SF1901: Sannolikhetslära och statistik
Föreläsning 2. Betingad sannolikhet & Oberoende

Jan Grandell & Timo Koski

21.01.2016

KTH Matematik

Lärandemål

- Betingad sannolikhet (definition, betydelse)
- Lagen om total sannolikhet
- Bayes' sats (definition, använding, betydelse)
  - a priori sannolikhet
  - posteriori sannolikhet
- Oberoende händelser

KTH Matematik

Repetition: Kolmogorovs axiomsystem

Ett *sannolikhetsmått* P är en funktion av händelser, sådan att:

(a) $0 \leq P(A) \leq 1$;

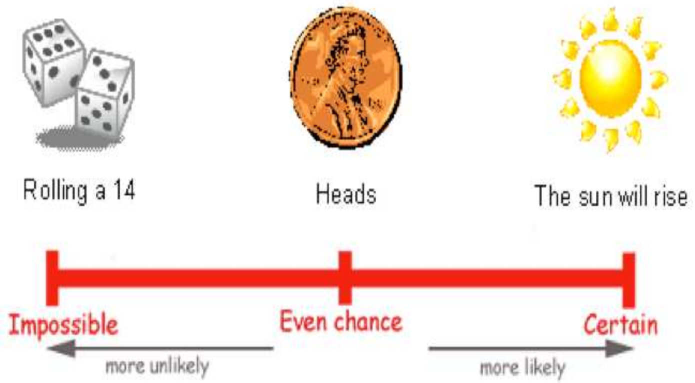
(b) $P(\Omega) = 1$

(c) om A och B är disjunkta händelser, så gäller

$$P(A \cup B) = P(A) + P(B)$$

KTH Matematik

Sannolikhet. En händelse med sannolikhet $= 1$ kallas en **säker** händelse.



KTH Matematik

Repetition: Regler för sannolikhetskalkyl

Sats

$P(A^*) = 1 - P(A).$

Sats

$P(A \cup B) = P(A) + P(B) - P(A \cap B).$

De Morgans regler

Sats

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Sats

$$(A \cap B)^* = A^* \cup B^*$$

KTH Matematik

En sats till

Sats

Låt A och B vara två händelser och $A \subset B$. Då gäller

$$P(A) \leq P(B).$$

Bevis: $B = A \cup (B \cap A^*)$ (rita ett Venndiagram). Det är klart att

$$A \cap (B \cap A^*) = \emptyset$$

Därför ger axiom (c)

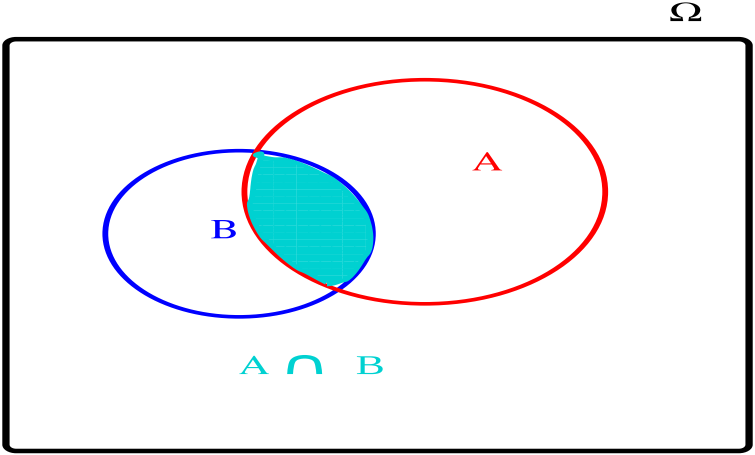
$$P(B) = P(A \cup (B \cap A^*)) = P(A) + P(B \cap A^*) \geq P(A),$$

ty $P(B \cap A^*) \geq 0$.

KTH Matematik

Inledning (betingad sannolikhet)

Låt A och B vara två händelser, dvs $A, B \subset \Omega$. Vad är sannolikheten, betecknad med $P(B \mid A)$, för B då vi vet att A har inträffat?



Vi är ledda oss till följande definition.

Definition

Låt A och B vara två händelser. Antag att $P(A) > 0$. Sannolikheten för B betingat av A betecknas med $P(B \mid A)$ och definieras som

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}.$$

Betingad sannolikhet är ett sannolikhetsmått:

- $P(B \mid A) \geq 0$ är klart. $A \cap B \subseteq A$, och satsen ovan ger $P(A \cap B) \leq P(A)$ d.v.s. $\frac{P(A \cap B)}{P(A)} \leq 1$.
- $P(A \mid A) = \frac{P(A \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$ m.a.o. A är utfallsrummet.
- Tag $B_1 \cap B_2 = \emptyset$.

$$P(B_1 \cup B_2 \mid A) = \frac{P((B_1 \cup B_2) \cap A)}{P(A)}$$

$$= \frac{P((B_1 \cap A) \cup (B_2 \cap A))}{P(A)} = \frac{P((B_1 \cap A)) + P((B_2 \cap A)}{P(A)}$$

and by axiom (c)

$$= \frac{P(B_1 \cap A)}{P(A)} + \frac{P(B_2 \cap A)}{P(A)} = P(B_1 \mid A) + P(B_2 \mid A).$$

KTH Matematik

Exempel

(Kast med röd och vit tärning)
$A =$ *summan av ögonen är högst 4.*
$B_k =$ *vita tärningen visar* k *ögon.*
$P(B_k \mid A) = 0$ *om* $k \geq 4$.
Möjliga utfall, m, är 36: $(v, r), v, r = 1, \ldots 6$, *dvs* $(1, 1), (1, 2), \ldots (6, 6)$.
Gynnsamma utfall för A, är 6*: (1,1), (1,2), (1,3), (2,1), (2,2), (3,1).*
Gynnsamma utfall för $A \cap B_k$, *är* $4 - k$*:* $(v, r), v = k, r = 1, \ldots 4 - k$,
dvs $(k, 1), (k, 2), \ldots (k, 4 - k)$ *om* $k < 4$.

KTH Matematik

Betingad sannolikhet: exempel forts.

Exempel

Klassiska sannolikhetsdefinitionen ger

$$P(A) = \frac{6}{36} \quad och \quad P(A \cap B_k) = \frac{4-k}{36}.$$

Detta ger, för $k < 4$,

$$P(B_k \mid A) = \frac{4-k}{6} = \begin{cases} \frac{3}{6} = \frac{1}{2} & k = 1 \\ \frac{2}{6} = \frac{1}{3} & k = 2 \\ \frac{1}{6} & k = 3. \end{cases}$$

KTH Matematik

Ofta är det lättare att ange värden till betingade sannolikheter än till obetingade, och vi utnyttjar definitionen "baklänges".

Exempel

En ohederlig person har två tärningar, en äkta och en falsk som alltid ger 6 ögon. Han väljer slumpmässigt den ena. Vad är sannolikheten för 5 resp. 6 ögon. Låt oss betrakta fallet med sex ögon. Intuitivt bör gälla att sannolikheten är

$$\frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot 1 = \frac{1}{12} + \frac{6}{12} = \frac{7}{12}.$$

Lagen om total sannolikhet

Mera systematiskt gäller följande sats

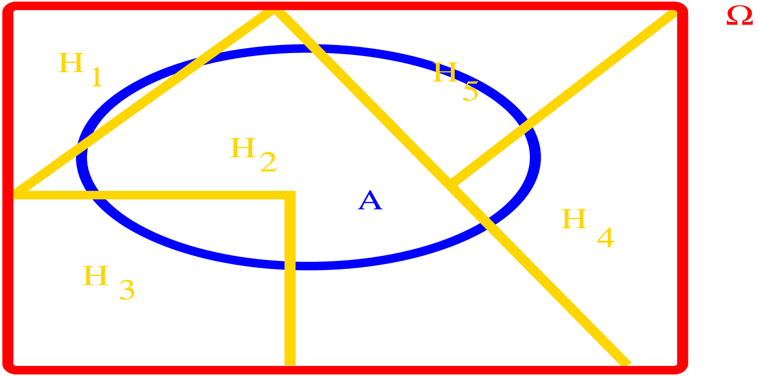
Sats

(Lagen om total sannolikhet)
Om $H_1, \ldots, H_n$ är disjunkta händelser, har positiv sannolikhet och uppfyller hela Ω, så gäller för varje händelse $A \subset \Omega$ att

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A \mid H_i).$$

KTH Matematik

Lagen om total sannolikhet



Lagen om total sannolikhet

Bevis. Vi utnyttjar en av De Morgans regler

$$P(A) = P(A \cap \Omega) = P(A \cap (H_1 \cup \ldots \cup H_n)) = P((A \cap H_1) \cup \ldots \cup (A \cap H_n))$$

och sedan Kolmogorovs axiom (c) från första föreläsningen, ty $(A \cap H_i) \cap (A \cap H_j) = \emptyset$ om $i \neq j$,

$$= \sum_{i=1}^{n} P(A \cap H_i) = \sum_{i=1}^{n} P(H_i) P(A \mid H_i).$$

$\square$

KTH Matematik

Bayes' sats

Vi ska nu ge en viktig sats om "vändning" av händelserna i betingade sannolikheter.

Sats

(Bayes' sats) *Under samma villkor som i lagen om total sannolikhet gäller*

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{\sum_{j=1}^{n} P(H_j)P(A \mid H_j)}.$$

Bevis.

$$P(H_i \mid A) = \frac{P(H_i \cap A)}{P(A)} = \frac{P(H_i \cap A)}{P(H_i)} \cdot \frac{P(H_i)}{P(A)} = P(A \mid H_i) \cdot \frac{P(H_i)}{P(A)}.$$

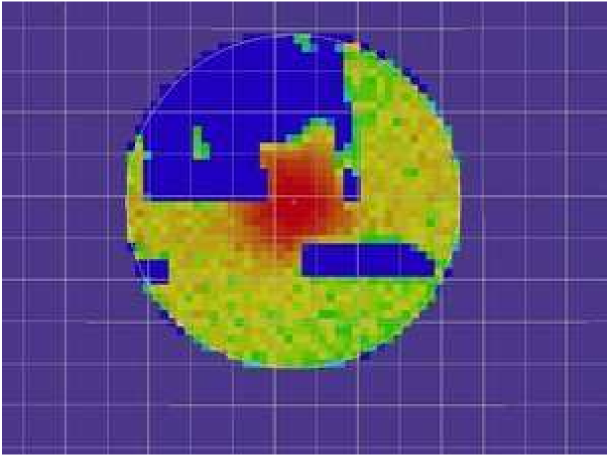
Lagen om total sannolikhet tillämpad på $P(A)$ ger resultatet.

KTH VETENSKAP OCH KONST
KTH Matematik

Bayes' sats: tolkning

Sannolikheten $P(H_i)$ kallas **a priori sannolikhet** för H_i. Sannolikheten $P(H_i \mid A)$ kallas **posteriori sannolikhet** för H_i. Bayes' sats visar alltså hur a priori sannolikhet omvandlas till posteriori sannolikhet.

Bayesian search theory is the application of Bayesian statistics to the search for lost objects. It has been used several times to find lost sea vessels, for example the USS Scorpion. It also played a key role in the recovery of the flight recorders in the Air France Flight 447 disaster of 2009.



Logiska resonemang & sannolikhet

Anta

$$P(B|A) > P(B).$$

D.v.s., om A inträffat (är sann), blir B mer sannolik. Men då fås

$$\frac{P(A)\,P(B|A)}{P(B)} > P(A)$$

d.v.s.

$$P(A \mid B) > P(A).$$

eller: om B inträffar, blir A mer sannolik. Till exempel, om det regnar, är en ytbeläggning blöt. Nu ser vi att ytbeläggningen är blöt, det är alltså mer sannolikt att det regnat.

Antag $P(B|A) > P(B)$. Då gäller:

1. Om B^* är sant, då blir A mindre sannolik.
2. Om A^* är sant, då blir B mindre sannolik.

En mjukversion av den vanliga logiken.

Drill på betingad sannolikhet

$0 < \mathbf{P}(A) < 1$. Vilket av följande påståenden är *felaktigt* ?

a) $\mathbf{P}(\Omega|A) = 1$.

b) $\mathbf{P}(A|\Omega) = \mathbf{P}(A)$.

c) $\mathbf{P}(A^*|A) = 0$.

d) $\mathbf{P}(A|A) = \mathbf{P}(A)$.

KTH Matematik

Thomas Bayes ca. 1701 – 1761

Bayes' sats: exempel

Låt oss gå tillbaka till exemplet om falskspelaren. Sätt
$A = 6$ ögon.
$H_1 =$ äkta tärningen.
$H_2 =$ falska tärningen.
Då gäller

$$P(A) = P(H_1)P(A \mid H_1) + P(H_2)P(A \mid H_2) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot 1 = \frac{7}{12},$$

som i exemplet. Bayes' sats ger vidare

$$P(H_1 \mid A) = \frac{P(H_1 \cap A)}{P(A)} = P(A \mid H_1) \cdot \frac{P(H_1)}{P(A)} = \frac{1}{6}\frac{1}{2}\frac{12}{7} = \frac{1}{7}$$

och

$$P(H_2 \mid A) = \frac{P(H_2 \cap A)}{P(A)} = P(A \mid H_2) \cdot \frac{P(H_2)}{P(A)} = 1 \cdot \frac{1}{2}\frac{12}{7} = \frac{6}{7}$$

vilket kanske inte är lika lätt att inse rent intuitivt.

Bayes' sats: screening för cancer

Vi räknar på det exempel som ges av John Allen Paulos: The Math behind Screening Tests. What a positive result really means. *Scientific American* , January 2012. http://www.scientificamerican.com/article.cfm?id=weighing-the-positives

Låt oss anta att 0.4 % av befolkningen drabbas av en form av cancer. Ett diagnostiskt test ger positivt resultat i 99.5 % av fallen, om en individ har cancer och ger positivt resultat i 1 % av fallen, om en individ inte har cancer. Om vi plockar på måfå en vuxen individ ur en stor population och testet ger ett postivt svar, vad är sannolikheten för att individen har den aktuella formen av cancer ? Beteckna $A_1 =$ individen har cancer. $A_2 =$ individen har inte cancer. $B =$ positivt testresultat. Vi har

$P(A_1) = 0.004, P(A_2) = 0.996, P(B|A_1) = 0.995, P(B|A_2) = 0.01$

Sökt är $P(A_1|B)$.

KTH Matematik

Bayes' sats: medicinsk diagnostik

$$P(A_1) = 0.004, P(A_2) = 0.996, P(B|A_1) = 0.995, P(B|A_2) = 0.01$$

Bayes' sats ger $P(A_1|B)$ som

$$P(A_1|B) = \frac{P(B|A_1)\,P(A_1)}{P(B|A_1)\,P(A_1) + P(B|A_2)\,P(A_2)} =$$

$$= \frac{0.995 \cdot 0.004}{0.995 \cdot 0.004 + 0.01 \cdot 0.996} = 0.2855.$$

En oväntat låg posteriori sannolikhet ! Varför är detta ändå ett rimligt svar ? Vilken slutsats drar vi om screening ?

KTH Matematik

Bayes' sats & inlärning

Vi har N slantar i en urna. $N-1$ av dessa är hederliga i den meningen att de har två olika sidor, krona och klave. En av slantarna är falsk i den meningen att den har krona på båda sidor.
Någon plockar på måfå en av slantarna utan att vi får inspektera densamma, och singlar den k ggr. Utfallet blir k kronor. Vad är sannolikheten att den valda slanten är den falska?

Beteckningar: H_1 = hederlig slant, H_2 = falsk slant. A_k = händelsen med k st. kronor i k singlingar.
Sannolikheter enligt uppgiften: $P(H_1) = \frac{N-1}{N}$ (antalet gynnsamma fall/totalantalet fall). $P(H_2) = \frac{1}{N}$. $P(A_k \mid H_1) = \frac{1}{2^k}$ (antalet gynnsamma fall/totalantalet fall, senare alt. p.g.a oberoende) och $P(A_k \mid H_2) = 1$. Sökt: $P(H_2 \mid A_k)$.

KTH Matematik

Bayes' sats & inlärning

H_1 = hederlig slant, H_2 = falsk slant. A_k = händelsen med k st. kronor i k singlingar.
$P(H_1) = \frac{N-1}{N}$, $P(H_2) = \frac{1}{N}$. $P(A_k \mid H_1) = \frac{1}{2^k}$ och $P(A_k \mid H_2) = 1$. Sökt: $P(H_2 \mid A_k)$.

Lagen om total sannolikhet (LTS) ger

$$P(A_k) = P(A_k \mid H_1) P(H_1) + P(A_k \mid H_2) P(H_2)$$

$$= \frac{1}{2^k}\frac{N-1}{N} + \frac{1}{N} = \frac{2^k + N - 1}{2^k N}$$

Bayes' sats ger

$$P(H_2 \mid A_k) = \frac{P(A_k \mid H_2) P(H_2)}{P(A_k)}$$

$$= \frac{2^k}{2^k + N - 1}.$$

KTH Matematik

Bayes' sats & inlärning

H_1 = hederlig slant, H_2 = falsk slant. A_k = händelsen med k st. kronor i k singlingar.
$P(H_1) = \frac{N-1}{N}$, $P(H_2) = \frac{1}{N}$, $P(A_k \mid H_1) = \frac{1}{2^k}$ och $P(A_k \mid H_2) = 1$.

$$P(H_2 \mid A_k) = \frac{2^k}{2^k + N - 1}.$$

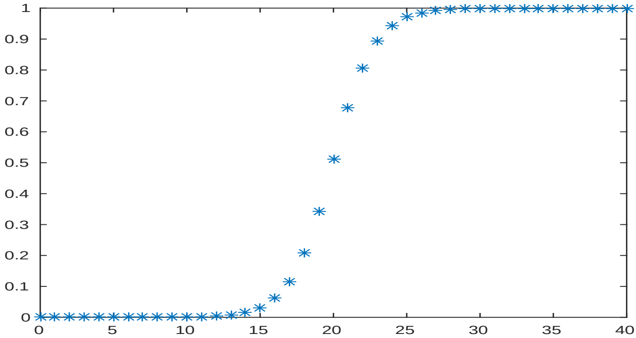
Observera att

$$P(H_2 \mid A_0) = \frac{1}{N} = P(H_2).$$

KTH Matematik

I figuren plottas med $*$ den sökta sannolikheten som funktion av k för $k = 0, \ldots, 40$ och $N = 1000000$

$$P\left(H_2 \mid A_k\right) = \frac{2^k}{2^k + N - 1}.$$



Oberoende händelser

Intuitivt är två händelser A och B oberoende om inträffandet av A inte ger någon information om huruvida B inträffar eller ej. I formler[1] betyder detta

$$P(B \mid A) = P(B).$$

Allmänt gäller ju

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}, \quad \text{om } P(A) > 0.$$

[1] J.f.r med $P(H_2 \mid A_0) = P(H_2)$ ovan.

Oberoende händelser

Multiplikation med $P(A)$ leder oss till följande definition:

Definition

Två händelser A och B är oberoende om

$$P(A \cap B) = P(A)P(B).$$

Definitionen ovan kräver inget villkor om positiva sannolikheter.

Oberoende händelser

Det är inte självklart hur oberoende skall definieras för flera händelser.

Definition

Tre händelser A, B och C är oberoende om

$$
\begin{aligned}
P(A \cap B) &= P(A)P(B) \\
P(A \cap C) &= P(A)P(C) \\
P(B \cap C) &= P(B)P(C) \\
P(A \cap B \cap C) &= P(A)P(B)P(C).
\end{aligned}
$$

Endast $P(A \cap B \cap C) = P(A)P(B)P(C)$ räcker inte, vilket inses om vi sätter $A = B$ och $C = \emptyset$.

KTH Matematik

Inte heller räcker parvis oberoende, vilket ses av följande exempel:
Kast med röd och vit tärning:
$A =$ vita tärningen visar jämnt antal ögon.
$B =$ röda tärningen visar jämnt antal ögon.
$C =$ jämn ögonsumma.
A och B är oberoende av "försöksskäl". Vidare gäller

$$P(A \cap C) = P(A \cap B) = P(A)P(B) = \frac{1}{4} \quad \text{och} \quad P(A)P(C) = \frac{1}{4}.$$

Således är A och C oberoende. Pss. följer att B och C är oberoende. Eftersom $A \cap B \Rightarrow C$ vore det inte rimligt att anse att A, B och C är oberoende.

Allmänt: Oavsett vilka händelser vi plockar ut så skall sannolikheten för snittet vara produkten av sannolikheterna.
Man kan visa att om $A_1, \ldots, A_n$ är oberoende, så är även $A_1^*, \ldots, A_n^*$ oberoende. Detta kan verka helt självklart, med är inte helt lätt att visa. Vi nöjer oss med fallet $n = 2$.

Oberoende händelser

Vi har

$$P(A^* \cap B^*) = P((A \cup B)^*) = 1 - P(A \cup B)$$

$$= 1 - P(A) - P(B) + P(A)P(B) = 1 - P(A) - P(B)(1 - P(A))$$

$$= (1 - P(A))(1 - P(B)) = P(A^*)P(B^*).$$

KTH Matematik

Oberoende händelser

Sats

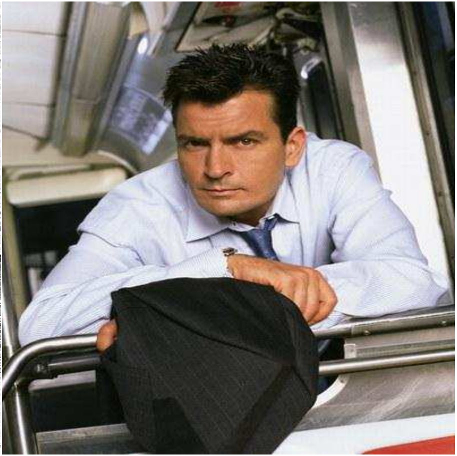
Låt händelserna $A_1, \ldots, A_n$ vara oberoende. Sätt $B = \bigcup_1^n A_i$, dvs. minst en av händelserna $A_1, \ldots, A_n$ inträffar. Då gäller

$$P(B) = 1 - (1 - P(A_1))(1 - P(A_2)) \ldots (1 - P(A_n)).$$

Bevis.

$$P(B) = 1 - P(B^*) = 1 - P\left(\bigcap_1^n A_i^*\right) = 1 - \prod_1^n P(A_i^*) = 1 - \prod_1^n (1 - P(A_i))$$

$\square$



KTH Matematik

Stokastiska variabler: inledning

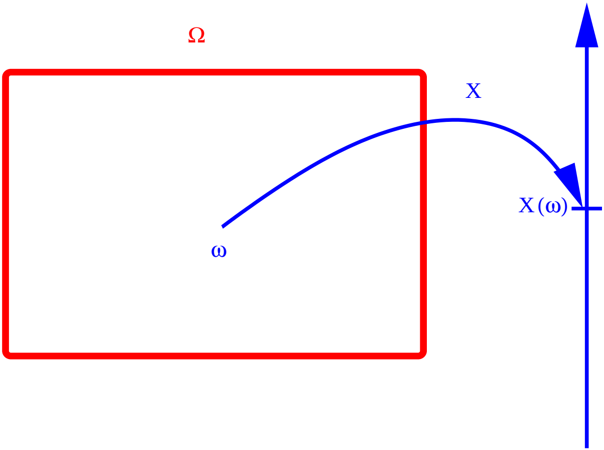
nästan alla situationer som vi betraktar, kommer resultaten av slumpförsöken att vara tal, kontinerliga mätvärden eller antal. Det är praktiskt att anpassa beteckningarna till detta.

Definition

En stokastisk variabel s.v. (eller en slumpvariabel) X är en funktion från Ω till reella linjen.

Lite löst kommer vi att uppfatta X som en beteckning för resultatet av ett slumpförsök.
För ett tärningskast kan X anta ett av värdena 1, 2, 3, 4, 5 eller 6.

KTH Matematik


Ω
X
X(ω)
ω