varepsilon) \leq E(X)/\varepsilon$ (ezt nevezik Markov egyenlőtlenségnek).
- $g(x)=x$, X helyett $(X-EX)^2$-re alkalmazva: $P((X-EX)^2 \geq \varepsilon^2) \leq E(X-EX)^2/\varepsilon^2$, ami egyszerűsítve $P(|X-EX| \geq \varepsilon) \leq D^2(X)/\varepsilon^2$ (elnevezés:Csebisev egyenlőtlenség).
- Megjegyzés. A fenti egyenlőtlenségek élesek, azaz minden ε-ra megadható olyan valószínűségi változó, amelyre $P(X \geq \varepsilon)=E(X)/\varepsilon$. (Két értéket vesz fel: 0, ε).

Alkalmazások

- Az eredmények a gyakorlatban mégsem adnak kellően pontos becslést, mert a tényleges eloszlások tulajdonságait nem veszi figyelembe. Ezért, ha ismerjük az adott változó eloszlását, mindig abból adjuk meg a valószínűségek értékét.
- Példa: hányszor kell egy szabályos érmét feldobni, hogy a fejek relatív gyakorisága legalább 0.99 valószínűséggel ne térjen el 0.05-nél jobban 0.5-től? Csebisev egyenlőtlenségből: $P(|X-0.5| \geq 0.05) \leq D^2(X)/\varepsilon^2 = 400/4n \leq 0.01$ elég, amiből $n \geq 10000$ adódik. A binomiális eloszlásból adódó pontos érték: 670. Szimuláció

Nagy számok törvényei

- Legyenek $X_1, X_2, \ldots, X_n$ független, azonos eloszlású valószínűségi változók. Tegyük fel, hogy $\sigma^2 = D^2(X)$ véges ($m := E(X_i)$). Ekkor minden ε, $\delta > 0$-hoz megadható olyan n_0, hogy $n > n_0$ esetén $P(|(X_1+X_2+\ldots+X_n)/n - m| \geq \varepsilon) \leq \delta$.
- Bizonyítás. $E(X_1+X_2+\ldots+X_n)/n = m$, és $D^2[(X_1+X_2+\ldots+X_n)/n] = \sigma^2/n$. A Csebisev egyenlőtlenség miatt $P(|(X_1+X_2+\ldots+X_n)/n - m| \geq \varepsilon) \leq \sigma^2/\varepsilon^2 n$, ami 0-hoz tart, azaz elég nagy n-re kisebb lesz δ-nál.
- Elnevezés: $(X_1+X_2+\ldots+X_n)/n \to m$ sztochasztikusan (az ilyen konvergenciát bizonyító tételeket gyenge tételnek nevezzük).

Megjegyzések.

- A tétel feltételei gyengíthetőek: elég, ha a független, azonos eloszlású változók várható értéke véges.
- Az állítás is erősíthető: 1 valószínűségű konvergencia is bizonyítható (ez azt jelenti, hogy

$$P\{\omega: (X_1+X_2+\ldots+X_n)/n \to m\} = 1.$$

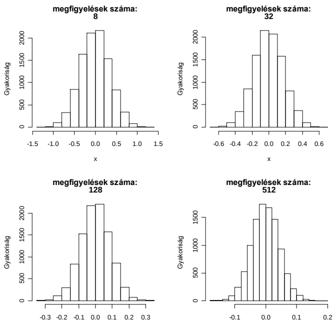
- Ha $\Omega=[0,1]$ és P(A) az A "hossza", akkor az 1 valószínűségű konvergencia lényegében a szokásos pontonkénti konvergencia. Ez nem következik a sztochasztikus konvergenciából:

  Legyen pl. $X_{2^n+k}(z)=1$, ha $k/2^n < z < (k+1)/2^n$ és 0 különben. Ekkor $X_n \to 0$ sztochasztikusan, de $P\{z: X_n \to 0\} = 0$

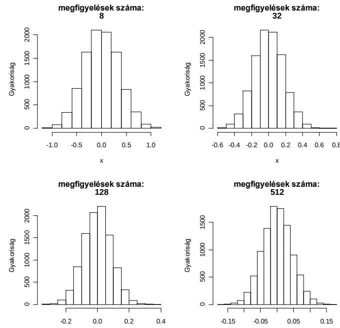
Bernoulli tétele

- A nagy számok törvényének legelső verzióját még Bernoulli bizonyította, indikátorváltozókra: eszerint azonos körülmények között elvégzett független kísérleteknél tetszőleges esemény relatív gyakorisága tart az esemény valószínűségéhez. (Az előző speciális esete: X indikátorváltozó.)

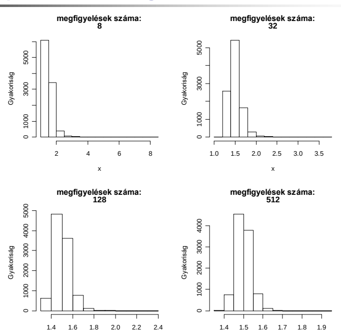
Független, standard normális eloszlású megfigyelések átlaga



Független, azonos, egyenletes eloszlású megfigyelések átlaga (m=0, σ=1)



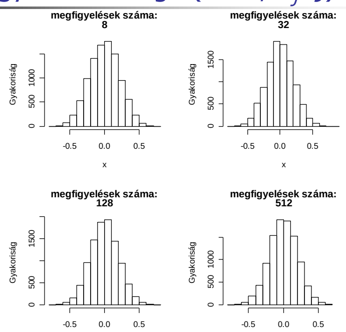
Független, azonos, Pareto (k=3) eloszlású megfigyelések átlaga



Összefüggő változók

- Valami feltétel kell: $X_1=X_2=\ldots=X_n$ esetén $(X_1+X_2+\ldots+X_n)/n = X_1$ és így nem konvergál konstanshoz.
- Tétel (Bernstein). Legyen $(X_1, X_2, \ldots, X_n)$ olyan, hogy $D^2(X_1)+D^2(X_2)\ldots+ D^2(X_n)<kn$, valamint tegyük fel, hogy van olyan $h: \boldsymbol{N}\to \boldsymbol{R}_+$ függvény, melyre $|R(X_i,X_j)|\leq h(|i-j|)$ és $\frac{1}{n}\sum_{i=1}^{n} h(i) \to 0$
  ekkor $(X_1+X_2+\ldots+X_n)/n - (m_1+m_2+\ldots+m_n)/n \to 0$ sztochasztikusan.

Független, egyenletes eloszlású megfigyelések átlaga (m=0, $\sigma_j=j$)



Összefoglalás

- A konvergencia sebessége a független esetben a szórásnégyzettől függ.
- Összefüggő esetben pedig a korreláció a döntő.
- A törvény nem jelenti azt, hogy az eddig nem szerepelt értékek a jövőben a vártnál gyakoribbak lesznek, hanem csupán az eloszlás szerint kapott nagyszámú érték állítja helyre a várt gyakoriságokat.