Bayes–adatfeldolgozás I

Kálvin Sándor

2014 november 10

A mérések szerepe a fizikában

Elméleti fizika

Tesztelhető elméletek megalkotása.
Az mérési adatok feldolgozásából kapott eredmények birtokában az elméletek módosítása.

Kísérleti fizika

Kísérletek elvégzése és a mérési eredményekből ...

- fizikai paraméterek meghatározása
  paraméterbecslés
- következtetés az elméleti modellek helyességéről
  modellválasztás

A két terület nem különül el egymástól.

Következtetés

A rendelkezésre álló igaz kijelentésekből más eseményekről fogalmazunk meg állításokat.

- A rendelkezésre álló információ elég ahhoz, hogy biztos kijelentéseket tegyünk akkor deduktív logikát alkalmazunk.

- A rendelkezésre álló információ nem elég ahhoz, hogy biztos kijelentéseket tegyünk akkor induktív logikát alkalmazunk. A következtetések bizonytalanok.

Tudásunk mindig hiányos, a mérések pontatlanok.

induktív következtetés kell használnunk

Hiányos ismeretekből $\rightarrow$ nem egyértelműek a következtetések.

statisztikus vagy valószínűségi következtetés

Valószínűségi következtetés

A valószínűségi következtetést használjuk a mindennapi életben (is)
Mekkora az esélye, hogy ...

- holnap sütni fog a nap
  (Kék az ég. Tegnap sütött.) – rendelkezésre álló információk
  /Lemenjek a strandra?/ – döntéshozatal
- a vonat 5 percnél kevesebbet fog késni
  (Ma pontosan érkezett. Tegnap késett.)
  /Melyik vonatra vegyek jegyet?/
- a plazma hőmérséklete 10keV és 12keV között van
  (TS mérés 13keV +/- 4keV)
- { Elmélet-1: – 50MW ... → H-mode
  Elmélet-2: – 40MW ... → L-mode }
  (Mérés– 45MW ... H-mode ...)
  Melyik elmélet valószínűbb?

Deduktív következtetés

- Egy elmélet kísérletileg ellenőrizhető jóslatait alkotjuk meg.
- Elég információ áll rendelkezésre, hogy biztos kijelentést tegyünk.
- Az elméletet igaznak tételezzük fel,
  pontosan ismerjük a mérési eljárást, ...
  ismerjük a statisztikus és szisztematikus hibákat.
- A plazma különböző fizikai paraméterei esetén
  $\rightarrow$ milyen mérési adatokat várunk.
- Az okokból az okozatokat határozzuk meg.

Induktív következtetés

- A kísérletek eredményeiből az elméletek igaz voltára, vagy fizikai paramétereinek értékere következtetünk.
- Nem áll rendelkezésünkre elég információ, hogy biztos kijelentéseket tegyünk.
  Nem tudjuk pontosan meghatározni a paraméterek értékét.
  Nem tudjuk egyértelműen eldönteni melyik elmélet igaz.
- Az okozatokból következtetünk a lehetséges okokra.
- Csak valószínűségi következtetés tehetünk.

Deduktív, induktív következtetés

Deduktív

- A rózsa piros.
- Ez a virág rózsa.
- ◇ Ez a virág piros.

Elég információ áll rendelkezésre.
Biztos kijelentést tehetünk.

Induktív

- A rózsa piros.
- Ez a virág piros.
- ◇ Ez a virág lehet hogy rózsa.

Nem áll elég információ rendelkezésre.
Valószínűségi következtetés tehetünk.
Ez a virág valószínűbb hogy rózsa.

Meg tudjuk határozni, hogy mennyivel valószínűbb, hogy rózsa?

Jelölések

A, B, H ... állítás, elmélet, esemény, ...
$\overline{A}$ – logikai negáció
$A + B$ – logikai VAGY
AB vagy (A, B) – logikai ÉS
(B lehet igaz vagy hamis)

$A|B$ feltételes állítás – A ha B fennáll
(B mindig igaz)
Összetett állítások
$C = \overline{A + B} \quad D = (A + B)(A - C)$ ...
$C = A \Rightarrow B$
A logikai ÉS és a negáció alkalmazásával bármely összetett állítás felírható.

Implikáció

$A \Rightarrow B$ – ha A akkor B

A – ok B – okozat

ekvivalens azzal az egyenlettel, hogy $\quad A = A, B$

ekvivalens azzal az összetett állítással, hogy $\quad \overline{A} + B$

$$A \Rightarrow B \equiv \overline{B} \Rightarrow \overline{A}$$

A	B	A $\Rightarrow$ B	A,B
I	I	I	I
I	H	H	H
H	I	I	H
H	H	I	H

Deduktív következtetés

$A \Rightarrow B$

Megfigyeljük A-t és ez igaz $\rightarrow$ B igaz

$A \Rightarrow B \equiv \overline{B} \Rightarrow \overline{A}$

Megfigyeljük B-t és ez hamis $\rightarrow$ A hamis

Implikáció

A	B	A ⇒ B
I	I	I
I	H	H
H	I	I
H	H	I

Induktív következtetés

- Legyen $A \Rightarrow B$ és megfigyeljük B-t.
  Mit tudunk következtetni A-ról?
- Ha B hamis akkor A is hamis.
  $\overline{B} \Rightarrow \overline{A}$
- Ha B igaz akkor A nem biztos, hogy igaz,
  (nem igaz, hogy $B \Rightarrow A$)
  de A valószínűbb hogy igaz.
- Ha A hamis akkor B nem biztos, hogy hamis,
  de B kevésbé valószínűbb hogy igaz.

Valószínűség a Bayes elméletben

A Bayes elméletben a valószínűség azt fejezi ki, hogy mennyire hiszünk egy esemény igaz voltában az eseménnyel kapcsolatos információk birtokában.
Jelöljük $p(A|I)$-vel az A esemény hihetőségét az I információk birtokában. Valószínűség minden eseményhez, állításhoz hozzárendelhető.
Mindig feltételes.

A Bayes valószínűség elmélet úgy tekinthető mint a logika (Boole algebra) kiterjesztése olyan esetekre amikben nem csak igaz vagy hamis értékek szerepelnek.

Cox axiómák

(Nem axiómák inkább elvárások.)

- Legyen tranzitív $\rightarrow$ hihetőség valós szám $p(A|I)$
- $p(A|I) = 1$, ha az A esemény igaz $p(A|I) = 0$, ha az A esemény hamis
- Ha $p(A|I)$ adott $\rightarrow$ $p(\overline{A}|I)$ is adott
- Ha $p(A|B, I)$ adott és $p(B|I)$ is adott
  $\rightarrow$ $p(AB|I)$ is adott
- A levonható következtetés $p(A|I)$-ról csak I-től függhet nem a logikai lepésektől ahogy meghatározzuk $p(A|I)$-t.

I a rendelkezésre álló releváns információk.
$\rightarrow$ az esemény hihetőségéhez rendelt szám a valószínűségelmélet szabályait követi.
$p(A|I)$ a Bayes értelemben vett valószínűség.

Mi a valószínűség?

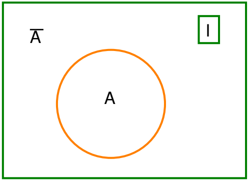
- Klasszikus
  - $p(A|I)$ az A esemény relatív gyakorisága a végtelen sok I eseményben
  - A valószínűségi változó – $p(A|I)$ az A tulajdonsága
- Bayes
  - $p(A|I)$ az A esemény hihetősége, hogy A igaz, ha az I információk állnak rendelkezésre
  - A tetszőleges állítás – $p(A|I)$ ismeretünk A-ról

Valószínűségi kalkulus

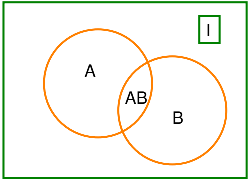
A valószínűségi számítások a következű két szabályon alapulnak. Ezek vezethetőek le a Cox axiómákból.

- összeg szabály
  $p(A|I) + p(\overline{A}|I) = 1$
- szorzat szabály
  $p(A, B|I) = p(A|B, I)p(B|I) = p(B|A, I)p(A|I)$

A



A



Valószínűségi kalkulus – Venn diagram – normált terület
$p(A) = A_t/I_t \quad p(\overline{A}) = \overline{A}_t/I_t \quad p(AB) = AB_t/I_t \quad p(A|B) = AB_t/B_t$

Valószínűségi kalkulus

Az összeg és a szorzatszabályból levezethetőek a következő állítások.

Kiterjesztett összeg szabály
$p(A+B|I) = p(A|I) + p(B|I) - P(A,B|I)$

Ha A és B egymást kizáró, azaz $AB = 0 \rightarrow$
$p(A+B|I) = p(A|I) + p(B|I)$

Ha A és B független, azaz $p(A|B,I) = p(A|I) \rightarrow$
$p(A,B|I) = p(A|I)p(B|I)$

Bayes-tétel

$$p(A|B,I) = \frac{p(B|A,I)p(A|I)}{p(B|I)}$$

Valószínűségi kalkulus

A H állítást bontsuk fel teljes, egymást kölcsönösen kizáró $\{H_i\}$ eseményrendszerre.
A $\{H_i\}$ események közűl egy és csak egy igaz. $\rightarrow$ H igaz

$$H = \sum_i H_i \qquad H_i, H_j = \delta_{ij}$$

$$\sum_i p(H_i|I) = p(\sum_i H_i|I) = p(H|I) = 1$$

Marginalizáció
$\sum_i p(A, H_i|I) = p(A \sum_i (H_i)) = p(A, H|I) = p(A|I)$
Kezelhetőek az u.n. érdektelen paraméterek.
Felbontás
$P(A|I) = \sum_i p(A|H_i, I) P(H_i|I)$
Összetett esemény valószínűsége egyszerűbb események valószínűségéből határozható meg.

Bayes elmélet: a megismerés leírása

A H esemény legyen egy fizikai elmélet, paraméter, ...
A D esemény legyen a mérési adat

$$p(H|D,I) = \frac{p(D|H,I)p(H|I)}{p(D|I)}$$

$p(H|I)$ a prior – tudásunk az adatok kiértékelése előtt.
$p(D|H,I)$ a likelihood – mérési adatok valószínűsége, ha az elmélet igaz.
$p(H|D,I)$ a poszterior –
tudásunk az elméletről a mért adat figyelembevételével.
$p(D|I)$ teljes valószínűség – normálási állandó.

$p(H|I)$ – szubjektív valószínűség
Racionálisan gondolkodó egyének azonos információk birtokában azonos valószínűségeket rendelnek az eseményekhez.

Valószínűségeloszlás

Alkalmazzuk a Bayes tételt egy teljes egymást kölcsönösen kizáró teljes $\{H_i\}$ eseményrendszerre.

$$p(H_i|D,I) = \frac{p(D|H_i,I)p(H_i|I)}{p(D|I)}$$

$\sum p(H_i|D,I) = 1$

$P(D|I) = \sum_i p(D|H_i,I)p(H_i|I)$
A nevező normálási állandó.

A $\{H_i\}$ eseményrendszer poszterior valószínűségeloszlását kapjuk.

Ebben az értelemben a likelihood a hipotézisek függvénye adott adatok esetén.

Folytonos eseménytér

Legyen az A esemény, hogy az a folytonos paraméter az $[b, c]$ intervallumba esik.
Az A esemény valószínűsége
$p(A|I) = \int_b^c pdf(a|I)da$, ahol

$$pdf(a|I) = \lim_{\delta a \to 0} \frac{p(a \leq x < a + \delta a|I)}{\delta a}$$

$pdf()$ a valószínűségi sűrűségfüggvény.
Definiálható teljes EKK eseményrendszert
$A_1 = a\epsilon[a_1, a_2]$
$A_2 = a\epsilon[a_2, a_3]$
.
$A_N = a\epsilon[a_{N-1}, a_N]$
Tartson bármely intervallum hossza nullához.
A valószínűségi kalkulus használatával a fenti kifejezések beláthatóak.

Folytonos eseménytér

Szorzat szabály
$pdf(a,b|I) = pdf(a|b,I)pdf(b|I)$

Normált
$\int pdf(a|I)da = 1$

Marginalizáció
$pdf(a|I) = \int pdf(a,b|I)db$

Bayes tétel

$$pdf(a|D,I) = \frac{pdf(D|a,I)pdf(a|I)}{\int pdf(D|a,I)pdf(a|I)da}$$

Folytonos eseménytér

Diszkrét eseménytér – Folytonos eseménytér
Összegzés – Integrálás
A továbbiakban nem teszünk különbséget jelölésben p és pdf között, változója egyértelműen meghatározza, hogy melyiket kell használni.
A valószínűségi kalkulus valószínűségekre alkalmazható nem a valószínűségi sűrűségfüggvényre.
Valószínűséget úgy kapunk, ha a $p(x|I)$sűrűségfüggvényt szorozzuk dx-szel.
A valószínűség dimenziótlan, így a sűrűségfüggvény dimenziója a változó dimenziójának a reciproka.
Integrálni csak a csak a feltételt jelző vonaltól balra lévő változóra lehet. (Lásd a sűrűségfüggvény definíciója.)
A likelihood nem sűrűségfüggvény, így nem is normált.

Példa

A Bayes módszer alkalmazása.

AIDS teszt

Kovács AIDS tesztet csináltat és a teszt eredménye $+$.
Legyen a teszt hatékonysága 98%

Mi a valószínűsége, hogy Kovács AIDS-es?
Nyilvánvaló, hogy nem 98%.

A probléma fenti megfogalmazása nem egyértelmű.

Példa I

A probléma pontos megfogalmazásához meg kell adni:
Egy teljes egymást kölcsönösen kizáró eseményrendszert.

K Kovács AIDS-es

$\overline{K}$ Kovács nem AIDS-es

A likelihoodot az eseménytéren.

$p(+|K, I)$ a teszt $+$, ha AIDS-es - legyen 98%

$p(+|\overline{K}, I)$ a teszt $+$, ha nem AIDS-es - legyen 3%

A prior valószínűséget,
ismeretünket arról, hogy Kovács AIDS-es mielőtt a tesztet kiértékelnénk.
Legyen ez az AIDS-es betegek aránya a teljes lakosság körében.

$$p(K|I) = 1\%$$

Példa II

Mi a valószínűsége, hogy Kovács AIDS-es, ha a teszt eredménye $+$? Meg kell határozni a $p(K|+,I)$ poszterior valószínűséget.

Bayes tétel

$$p(K|+,I) = \frac{p(+|K,I)p(K|I)}{p(+|K,I)p(K|I) + p(+|\overline{K},I)p(\overline{K}|I)}$$

$$p(K|+,I) = \frac{0.98 \cdot 0.01}{0.98 \cdot 0.01 + 0.03 \cdot 0.99} = 0.248$$

Példa III

Vegyünk extrém esetet.
A test mindig negatív, ha az egyén nem AIDS-es.
$p(+|\overline{K}, I) = 0$

$$p(K|+, I) = \frac{0.98 \cdot 0.01}{0.98 \cdot 0.01 + 0.0 \cdot 0.99} = 1$$

Azaz Kovács biztos,hogy AIDS-es.

Az eredményt megkaphatjuk a Boole algebra felhasználásával is.
I információ $\rightarrow \overline{K} \Rightarrow -$
$+ \Rightarrow K$
$p(K|+, I) = 1$

Példa IV

Vegyünk extrém esetet
A test mindig $+$, ha az egyén AIDS-es.
$p(+|K, I) = 1$

$$p(K|+, I) = \frac{1.00 \cdot 0.01}{1.00 \cdot 0.01 + 0.03 \cdot 0.99} = 0.252$$

Azaz Kovács nem biztos AIDS-es.

Boole algebra
I információ $\rightarrow K \Rightarrow +$
de nem igaz, hogy $+ \Rightarrow K$

Pelda V

Vegyünk extrém esetet
Senki sem AIDS-es
$p(K|I) = 0$

$$p(K|+, I) = \frac{0.98 \cdot 0.0}{0.98 \cdot 0.0 + 0.03 \cdot 1.0} = 0$$

A prior bizonyosságát semmilyen mérés nem tudja megváltoztatni.

Példa VI

Kovács másodszori tesztje is $+$.
A Bayes elméletben most a prior az első teszt poszteriorja.
$P(K|I) = 0.25$

$$p(K|++,I) = \frac{0.98 \cdot .25}{0.98 \cdot 0.25 + 0.03 \cdot 0.75} = 0.92$$

Tekintsük a két tesztet mint egy mérést.
A likelihood – a két $+$ teszt valószínűsége
$p(++|K,I) = 0.98^2$
$p(++|\overline{K},I) = 0.03^2$

$$p(K|++,I) = \frac{0.98^2 \cdot 0.01}{0.98^2 \cdot 0.01 + 0.03^2 \cdot 0.99} = 0.92$$

A két eredmény megegyezik.
Ez a Cox axiómák egyik eleme.