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

2019/2020 1. félév

Zempléni András

zempleni@caesar.elte.hu

http://zempleni.elte.hu

Példák

- Az elfajult eloszlás szórásnégyzete:

$D^2(X)= E(X^2)-E^2(X)=c^2-c^2=0.$

- Megfordítás: ha $D^2(X)= 0$, akkor X 1 valószínűséggel konstans.
  Biz.: $(X-E(X))^2 \geq 0$, várható értéke 0, tehát ő maga is 1 valószínűséggel 0, azaz $X=E(X)$ 1 valószínűséggel.
- A p valószínűségű A esemény indikátorának szórásnégyzete:
  $D^2(X)= E(X^2)-E^2(X)= p - p^2= p(1-p)$. Azaz $p=0.5$ esetén a legnagyobb a szórásnégyzet.
- A kockadobás szórásnégyzete:
  $D^2(X)= E(X^2)-E^2(X)=(1+4+...+36)/6 - 49/4=91/6 - 49/4 =(182-147)/12=35/12.$

Az (n,p) paraméterű binomiális eloszlás szórásnégyzete

$$E(X^2)=\sum_{k=1}^{n} k^2\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k=1}^{n} npk\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}=$$

$$=\sum_{k=1}^{n} np(k-1+1)\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}=$$

$$=\sum_{k=2}^{n} n(n-1)p^2\binom{n-2}{k-2}p^{k-2}(1-p)^{n-k}+\sum_{k=1}^{n} np\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}=$$

$$=n(n-1)p^2+np$$

$$D^2(X)=E(X^2)-E^2(X)=n(n-1)p^2+np-n^2p^2=np-np^2=np(1-p)$$

Példák 2.

A Poisson eloszlás szórásnégyzete:

$$E(X^2)=\sum_{k=1}^{\infty} k^2\lambda^k\frac{e^{-\lambda}}{k!}=\sum_{k=1}^{\infty} k\lambda^k\frac{e^{-\lambda}}{(k-1)!}=\lambda\sum_{k=1}^{\infty}(k-1+1)\lambda^{k-1}\frac{e^{-\lambda}}{(k-1)!}=$$

$$=\lambda^2\sum_{k=2}^{\infty}\lambda^{k-2}\frac{e^{-\lambda}}{(k-2)!}+\lambda=\lambda^2+\lambda.$$

Ebből

$$D^2(X)=\lambda^2+\lambda-\lambda^2=\lambda.$$

Azaz a Poisson eloszlás várható értéke és szórásnégyzete megegyezik.

A szórás

- Szórásnégyzet mértékegysége az eredeti X mértékegységének a négyzete (azaz pl. a buszok követési időközénél négyzetperc). Ez nem teszi egyszerűvé interpretációját.
- Szórás: $D(X)$ a szórásnégyzet pozitív négyzetgyöke. Ez már a megfelelő mértékegységű, $D(aX)=|a|D(X)$.

Összeg szórásnégyzete

- $D^2(X+Y)=E[(X+Y-E(X+Y))^2]=$
  $E[(X-E(X)+Y-E(Y))^2]=E[(X-E(X))^2]+E[(Y-E(Y))^2]+$
  $+2E[(X-E(X))(Y-E(Y))]=D^2(X)+D^2(Y)+$
  $+2E[(X-E(X))(Y-E(Y))]$
- Példák:
  - $X=Y$ esetén $D^2(X+Y)= D^2(2X)=4\,D^2(X)$
  - $X=-Y$ esetén $D^2(X+Y)= D^2(0)=0$

azaz nem csak X és Y egydimenziós eloszlásától, hanem az együttes viselkedésüktől, azaz az együttes eloszlásuktól is függ az összegük szórásnégyzete.

A független val. változók esete

- *Állítás.* ha X,Y függetlenek és véges várható értékűek, akkor E(XY)=E(X)E(Y).
- *Bizonyítás.*

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

ami a függetlenség miatt így írható:

$$=\sum_k x_k P(X=x_k)\sum_m y_m P(Y=y_m)=E(X)E(Y).$$

Kovariancia

- Definíció. Az X és Y kovarianciája:

cov(X,Y):=E[(X-E(X))(Y-E(Y))]

- Kiszámítása: cov(X,Y)= E[XY-XE(Y)-YE(X)+E(X)E(Y)]=E(XY)-E(X)E(Y)
- Az előzőek értelmében cov(X,Y)=0, ha X és Y függetlenek.
- Megj.: Abból, hogy cov(X,Y)=0 nem következik, hogy függetlenek: legyen X szimmetrikus a 0=ra (pl. P(X=1)=P(X=-1)=P(X=0)=1/3) és $Y=X^2$. Ekkor cov(X,Y)=E(X^3)-E(X)E(X^2)=0-0, hiszen E(X^3)=E(X)=0.
- A kovariancia szimmetrikus: cov(X,Y)= cov(Y,X)
- cov(X,X)= D^2(X)

Összeg szórásnégyzete 2

- D^2(X+Y)=D^2(X)+D^2(Y)+2cov(X,Y)
- Speciálisan: D^2(X+Y)=D^2(X)+D^2(Y), ha X és Y függetlenek (elég, hogy cov(X,Y)=0).
- n tagú összegre:

$$D^2(X_1+\ldots+X_n)=\sum_{i=1}^{n} D^2(X_i)+2\sum_{1\le i<j\le n}\operatorname{cov}(X_i,X_j)$$

- Spec.: $D^2(X_1+X_2+\ldots+X_n)= D^2(X_1)+D^2(X_2)+\ldots+D^2(X_n)$, ha a tagok páronként függetlenek.

Korrelációs együttható

- A kovariancia skálafüggő: cov(*a*X,*b*Y)=*ab*cov(X,Y)
- A változók közötti lineáris kapcsolat erősségét mérő mennyiség a *korrelációs együttható*:

$$R(X,Y)=\frac{\operatorname{cov}(X,Y)}{D(X)D(Y)}$$

- Tulajdonságai:
  - R(X,Y)=0, ha X és Y függetlenek (ez sem fordítható meg)
  - Ez alapján definíció szerint legyen R(X,Y)=0, ha X vagy Y elfajult eloszlású.
  - R(X,aX+b)=1, ha a>0, mert cov(X,aX+b)=aD^2(X).

A korreláció tulajdonságai

- |R(X,Y)|≤1 és |R|=1 akkor és csak akkor, ha X=aY+b 1 valószínűséggel (a≠0, b∈***R***).
- Ehhez:

$$X^*=\frac{X-E(X)}{D(X)},\quad Y^*=\frac{Y-E(Y)}{D(Y)}$$

a standardizált változók. E(X*)=E(Y*)=0, D(X*)=D(Y*)=1. R(X,Y)=E(X*Y*).
0≤E $(X^*\pm Y^*)^2$=E(X^{*2})±2E(X*Y*)+E(Y^{*2}) =2± 2E(X*Y*), tehát |R(X,Y)|≤1.
Ebből: R=1 akkor és csak akkor, ha 0=E $(X^*-Y^*)^2$, azaz X*=Y* 1 valószínűséggel. Ekkor X=aY+b , a>0.
R=-1 akkor és csak akkor, ha 0=E $(X^*+Y^*)^2$, azaz X*=-Y* 1 valószínűséggel. Ekkor X=aY+b , a<0.

Példák

- A polinomiális eloszlás koordinátái közötti korreláció: n=1-re. (Állítás: ez ugyanaz minden n-re.)
- E(X_1)=p_1, E(X_2)=p_2, E(X_1X_2)=0,
- $D^2(X_1)= p_1(1-p_1)$, $D^2(X_2)= p_2(1-p_2)$. Ebből

$$R(X_1,X_2)=\frac{E(X_1X_2)-E(X_1)E(X_2)}{D(X_1)D(X_2)}=\frac{-p_1p_2}{\sqrt{p_1(1-p_1)p_2(1-p_2)}}$$

- Spec.: p:=p_1=p_2, esetén R=- p/(1-p).

Y közelítése X függvényével

- Gyakori eset, hogy nem ismerjük a számunkra érdekes mennyiség (Y) pontos értékét (pl. holnapi részvény-árfolyam, vízállás, időjárás). Van viszont információnk hozzá kapcsolódó mennyiségről (X, mai értékek).
- Feladat: olyan f_0 megtalálása, amelyre $f_0(X)$ a lehető legjobb közelítése Y-nak.
- Matematikailag: f_0 a megoldása a $\min_f E(Y-f(X))^2$ szélsőérték-problémának (legkisebb négyzetes becslés).
- Ha az együttes eloszlás ismert (nem teljesen reális, de a megfigyelések alapján közelíthető), akkor megoldható a feladat.

A várható érték optimumtulajdonsága

Állítás. A $\min_a E(Y-a)^2$ feladat megoldása a=E(Y).

Bizonyítás. $E(Y-a)^2= E(Y^2)-2aE(Y)+a^2$

a szerint deriválva adódik, hogy valóban E(Y) a minimumhely.

A minimum értéke $D^2(Y)$.

Ugyanígy: X tetszőleges értéke esetén E(Y|X=x) adja a minimumot.

Példa

- Annyi érmével dobtunk újra, amennyi fejet kaptunk 2 érmével dobva. Csak azt tudjuk, hogy hány fejet kaptunk a második dobásnál. Közelítsük ennek segítségével az első dobás eredményét.
- Például F=0 esetre:

$$E(X\mid F=0)=\frac{\sum_{i=0}^{2} iP(X=i,F=0)}{P(F=0)}=\frac{\sum_{i=0}^{2} iP(F=0\mid X=i)P(X=i)}{\sum_{i=0}^{2} P(F=0\mid X=i)P(X=i)}$$

- Az eredmények: E(X|F=2)=2, E(X|F=1)=4/3, E(X|F=0)=2/3.

Optimum a lineáris függvények körében

$$\min_{a,b} E[Y-(aX+b)]^2$$

- Egyszerűbben megoldható
- Nem kell az együttes eloszlás
- A megoldás deriválással:

$$E[Y-(aX+b)]^2=E(Y^2)+a^2E(X^2)+2abE(X)+ +b^2-2aE(XY)-2bE(Y)$$

$$\frac{\partial}{\partial a}E[Y-(aX+b)]^2=2aE(X^2)+2bE(X)-2E(XY)=0$$

$$\frac{\partial}{\partial b}E[Y-(aX+b)]^2=2b+2aE(X)-2E(Y)=0$$

$$aE(X^2)=E(XY)-bE(X) \qquad b=E(Y)-aE(X)$$

$$aE(X^2)=E(XY)-(E(Y)-aE(X))E(X)$$

$$a=\frac{E(XY)-E(X)E(Y)}{E(X^2)-E^2(X)} \qquad b=E(Y)-\frac{E(XY)-E(X)E(Y)}{E(X^2)-E^2(X)}E(X)$$

Az aX+b egyenes tulajdonságai

- Ez a legkisebb négyzetes eltérést adó a lineáris függvények között (a fenti megoldás valóban minimum)
- Elnevezés: regressziós egyenes
- Átmegy az (E(X),E(Y)) ponton
- Példa: Kockával dobunk, majd ha *k* az eredmény, az 1,...,*k* cédulák közül húzunk egyet. Nem tudjuk a húzás eredményét, csak a kockadobásét. Hogyan tippeljünk a húzott számra (a legkisebb négyzetes eltérést adó becslést keressük)? E(h|K=k)=(k+1)/2 az univerzálisan legjobb közelítés, tehát a legjobb lineáris közelítés is.