

# SF1901: Sannolikhetslära och statistik

## Föreläsning 2. Betingad sannolikhet & Oberoende

Jan Grandell & Timo Koski

21.01.2015



KTH Matematik

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

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



Rolling a 14



Heads



The sun will rise



# Repetition: Regler för sannolikhetskalkyl

Sats

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

Sats

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



KTH Matematik

## 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

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

## 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)

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

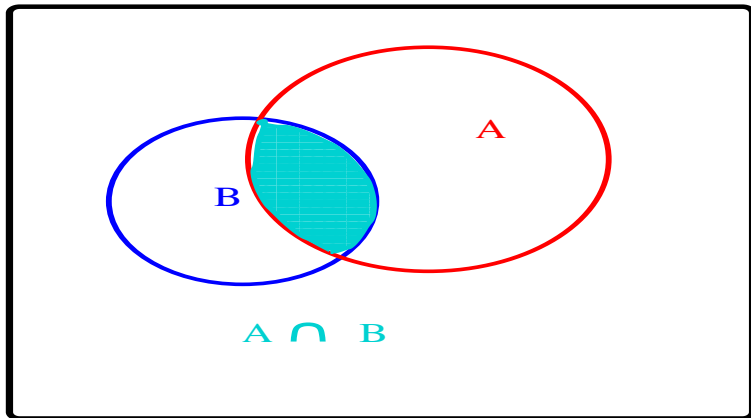
ty  $P(B \cap A^*) \geq 0$ .



# Inledning (betingad sannolikhet)

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

$\Omega$



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  $P(B | A)$  och definieras som

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



# Betingad sannolikhet är ett sannolikhetsmått:

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

$$\begin{aligned} P(B_1 \cup B_2 | A) &= \frac{P(B_1 \cup B_2 \cap A)}{P(A)} \\ &= \frac{P((B_1 \cap A) \cup (B_2 \cap A))}{P(A)} = \frac{P((B_1 \cap A)) + P((B_2 \cap A))}{P(A)} \end{aligned}$$

and by axiom (c)

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



## Exempel

*(Kast med röd och vit tärning)*

$A =$  *summan av ögonen är högst 4.*

$B_k =$  *vita tärningen visar  $k$  ögon.*

$P(B_k | A) = 0$  *om  $k \geq 4$ .*

*Möjliga utfall,  $m$ , är 36:  $(v, r)$ ,  $v, r = 1, \dots, 6$ , dvs  $(1, 1), (1, 2), \dots, (6, 6)$ .*

*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  $4 - k$ :  $(v, r)$ ,  $v = k, r = 1, \dots, 4 - k$ ,  
dvs  $(k, 1), (k, 2), \dots, (k, 4 - k)$  om  $k < 4$ .*

## Exempel

*Klassiska sannolikhetsdefinitionen ger*

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

*Detta ger, för  $k < 4$ ,*

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

Ofta är det lättare att ange värden till betingade sannolikheter än till obetingade, och vi utnyttar 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*

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

Mera systematiskt gäller följande sats

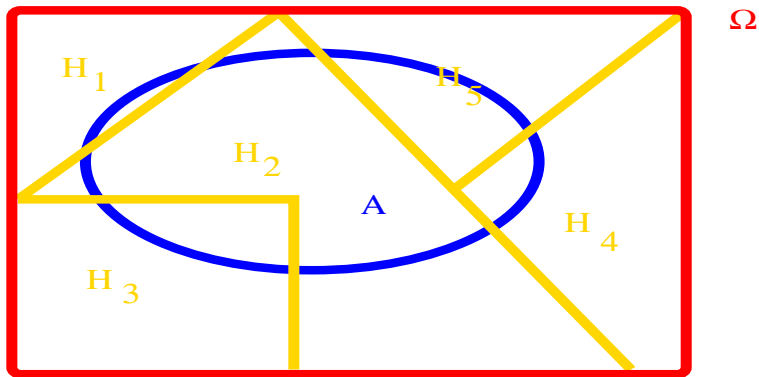
## Sats

### (Lagen om total sannolikhet)

*Om  $H_1, \dots, H_n$  är disjunkta händelser, har positiv sannolikhet och uppfyller hela  $\Omega$ , så gäller för varje händelse  $A \subset \Omega$  att*

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

# Lagen om total sannolikhet



**Bevis.** Vi utnyttjar en av De Morgans regler

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

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

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



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

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

**Bevis.**

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

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



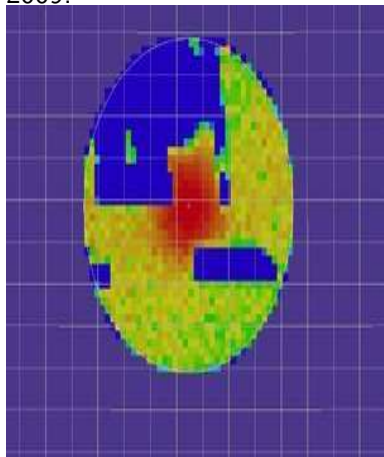


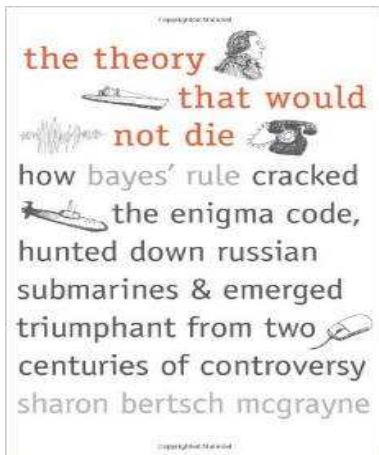
# Bayes' sats: tolkning

Sannolikheten  $P(H_i)$  kallas **a priori sannolikhet** för  $H_i$ . Sannolikheten  $P(H_i | A)$  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.





# 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

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

som i exemplet. Bayes' sats ger vidare

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

och

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

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 positivt 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  $P(A_1|B)$ .



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

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

$$\begin{aligned} P(A_1|B) &= \frac{P(B|A_1) P(A_1)}{P(B|A_1) P(A_1) + P(B|A_2) P(A_2)} = \\ &= \frac{0.995 \cdot 0.004}{0.995 \cdot 0.004 + 0.01 \cdot 0.996} = 0.2855. \end{aligned}$$

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

Intuitivt är två händelser  $A$  och  $B$  oberoende om inträffandet av  $A$  inte ger någon information om huruvida  $B$  inträffar eller ej. I formuler betyder detta

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

Allmänt gäller ju

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



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

## Definition

*Two events  $A$  and  $B$  are independent if*

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

Definitionen ovan kräver inget villkor om positiva sannolikheter.

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

## Definition

*Tre händelser  $A$ ,  $B$  och  $C$  är oberoende om*

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

$$P(A \cap C) = P(A)P(C)$$

$$P(B \cap C) = P(B)P(C)$$

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

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

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  $A_1, \dots, A_n$  är oberoende, så är även  $A_1^*, \dots, A_n^*$  oberoende. Detta kan verka helt självklart, med är inte helt lätt att visa. Vi nöjer oss med fallet  $n = 2$ .

Vi har

$$\begin{aligned}P(A^* \cap B^*) &= P((A \cup B)^*) = 1 - P(A \cup B) \\&= 1 - P(A) - P(B) + P(A)P(B) = 1 - P(A) - P(B)(1 - P(A)) \\&= (1 - P(A))(1 - P(B)) = P(A^*)P(B^*).\end{aligned}$$

## Sats

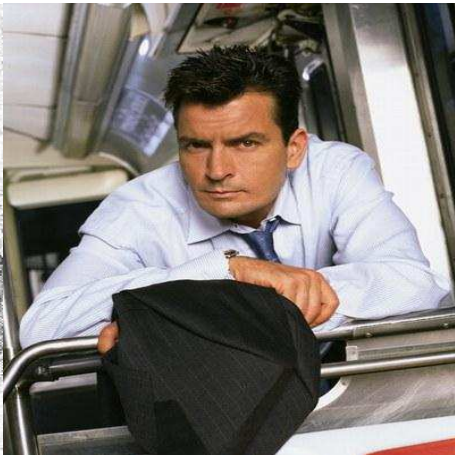
Låt händelserna  $A_1, \dots, A_n$  vara oberoende. Sätt  $B = \bigcup_1^n A_i$ , dvs. minst en av händelserna  $A_1, \dots, A_n$  inträffar. Då gäller

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

**Bevis.**

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





KTH Matematik

# What does it mean when X smiles at you every time hen sees you? A Bayesian analysis

$$p(A | E) = \frac{p(E | A) \cdot p(A)}{p(E)}.$$

$A = \text{like}$ ,  $E = \text{smile}$ .  $p(A | E) = p(\text{like}|\text{smile})$  is what you want to know; the probability X likes you given the fact that X smiles at you.

$p(E | A) = p(\text{smile}|\text{like})$  is the probability that X will smile given that X sees someone 'hen' likes.  $p(A) = p(\text{like})$  is the probability that X likes a random person.  $p(E) = p(\text{smile})$  is the probability that X will smile at a random person.

For example, suppose X just smiles at everyone. Then intuition says that fact that X smiles at you doesnt mean anything one way or another.

Indeed,  $p(\text{smile}|\text{like}) = 1$  and  $p(\text{smile})=1$ , and we have  $p(\text{like}|\text{smile}) = p(\text{like})$





nästan alla situationer som vi betraktar, kommer resultaten av slumpförsöken att vara tal, kontinuerliga 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  $\Omega$  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.

# Stokastiska variabler

