Valószínűségszámítás előadás informatika BSC/A szakosoknak és matematika BSC-seknek

2019/2020 1. félév

Zempléni András

zempleni@caesar.elte.hu

http://zempleni.elte.hu

Valószínűségi vektorváltozók

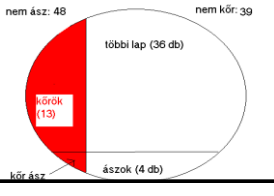
- Gyakran nem csak egy mennyiséget vizsgálunk a kísérletünk/megfigyelésünk során.
- Példa: hőmérséklet, csapadék, szélerő stb.
- $\underline{X}: \Omega \to \boldsymbol{R}^n$ függvény valószínűségi vektorváltozó, ha $\{\omega: \underline{X}(\omega) \in B\} \in \mathcal{A}$ minden B n-dimenziós Borel halmazra.
- $\underline{X} = (X_1,\ldots,X_n)$ pontosan akkor valószínűségi vektorváltozó, ha a koordinátái valószínűségi változók.

Vektorváltozók eloszlása

- Legyen B tetszőleges n-dimenziós Borel halmaz.
  $Q_{\underline{X}}(B) := P\{\omega: \underline{X}(\omega) \in B\}$ valószínűséget ad meg $\boldsymbol{R}^n$ Borel halmazain. Ez az $\underline{X}$ *eloszlása*.
- Speciális eset: ha diszkrét, akkor a $p_i := P(\underline{X} = \underline{x}_i)$ valószínűségek meg is határozzák $\underline{X}$ eloszlását.
- A kétdimenziós esetben a diszkrét valószínűségi változó eloszlása legegyszerűbben táblázattal adható meg:

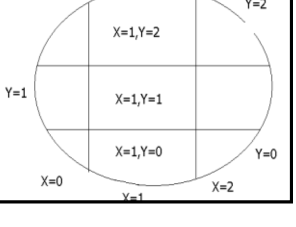
Példa: kétszer húzunk visszatevéssel egy francia kártyacsomagból

kőr\ ász	0	1	2
0	$(36/52)^2$	$2\cdot3\cdot36/(52)^2$	$(3/52)^2$
1	$2\cdot12\cdot36/(52)^2$	$2\cdot(12\cdot3+1\cdot36)/(52)^2$	$2\cdot3/(52)^2$
2	$(12/52)^2$	$2\cdot12/(52)^2$	$(1/52)^2$



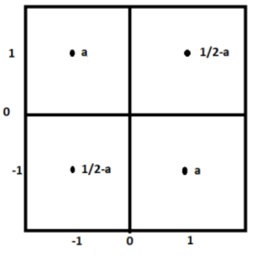
A peremeloszlások

- (X,Y) eloszlásából (elnevezés: *együttes eloszlás*) következtethetünk az egyes változók eloszlására:
  P(X=1)=P(X=1,Y=0)+P(X=1,Y=1)+P(X=1,Y=2)
- Az egyes koordináták (alacsonyabb dimenziós vektorok) eloszlása: *peremeloszlás*.
- Az együttes eloszlás tehát meghatározza a peremeloszlást (a megfordítás természetesen nem igaz) .



Kapcsolat az együttes- és a peremeloszlás között

- A peremeloszlások viszont nem határozzák meg az együttest:
  (tetsz. |a|<1/4-re
  P(X=1)=P(X=-1)=1/2
  P(Y=1)=P(Y=-1)=1/2)
- Kivétel: ha a komponensek függetlenek



Példa

- Polinomiális eloszlás: a kísérletünknek r különböző kimenetele lehet: $p_1, p_2, \ldots, p_r$ valószínűségűek ($p_1+p_2+\ldots+p_r=1$). n független, azonos körülmények között végrehajtott kísérlet során az i-edik esemény bekövetkezésének száma X_i. Az együttes eloszlásuk:

$$P(X_1 = k_1, \ldots, X_r = k_r) = \frac{n!}{k_1! k_2! \ldots k_r!} p_1^{k_1} p_2^{k_2} \ldots p_r^{k_r} \quad (k_1 + k_2 + \ldots + k_r = n)$$

Tulajdonságok

- Speciális esetek:
  - r=1 triviális
  - r=2: binomiális eloszlás (X_1, $X_2=n-X_1$)
  - Az egydimenziós peremeloszlások (X_i eloszlása) binomiális (n, p_i) paraméterrel
- Példa: n kockadobásnál az egyes értékek gyakorisága (itt r = 6).
- A koordináták nem függetlenek!

  (Pl. X_1 = n esetén a többi koordináta 0.)

Összeg várható értéke

- X, Y tetszőleges, véges várható értékűek. Ekkor E(X+Y)= E(X)+E(Y). Bizonyítás a diszkrét esetre:

$$E(X+Y) = \sum_{k,m} (x_k + y_m) P(X = x_k, Y = y_m) = \sum_{k,m} x_k P(X = x_k, Y = y_m) +$$

$$+ \sum_{k,m} y_m P(X = x_k, Y = y_m) = \sum_k x_k P(X = x_k) + \sum_m y_m P(Y = y_m) = E(X) + E(Y).$$

- Indukcióval: $E(X_1+X_2+\ldots+X_n)= E(X_1)+E(X_2)+\ldots+E(X_n)$.

Alkalmazások

- A binomiális eloszlás várható értéke:

  $X= X_1 + X_2 \ldots + X_n$ ahol X_i az i-edik kísérletnél az A esemény indikátora.
- Az előző tulajdonság alapján

  $E(X)= E(X_1 + X_2 \ldots + X_n)=n \cdot E(X_1)= n \cdot p$.
- Ugyanígy a hipergeometrikus eloszlás várható értéke is $n \cdot p$ ($p=M/N$ a selejtarány).

Névjegy probléma

- n ember bedobja a névjegyét egy kalapba, ezután mindenki húz egyet véletlenszerűen. X: azon személyek száma, akik a saját névjegyüket húzzák.
- X_i: az i-edik ember a saját névjegyét húzza. $E(X_i)=P(X_i=1)= 1/n$.
- $X= X_1 + X_2 \ldots + X_n$ és így a várható érték additivitása alapján

  $E(X)= E(X_1 + X_2 \ldots + X_n)=n \cdot E(X_1)= n \cdot 1/n=1$.
- Szimuláció

A feltételes várható érték

- Kocka-érme kísérlet: annyi érmével dobunk, amennyit dobtunk egy kockával. Ha a kockával i a dobott szám, az érméknél kapott fejek száma binomiális (i,1/2) paraméterrel. Ekkor tehát a várható fej-szám $i/2$.
- Formálisan/általánosan: ez a feltételes eloszlás várható értéke

$$\sum_i x_i P(X = x_i \mid A) = E(X \mid A)$$

- Megjegyzés. Ha E(X) létezik, akkor E(X|A) is.

A teljes várható érték tétele

- Tegyük fel, hogy E(X) létezik, és legyen B_1, B_2, ..., pozitív valószínűségű eseményekből álló teljes eseményrendszer. Ekkor

$$E(X) = E(X \mid B_1)P(B_1) + E(X \mid B_2)P(B_2) + \ldots$$

- *Bizonyítás.*

$$\sum_i E(X \mid B_i)P(B_i) = \sum_i \sum_k x_k P(X = x_k \mid B_i)P(B_i)$$

Az abszolút konvergencia miatt az összegzés sorrendje felcserélhető:

$$= \sum_k x_k \sum_i P(X = x_k \mid B_i)P(B_i) = \sum_k x_k P(X = x_k) = E(X)$$

Példák

- Legyen $X_1, X_2, \ldots, X_n$ független, azonos eloszlású. N pedig tőlük független, pozitív egész értékű valószínűségi változó. Ekkor $Y=X_1+X_2+\ldots+X_N$ esetén $EY=E(N)E(X_1)$.
- Pascal eloszlás várható értéke.
- Érme-kocka kísérlet
- Kocka-érme kísérlet

Valószínűségi változók szórásnégyzete

- Nem mindegy, hogy mekkora a vizsgált véletlen mennyiség ingadozása.
- Jobb, ha a buszok pontosan 10 percenként jönnek, mintha időnként 3 jön egymás után, és aztán 30 percet kell várni.
- Az ingadozás számszerűsítése: a várható értéktől vett átlagos négyzetes eltérés, elnevezése: szórásnégyzet. Formálisan:

$$D^2(X) := E[(X-E(X))^2].$$

- Kiszámítása: $D^2(X) = E[X^2-2XE(X)+E^2(X)] = E(X^2)-2E(X)E(X)+E^2(X)$ a várható érték linearitása miatt. Azaz

$$D^2(X)=E(X^2)-E^2(X).$$

Tulajdonságok

- $D^2(X) \geq 0$, mert nemnegatív valószínűségi változó várható értéke.
- $D^2(aX+b)=a^2D^2(X)$, mert $D^2(aX+b)= =E[(aX+b-E(aX+b))^2]= E[(aX+b-aE(X)-b)^2]= =E[(aX-aE(X))^2]=a^2 E[(X-E(X))^2]$.
- Abból, hogy E(X) véges, még nem következik $D^2(X)$ végessége, hiszen ha $P(X=k)=c/k^3$ (egyértelműen megadható olyan c, amire ez eloszlás lesz) akkor E(X) véges, de $E(X^2)=c(1+1/2+\ldots+1/k+\ldots)$, ami végtelen.