Valószínűségszámítás és statisztika előadás informatika BSC/B szakosoknak

2021/2022 2. félév

Zempléni András

andras.zempleni@ttk.elte.hu

http://zempleni.elte.hu

Események függetlensége

- Ha a B esemény bekövetkezése nem befolyásolja az A valószínűségét, azaz $P(A|B)=\mathrm{P}(A)$, akkor azt mondjuk, hogy az A és B függetlenek. Ez így nem ideális definíció (nem szimmetrikus,

  $P(B)>0$ kell hozzá), ezért
- Definíció. Az A és B *események függetlenek*, ha $P(A\cap B)=P(A)P(B)$.

Tulajdonságok

- Ha A és B diszjunktak, akkor csak triviális ($P(A)=0$ vagy $P(B)=0$) esetben függetlenek.
- Ha A és B függetlenek, akkor komplementereik is függetlenek.
- Önmaguktól csak a triviális események függetlenek.
- $A \subset B$ esetén csak akkor függetlenek, ha legalább az egyik triviális.
- Tipikus eset függetlenségre: A az első, B a második kísérlet eredménye.

Valószínűségi változók 1.

- A legtöbbször nem maga a kísérlet kimenetele (a realizálódott elemi esemény) hanem egy számszerűsíthető eredmény az érdekes.
- Példa: ipari termelés – minőségellenőrzés: a kérdés az esetleges selejtesek száma, nem pedig az, hogy pontosan melyik elemeket is választottuk.
- Sok gyakorlati esetben nem is adódik természetesen az Ω halmaz (pl. időjárás megfigyelés).

Valószínűségi változók 2.

- Mintavételi példa: N termék, n elemű minta visszatevés nélkül. Ω elemszáma: $\binom{N}{n}$
- Selejtesek száma (X): 0 és n közötti szám.
- Matematikailag: $X : \Omega \to \boldsymbol{R}$ függvény
- Feltétel: legyen értelme pl. annak a valószínűségéről beszélni, hogy $X<a$. Azaz $\{\omega: X(\omega)<a\} \in \mathcal{A}$ kell, hogy teljesüljön minden a–ra. Hasonlóképpen más „természetes" feltételnek is legyen valószínűsége.

Diszkrét valószínűségi változók

- Definíció: az *X diszkrét valószínűségi változó*, ha értékkészlete (x_1,…, x_n …) legfeljebb megszámlálható.
- A valószínűségi változó definíciójából adódóan $\{\omega: X(\omega)= x_i\}=\{X=x_i\}\in\mathcal{A}$ azaz $p_i := P(X=x_i)$ értelmes. Ezek meg is határozzák X eloszlását.
- Véges vagy megszámlálható valószínűségi mezőn minden valószínűségi változó diszkrét.
- Általában nem célszerű a természetszerűen folytonos értékkészletű X diszkretizálása (egyszerűbbek a folytonos modellek).

Példák

- $X(\omega)=c$ minden ω-ra.

Elnevezés: elfajult eloszlás.

$P(X=c)=1$.

- X akkor 1, ha egy adott, p valószínűségű A esemény bekövetkezik és 0 különben (elnevezés: az A *esemény indikátora*).

  $P(X=0)=1-p$

  $P(X=1)=p$

Példák 2.

- Mintavételnél legyen *X* a mintában levő selejtesek száma.
  - Visszatevéses esetben (binomiális eloszlás):

$$P(X=k)=\binom{n}{k}\left(\frac{M}{N}\right)^{k}\left(1-\frac{M}{N}\right)^{n-k}\quad (k=0,\dots,n)$$

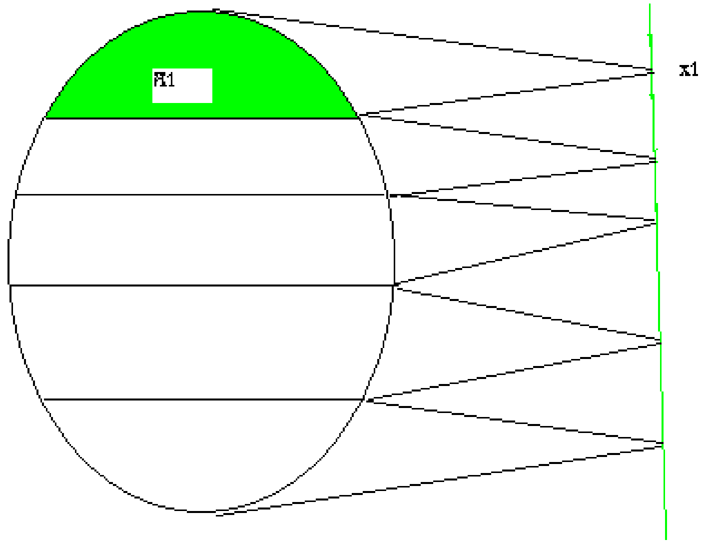
  - Visszatevés nélküli esetben:

(hipergeometriai eloszlás)

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\quad (k=0,\dots,n)$$

Teljes eseményrendszer

- Ha X diszkrét valószínűségi változó, akkor az $A_i=\{\omega : X(\omega)= x_i\}$ események teljes eseményrendszert alkotnak.



Valószínűségi változók függetlensége

- X és Y diszkrét valószínűségi változók függetlenek, ha

$$P(\{X = x_i\} \cap \{Y = y_k\}) = P(X = x_i) P(Y = y_k)$$

  teljesül minden i, k értékre. (Azaz az X-hez és az Y-hoz tartozó teljes eseményrendszerek függetlenek.)
- Megjegyzések:
  - az elfajult eloszlású valószínűségi változó minden valószínűségi változótól független.
  - Önmagától csak az elfajult eloszlású valószínűségi változó független.

Példák

- Függetlenek
  - Visszatevéses mintavételnél az egyes kísérletek eredményei
    - egymás utáni kockadobások
- Nem függetlenek
  - Visszatevés nélküli mintavételnél az egyes kísérletek eredményei
    - lottóhúzásnál az egymás utáni számok
- Kérdés: Két kockadobás összege független-e ugyanezen két dobás különbségétől?

Binomiális eloszlás alkalmazása

- Visszatevéses mintavétel más realizációja: független kísérletek azonos körülmények között. $P(A)=p$ esemény, végezzünk n (rögzített számú) független kísérletet.
- X: az A bekövetkezésének gyakorisága (pontosan hányszor jött ki az A). X eloszlása binomiális (n,p).
- $X= X_1 + X_2 \ldots + X_n$ ahol X_i az i-edik kísérletnél az A esemény indikátora. Ezek az indikátorok függetlenek is!

Geometriai (Pascal) eloszlás

- Független kísérletek azonos körülmények között. $P(A)=p>0$ valószínűségű esemény. Addig végzünk kísérleteket, míg A be nem következik.
- X: az első sikeres kísérlet sorszáma.

$p_k=P(X=k)=p(1-p)^{k-1}$ $(k=1,2,\ldots)$

Valóban valószínűségeloszlás $(p_1+p_2+\ldots=1)$

Azaz 0 a valószínűsége annak, hogy sosem kapunk fejet

Poisson eloszlás

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$$ (k=0,1,2,...; λ>0 paraméter). Valóban eloszlás. Grafikusan

Állítás. Ha a binomiális eloszlás paramétereire

$n \to \infty$ úgy, hogy $np \to \lambda$,

akkor a határérték éppen a λ paraméterű Poisson eloszlás.

Bizonyítás. $$\binom{n}{k} p^k (1-p)^{n-k} \approx \binom{n}{k} \frac{\lambda^k}{n^k} \left(1-\frac{\lambda}{n}\right)^{n-k} \to \frac{\lambda^k e^{-\lambda}}{k!}$$

Alkalmazások

- Első példa: lórugás áldozatainak száma a porosz hadseregben.
- Poisson folyamat: időben lejátszódó folyamatnál adott [*a*,*b*) intervallumba eső események száma ($X_{a,b}$) éppen $\lambda(b-a)$ paraméterű Poisson eloszlású, ha a folyamat
  - homogén: $X_{a,a+t}$ eloszlása csak t-től függ;
  - utóhatás nélküli: $X_{a,b}$ és $X_{b,c}$ függetlenek ha $a<b<c$;
  - nemelfajuló: $0<P(X_{a,b}=0)<1$.

A Poisson folyamat gyakorlati alkalmazásai

- Balesetek száma egy adott útszakaszon
- Adott értéknél erősebb viharok száma egy adott területen
- Számítógépes rendszer meghibásodásainak száma

Összefoglalás (diszkrét eloszlások)

- Binomiális eloszlás
  - Rögzített számú kísérletnél adott esemény gyakorisága (pl. 10 kockadobásból a hatosok száma)
  - Nagy mintaelemszámra, kicsi valószínűségnél a Poisson eloszlással közelíthető
- Pascal eloszlás
  - Addig kísérletezünk, míg egy adott esemény be nem következik, az első sikeres sorszáma (pl. az első hatost hányadik kockadobásnál kapjuk meg)
- Hipergeometriai eloszlás
  - Visszatevés nélküli mintavételnél adott típusú mintaelemek száma (pl. lottóhúzásnál az 5 találat valószínűsége)

Diszkrét valószínűségi változók várható értéke

- Szerencsejátékban a pontos nyeremény nem látható előre. De: az átlagos nyereményről szeretnénk tudni. (Kedvező-e a játék? Fair játék: az ár éppen a várható érték.)
- Példa: Dobókocka: annyi a nyereményünk, amennyit dobunk. Ennek átlagos értéke 1/6(1+2+…+6)=21/6=3.5
- De ha nem szabályos a kocka, például az egyes helyett is 6 van, akkor az átlagos nyeremény 1/6(2+…+5)+6/3=13/3.
- *Definíció*. A $p_i = P(X = x_i)$ eloszlással megadott valószínűségi változó *várható értéke* $\mathrm{E}(X) := p_1 x_1 + p_2 x_2 + \ldots$, ha a sor abszolút konvergens.

Példák

- Az elfajult eloszlás várható értéke:

$E(X) = cP(X = c) = c.$

- A p valószínűségű A esemény indikátorának várható értéke: $E(X) = 1P(X = 1) = p$
- *Az* $x_1, x_2, \ldots x_n$ számokon egyenletes eloszlás (mindegyik valószínűsége 1/n) várható értéke a számok számtani közepe.
- Az (n,p) paraméterű binomiális eloszlás várható értéke:

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

- Amerikai rulett. Ha k számra teszünk, a nyereményünk 36/k. A várható nettó nyeremény $(36/k)\cdot(k/38)-1= -2/38$.

Tulajdonságok

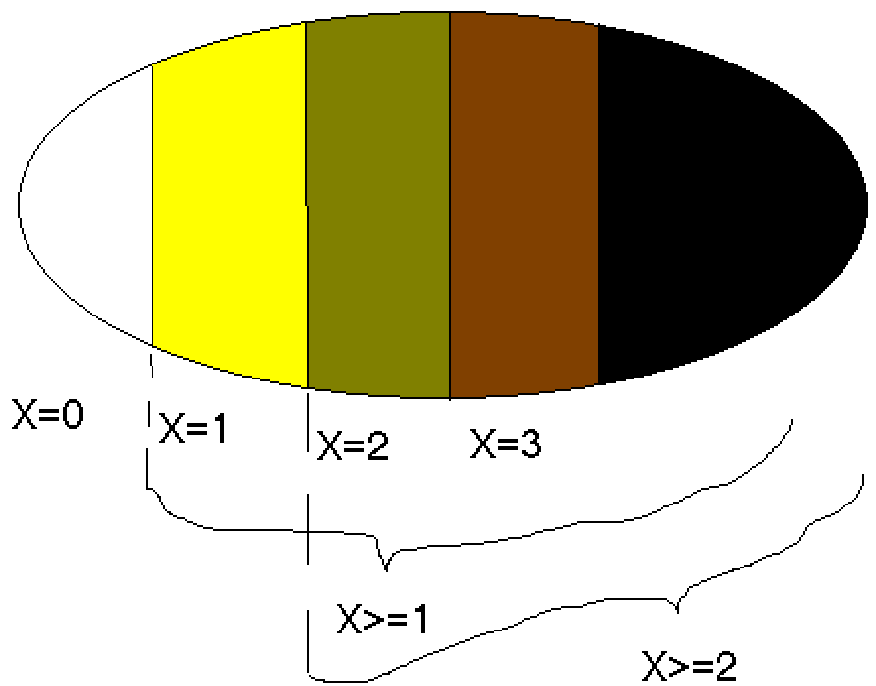
- Nem minden valószínűségi változónak van véges várható értéke:

  $$P(X=2^k)=(1/2)^k \quad k=1,2,\ldots$$

  esetén $E(X)=1+1+1+\ldots=\infty$.
- Azaz annak a játéknak az „ára”, ahol 2^k Ft-ot kapunk, ha szabályos érmével k-adikra dobunk először fejet: végtelen. Ez a Szt.Pétervári paradoxon; gyakorlatban persze nem reális így ez a játék, hiszen nincs az a bank, amely korlátlan pénzt tudna fizetni.
- Ha $E(X)$ véges, akkor az abszolút konvergencia miatt egyértelmű is.

Tulajdonságok 2.

- Ha $X \geq 0$ és $E(X)$ véges, akkor $E(X) \geq 0$.
- Ha $E(X)$ véges, akkor $E(aX+b)=aE(X)+b$ (a várható érték lineáris).
- Ha X nemnegatív egész értékű, akkor $E(X) = P(X \geq 1) + P(X \geq 2) + \ldots$



Alkalmazás: a Pascal eloszlás várható értéke

- P($X \geq k$)=(1-p)$^{k-1}$, így
- E(X)=1+(1-p)+(1-p)^2+…=1/p.
- Természetes eredmény: átlagosan a hatodik dobásra kapjuk az első hatost.
- Tulajdonság: a Pascal eloszlás örökifjú

$$P(X > k + l \mid X > l) = P(X > k)$$

(k,l tetszőleges természetes számok).