TAMS79: Föreläsning 1
Grundläggande begrepp

Johan Thim*

31 oktober 2018

1.1 Begrepp

Ett **slumpförsök** är ett försök där resultatet ej kan förutsägas deterministiskt. Slumpförsöket har olika möjliga **utfall**. Vi låter **Utfallsrummet** Ω vara mängden av alla möjliga utfall. En **händelse** är en delmängd av Ω, dvs en mängd av utfall. Men, alla möjliga delmängder av Ω behöver inte vara *tillåtna* händelser. För att precisera detta kräver vi att mängden av alla händelser (detta är alltså en mängd av mängder) är en så kallad σ-algebra. Vi definierar detta begrepp lite senare.

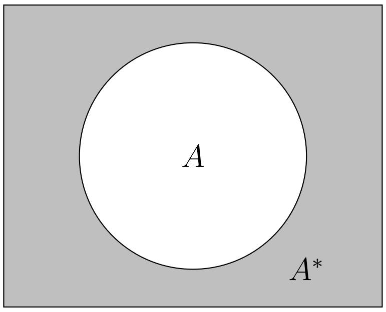
Exempel

1. Myntkast: $\Omega = \{\,\text{Krona},\ \text{Klave}\,\}$.
2. Tärning (T-6): $\Omega = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$. $A = \{1,\ 3,\ 5\}$ och $B = \{4\}$ är exempel på händelser.
3. Tiden till bilen går sönder: $\Omega = [0, \infty[\ = \{x \in \mathbf{R} : x \geq 0\}$. T ex $A = [0, 10[$ är händelsen att bilen går sönder innan 10 tidsenheter gått.

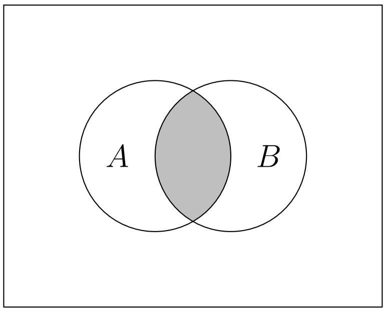
1.2 Mängdlära

En mängd M är en samling element utan ordning. Antalet element (kardinaliteten) i mängden betecknar vi $|M|$. Om mängden är ändlig är detta alltså bara hur många element som finns i mängden. Om mängden innehåller oändligt många element blir begreppet lite krångligare. En delmängd $A \subset M$ innehåller endast element från M (möjligen alla värden i M, eller inga). Mängder illustreras ofta med Venn-diagram. Nedan är hela rektangeln utfallsrummet Ω och de skuggade områderna olika delmängder.

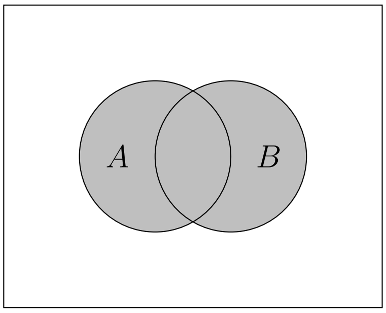
*johan.thim@liu.se



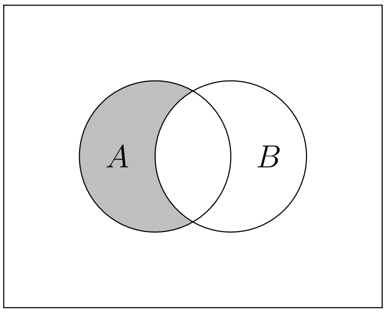
A är en händelse och A^* dess komplement. Komplementet A^* är händelsen att A **inte** inträffar. Komplementet A^* består av *alla* utfall (i Ω) som *inte* finns i händelsen A.



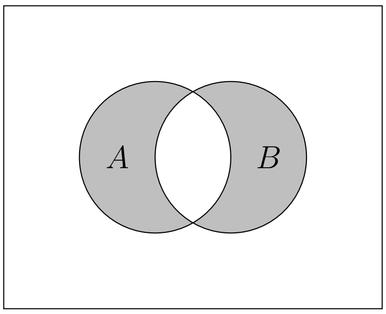
Snittet $A \cap B$ mellan händelserna $A, B \subset \Omega$. Händelsen att *både* A och B inträffar.
Om $A \cap B = \emptyset$ (tomma mängden) så kallas A och B för **oförenliga** (eller **disjunkta**). Två oförenliga händelser kan *ej* inträffa samtidigt.
Observera att $A \cap A^* = \emptyset$



Unionen $A \cup B$ mellan händelserna $A, B \subset \Omega$. Händelsen att *någon* av A och B inträffar (eller båda).
Observera att $A \cup A^* = \Omega$ (hela utfallsrummet).



Skillnaden $A \setminus B = A \cap B^*$. Alla utfall i A utom de som även ligger i B. Händelsen att A inträffar men *inte* B.
Observera att $A^* = \Omega \setminus A$.



Symmetriska skillnaden: $A \Delta B = (A \setminus B) \cup (B \setminus A)$. Händelsen att en av A och B inträffar, men *inte* båda. Exklusivt eller.

1.3 Sannolikhet

Händelser skulle som sagt vara element i en σ-algebra, vilket är ett objekt som definieras enligt följande.

σ-algebra

Definition. $\mathcal{F}$ är en σ-algebra på Ω om $\mathcal{F}$ består av delmängder av Ω så att

(i) $\Omega \in \mathcal{F}$.

(ii) om $A \in \mathcal{F}$ så är $A^* \in \mathcal{F}$.

(iii) om $A_1, A_2, \ldots \in \mathcal{F}$ så är unionen $A_1 \cup A_2 \cup \cdots \in \mathcal{F}$.

Det enklaste exemplet på en σ-algebra är $\mathcal{F} = \{\Omega, \emptyset\}$, dvs endast hela utfallsrummet och den tomma mängden. Av förklarliga skäl kommer vi inte så långt med detta. Ett annat vanligt exempel är att $\mathcal{F}$ består av *alla* möjliga delmängder till Ω; skrivs ibland $\mathcal{F} = 2^\Omega$, och kallas potensmängden av Ω. Denna konstruktion är lämplig när vi har diskreta utfall. Om Ω består av ett kontinuum så visar det sig dock att 2^Ω blir alldeles för stor för många tillämpningar.

Kolmogorovs Axiom: Sannolikhetsmått

Definition. Ett sannolikhetsmått på en σ-algebra $\mathcal{F}$ över ett utfallsrum Ω tilldelar ett tal mellan noll och ett, en sannolikhet, för varje händelse som är definierad (dvs tillhör $\mathcal{F}$). Formellt är P en mängdfunktion; $P\colon \mathcal{F} \to [0,1]$. Sannolikhetsmåttet P *måste* uppfylla Kolmogorovs axiom:

(i) $0 \leq P(A) \leq 1$ för varje $A \in \mathcal{F}$.

(ii) $P(\Omega) = 1$.

(iii) Om $A \cap B = \emptyset$ så gäller att $P(A \cup B) = P(A) + P(B)$.

Formellt är det alltid en trippel $(\Omega, \mathcal{F}, P)$ när vi diskuterar sannolikhet, men vi låter ofta $\mathcal{F}$ vara underförstådd.
Masstolkning: Ibland tolkas $P(A)$ som händelsen A:s sannolikhetsmassa. Ger en intuitiv bild av sannolikhetsfördelning mellan händelser (ofta grafiskt). Rita proportionerliga Venn-diagram!
Följder av dessa axiom innefattar följande.

Egenskaper för sannolikhetsmåttet

(i) $P(A^*) = 1 - P(A)$;

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

(iii) $P(A \cup B) \leq P(A) + P(B)$ (Booles olikhet);

(iv) $P(\emptyset) = 0$;

(v) $P(A \setminus B) = P(A) - P(A \cap B)$.

Bevis? Rita Venn-diagram!

1.3.1 Klassiska definitionen av sannolikhet

Ett försök där varje utfall har samma sannolikhet säges ha **likformig** sannolikhetsfördelning. Om Ω är ändlig, säg $|\Omega| = m$, så är $P(\omega) = 1/m$ för varje utfall $\omega \in \Omega$.

För en likformig sannolikhetsfördelning på Ω så gäller för en händelse $A \subset \Omega$ att

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{"gynsamma utfall"}}{\text{"möjliga utfall"}}.$$

Två tärningar

Vi kastar två rättvisa tärningar. Låt C vara händelsen att poängsumman blir sju. Vad är sannolikheten för C?

Lösning: Det finns 36 st möjliga utfall. Det första kastet ger sex möjligheter, och för vart och ett av dessa utfall får vi sex nya möjligheter vid andra kastet. Av dessa fall är följande "gynsamma":

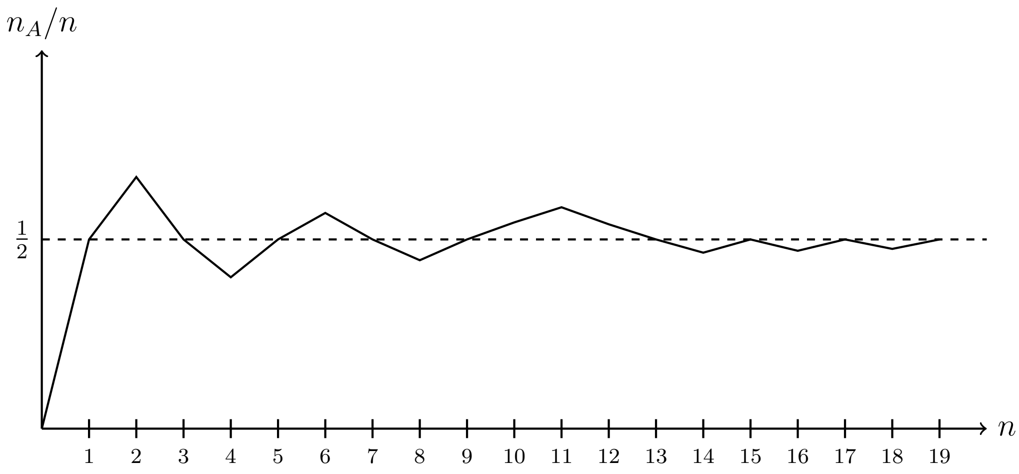
$$C = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}.$$

Vi representerar kasten som två koordinater, den första är från tärning ett och den andra från tärning två. Enligt den klassiska sannolikhetsdefinitionen får vi

$$P(C) = \frac{6}{36} = \frac{1}{6}.$$

1.3.2 Frekvenstolkning

Vi upprepar ett försök n gånger och räknar antalet n_A gånger som händelsen A inträffar. Den relativa frekvensen definieras som n_A/n. Om $n \to \infty$ förefaller det rimligt att $n_A/n \to P(A)$. Detta kallas frekvenstolkningen av sannolikhet. Som exempel, låt oss singla slant många gånger och räkna antalet kronor ($A = \{\text{Krona}\}$) och plotta den relativa frekvensen:



Om myntet är symmetriskt förväntar vi oss att $n_A/n \to 1/2$ (eller hur?).

1.4 Oberoende

Oberoende

Definition. Två händelser A och B kallas **oberoende** om $P(A \cap B) = P(A)P(B)$.

Observera att denna likhet *inte* gäller i allmänhet. Ta till exempel händelsen $A = \{\text{Krona}\}$ och $B = \{\text{Klave}\}$ vid ett myntkast. Klart att $A \cap B = \emptyset$ så $P(A \cap B) = 0$. Men

$$P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \neq 0.$$

Självklart är A och B inte oberoende. Observera även att om vi pratar om tre eller fler händelser blir definitionen av oberoende krångligare; se boken.

1.5 Kombinatorik

Multiplikationsprincipen: Om vi har en tvåstegsprocess av valmöjligheter, där vi i första steget har n_1 möjligheter och i det andra n_2 möjliga val, så finns det totalt sätt $n_1 \cdot n_2$ kombinationer.

Exempel

En tre-rätters meny har 2 förrätter, 3 varmrätter, och 4 efterätter. Hur många olika måltider kan man beställa om man vill ha förrätt, varmrätt och efterätt?

Enligt multiplikationsprincipen blir det $2 \cdot 3 \cdot 4 = 24$ olika måltider.

Något lite krångligare? Vi utnyttjar multiplikationsprincipen för att reda ut följande scenario.

Inbrottstjuven

Inbrottstjuven Ivar försöker öppna 10 st dörrar, $D_1, D_2, \dots, D_{10}$, och att Ivar har 80% sannolikhet att lyckas med varje dörr. Vad är sannolikheten att exakt sex st dörrar blir öppnade?

Lösning: Vi antar att öppnandet av olika dörrar är oberoende av varandra (är det rimligt?). En viss följd av resultat, t ex $D_1 = Y$, $D_2 = N$, $D_3 = Y, \dots D_{10} = N$ (med sex st Y, öppna dörrar, och fyra st N, misslyckade försök), har eftersom händelserna är oberoende sannolikheten

$$\begin{aligned} P(D_1 = Y, \dots, D_{10} = N) &= P(D_1 = Y)P(D_2 = N) \cdots P(D_{10} = N) \\ &= 0.8 \cdot 0.2 \cdot \dots \cdot 0.2 \\ &= 0.8^6 \cdot 0.2^4 \approx 4.194 \cdot 10^{-4}. \end{aligned}$$

Hur många sådana följder finns det? Vi har tio dörrar och skall välja ut sex st som öppnas:

Dörr 1	Dörr 2	Dörr 3	Dörr 4	Dörr 5	Dörr 6
10	9	8	7	6	5

Dörr 1 kan vi välja på 10 olika sätt. När vi sedan väljer dörr 2 finns det bara 9 kvar att välja på. Och så vidare. Ordningen på dörrarna är nu fixerad, och vi får (från multiplikationsprincipen)

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 151200$$

sådana val.
När de sex dörrarna är valda kan vi variera ordningen mellan dessa 6 på 6! olika sätt:

Dörr 1	Dörr 2	Dörr 3	Dörr 4	Dörr 5	Dörr 6
6	5	4	3	2	1

Vi kan nu ta bort ”multipla” dörrval (de kombinationer som bara skiljer sig åt med i vilken ordning sex st specifika dörrar ligger):

$$\frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}{6!} = \frac{10!}{6! \cdot 4!} = \binom{10}{6}.$$

Vi får alltså en binomialkoefficient!
Eftersom de olika sekvenserna av dörrval är oförenliga händelser (två olika val ger olika dörrsekvenser) får vi sannolikheten

$$\binom{10}{6} 0.8^6 0.2^4 \approx 0.088.$$

Det är alltså ungefär 8.8% chans att exakt sex stycken dörrar blir öppnade.

1.6 Betingad sannolikhet

Betingad sannolikhet

Definition. Låt $P(B) > 0$. Den betingade sannolikheten $P(A \mid B)$ för händelsen A, givet att händelsen B inträffar, definieras som $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$.

Om A och B är oberoende och $P(B) \neq 0$ ser vi att

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

Rimligt?

Exempel

Alla som lyssnar på hårdrock i någon form har säkert funderat över vilken av Slayer-låtarna *Angel of Death* och *Raining Blood* som är bäst[a]. Examinator funderade över detta och samlade in följande siffror på internet:

	Angel of Death	*Raining Blood*	Summa
`Returntothepit.com`	199	173	372
`MetalStorm.net`	47	43	90
Summa	246	216	462

[a] Självklart är *Angel of Death* den bästa av dessa två, men det är inte poängen!

Låt A = Angel of Death och B = `Returntothepit.com`. Från tabellen erhåller vi

$$P(A) = \frac{246}{462}, \qquad P(B) = \frac{372}{462} \qquad \text{och} \qquad P(B \cap A) = \frac{199}{462}.$$

Vi kan direkt beräkna $P(B \mid A)$ genom att titta enbart i första kolumnen (det vi menar med sannolikhet betingad på A = första kolumnen): $P(B \mid A) = \dfrac{199}{246}$. Använder vi definitionen istället blir det

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{199/462}{246/462} = \frac{199}{246}.$$

Ibland hjälper det att dela upp ett problem i mindre bitar där vi enklare kan finna sannolikheterna. Lagen om total sannolikhet ger oss en enkel möjlighet att pussla ihop dessa bitar igen efteråt.

Lagen om total sannolikhet

Sats. Låt $A, B_1, B_2, \ldots, B_n \subset \Omega$ vara händelser sådana att:

(i) $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$;

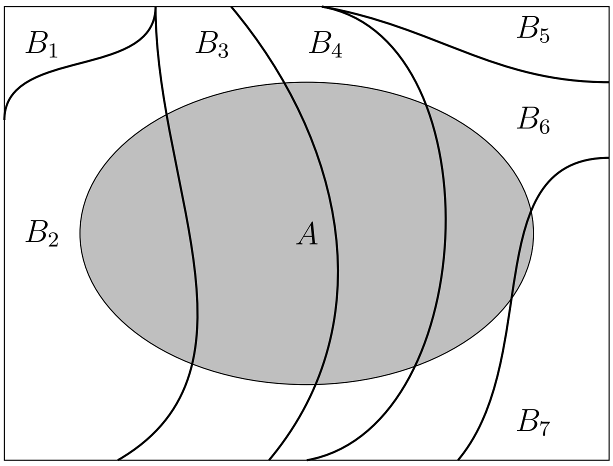
(ii) $P(B_k) \neq 0$ för alla $k = 1, 2, \ldots, n$;

(iii) $B_i \cap B_j = \emptyset$ om $i \neq j$.

Då gäller lagen om total sannolikhet:

$$P(A) = \sum_{k=1}^{n} P(A \,|\, B_k) P(B_k).$$

Figuren nedan visar ett exempel på hur situationen skulle kunna se ut.



Beviset? Ganska enkelt:

$$\sum_{k=1}^{n} P(A \,|\, B_k) P(B_k) = \sum_{k=1}^{n} \frac{P(A \cap B_k)}{P(B_k)} P(B_k) = \sum_{k=1}^{n} P(A \cap B_k) = P(A),$$

där vi använt definitionen av betingad sannolikhet samt Kolmogorovs tredje axiom. Händelserna $A \cap B_k$ är disjunkta för $k = 1, 2, \ldots, n$ eftersom $B_i \cap B_j = \emptyset$ om $i \neq j$.

Bayes sats

Sats. Med samma villkor som för lagen om total sannolikhet gäller

$$P(B_i \,|\, A) = \frac{P(B_i) P(A \,|\, B_i)}{\sum_{k=1}^{n} P(A \,|\, B_k) P(B_k)}$$

för varje $i = 1, 2, \ldots, n$.

Bayes sats är en följd av lagen om total sannolikhet samt definitionen av betingad sannolikhet.

Exempel

Tre maskiner tillverkar prylar. Maskin 1 står för 60%, M-2 för 25% och M-3 för 15%. Av de tillverkade prylarna är 5, 3 respektive 2% felaktiga från de olika maskinerna.
Hur stor är sannolikheten att en på måfå vald enhet är trasig? Om en enhet visar sig vara trasig, hur stor är sannolikheten att den kommer från maskin ett?

Lösning: Enligt lagen om total sannolikhet får vi

$$\begin{aligned} P(\text{trasig pryl}) &= P(M_1)\cdot P\big(\text{trasig} \mid M_1\big) + P(M_2)\cdot P\big(\text{trasig} \mid M_2\big) \\ &\quad + P(M_3)\cdot P\big(\text{trasig} \mid M_3\big) \\ &= 0.6\cdot 0.05 + 0.25\cdot 0.03 + 0.15\cdot 0.02 = 0.0405. \end{aligned}$$

Den andra frågan kan vi svara på mha Bayes sats:

$$P\big(M_1 \mid \text{trasig}\big) = \frac{P(M_1)\cdot P\big(\text{trasig} \mid M_1\big)}{P(\text{trasig})} = \frac{0.6\cdot 0.05}{0.0405} \approx 0.741.$$

Sensitivitet och specificitet

Ett forensiskt test för narkotivapåverkan har sensitivitet 0.9999 (positivt utslag vid påverkan) och specificitet 0.995 (negativt uslag om inte påverkad). Antag att den forensiska analytikern får tillbaka beskedet ”positivt utslag” vid en undersökning. Vad är sannolikheten att personen i fråga faktiskt var påverkad?
Betrakta två grupper. Om personen kommer från grupp ett bedöms sannolikheten att personen är påverkad till 20%, och i grupp två bedöms motsvarande sannolikhet till 0.1%.

Låt A vara händelsen att testet är positivt och B sannolikheten att personen är påverkad. Vi låter $P(B) = p$. Bayes sats medför att

$$\begin{aligned} P\big(B \mid A\big) &= \frac{P(B)P\big(A \mid B\big)}{P(B)P\big(A \mid B\big) + P(B^*)P\big(A \mid B^*\big)} = \frac{p\cdot 0.9999}{p\cdot 0.9999 + (1-p)\big(1 - P\big(A^* \mid B^*\big)\big)} \\ &= \frac{p\cdot 0.9999}{p\cdot 0.9999 + (1-p)0.05} = \frac{1}{1 + \left(\frac{1}{p} - 1\right)\frac{0.05}{0.9999}} \end{aligned}$$

Här har vi utnyttjat att $P(A\cap B^*) = P(A) - P(A\cap B)$.
Om $p = 0.2$ så är $P\big(B \mid A\big) \approx 0.83$ och om $p = 0.001$ så är $P\big(B \mid A\big) = 0.020$.

Observera att det alltså *inte* är 99.999% chans att positivt utslag innebär påverkan. Vad skulle krävas för att detta skulle gälla?