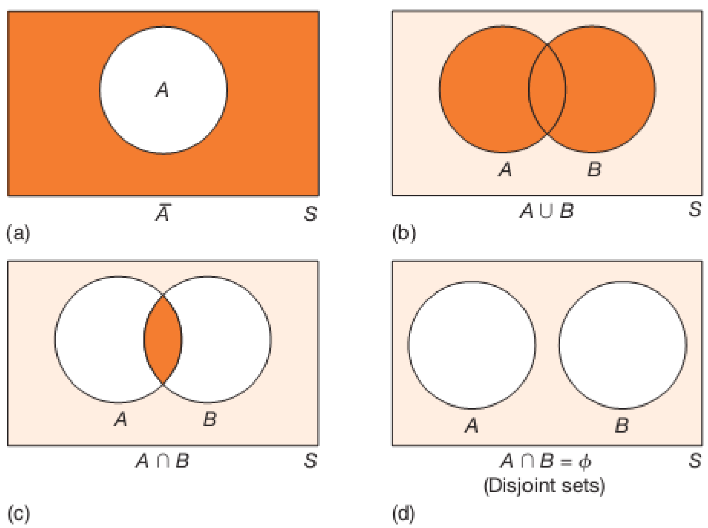
Introduktion til sandsynlighedsteori

Læsning:

Cooper/McGillem kap. 1.7-1.10

Opsummering - axiomer

1. $\Pr(A) \geq 0$
2. $\Pr(S) = 1$
3. Hvis $A \cap B = \emptyset$, så er $\Pr(A \cup B) = \Pr(A) + \Pr(B)$



Tænk på sandsynligheder som arealer:

- Arealet af S er 1.
- Sandsynligheden for hændelsen A i figur (a) er arealet af cirklen A .
- Sandsynligheden for hændelsen AUB i figur (d) er summen af arealerne af cirklerne A og B.
- Hvad er sandsynligheden for hændelsen AUB i figur (b)?

Opsummering - regneregler

- Relativ frekvens $\Pr(A) \approx N_A/N$

- Sandsynligheden for komplementet $\bar{A}$ $\Pr(\bar{A}) = 1 - \Pr(A)$

- Sandsynligheden for den simultane hændelse A <u>og</u> B, skrevet (A,B) eller (A∩B)
$$\begin{aligned}\Pr(\mathrm{A},\mathrm{B}) &= \Pr(\mathrm{B}) \cdot \Pr(\mathrm{A}|\mathrm{B}) \\ &= \Pr(\mathrm{A}) \cdot \Pr(\mathrm{B}|\mathrm{A})\end{aligned}$$

- Hændelserne A og B er uafhængige, hvis $\Pr(A,B) = \Pr(A) \cdot \Pr(B)$

- Sandsynligheden for den sammensatte hændelse A <u>eller</u> B, skrevet (A∪B)
$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

- Hvis A og B er indbyrdes disjunkte: $\Pr(A \cup B) = \Pr(A) + \Pr(B)$

Eksempel

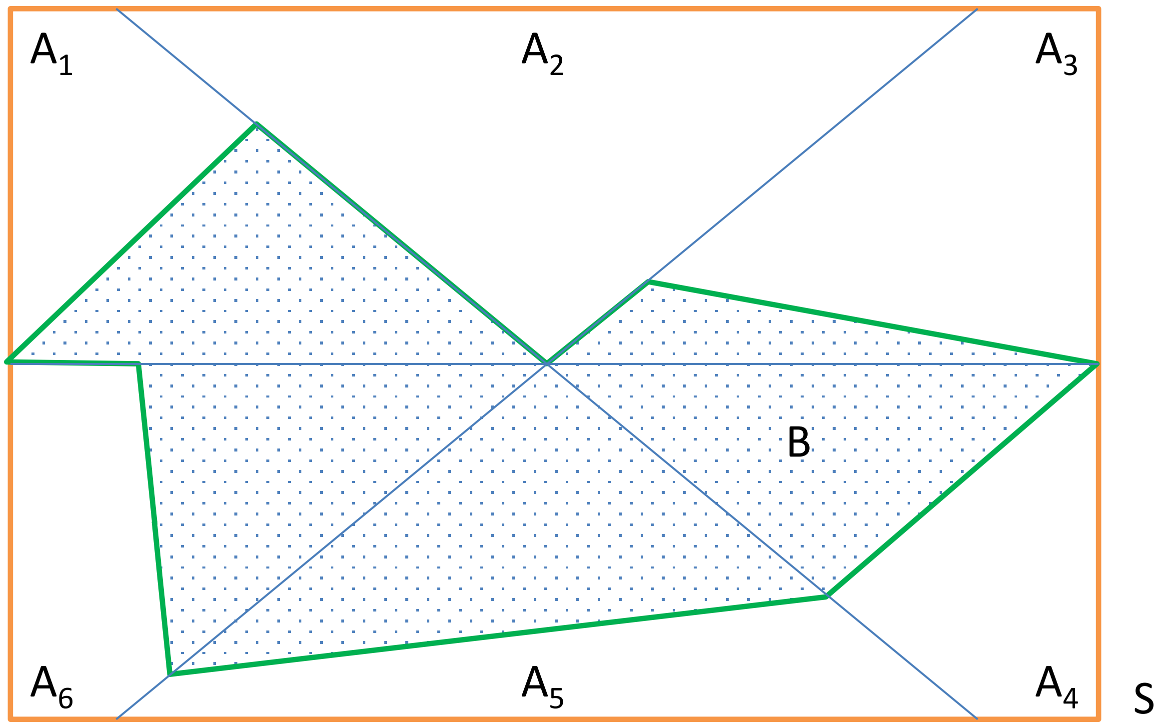
\ Bin	1	2	3	4	5	6	Total
10Ω	500	0	200	800	1200	1000	3700
100Ω	300	400	600	200	800	0	2300
1000Ω	200	600	200	600	0	1000	2600
Total	1000	1000	1000	1600	2000	2000	8600

Tabel 1-3

- Hændelser
  - A_i = vælg den i'te bin (i = 1,2,…,6), hvor vi har valgt $\Pr(A_i) = 1/6$.
  - B = træk en modstand på 10Ω
- Hvad er $\Pr(B)$?
- Hvad er $\Pr(B \cap A_i)$?
- Hvad er $\Pr(A_i|B)$?

Venn diagram

\ Bin	1	2	3	4	5	6	Total
10Ω	500	0	200	800	1200	1000	3700
100Ω	300	400	600	200	800	0	2300
1000Ω	200	600	200	600	0	1000	2600
Total	1000	1000	1000	1600	2000	2000	8600



A_i = vælg den i'te bin (i = 1,2,…,6), hvor vi har valgt $Pr(A_i) = 1/6$.
B = træk en modstand på 10Ω

Total probability

- $S = A_1 \cup A_2 \cup \cdots \cup A_n$, hvor A_i'erne er indbyrdes disjunkte.

- $B = (B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n)$

- $\Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \cdots + \Pr(B \cap A_n)$

$= \Pr(B|A_1)\Pr(A_1) + \Pr(B|A_2)\Pr(A_2) + \cdots + \Pr(B|A_n)\Pr(A_n)$

Eksempel – total probability

\ Bin	1	2	3	4	5	6	Total
10Ω	500	0	200	800	1200	1000	3700
100Ω	300	400	600	200	800	0	2300
1000Ω	200	600	200	600	0	1000	2600
Total	1000	1000	1000	1600	2000	2000	8600

Tabel 1-3

$$\begin{aligned}\Pr(B) &= \Pr(B|A_1)\Pr(A_1) + \Pr(B|A_2)\Pr(A_2) + \cdots + \Pr(B|A_6)\Pr(A_6) \\ &= \frac{500}{1000}\cdot\frac{1}{6} + 0\cdot\frac{1}{6} + \frac{200}{1000}\cdot\frac{1}{6} + \frac{800}{1600}\cdot\frac{1}{6} + \frac{1200}{2000}\cdot\frac{1}{6} + \frac{1000}{2000}\cdot\frac{1}{6} = 0{,}3833\end{aligned}$$

hvor

$$\Pr(B|A_1) = \frac{500}{1000},\ \Pr(B|A_2) = 0,\ \Pr(B|A_3) = \frac{200}{1000},$$

$$\Pr(B|A_4) = \frac{800}{1600},\ \Pr(B|A_5) = \frac{1200}{2000},\ \Pr(B|A_6) = \frac{1000}{2000}$$

Bayes regel

- Hvis $S = A_1 \cup A_2 \cup \cdots \cup A_n$, og A_i'erne er indbyrdes disjunkte:

$$\Pr(B) = \Pr(B|A_1)\Pr(A_1) + \Pr(B|A_2)\Pr(A_2) + \cdots + \Pr(B|A_n)\Pr(A_n)$$

- Observation:

$$\Pr(B \cap A_i) = \Pr(B|A_i)\Pr(A_i) = \Pr(A_i|B)\Pr(B)$$

- **Bayes regel:**

$$\Pr(A_i|B) = \frac{\Pr(B|A_i)\Pr(A_i)}{\Pr(B)}$$

Eksempel – Bayes regel

\ Bin	1	2	3	4	5	6	Total
10Ω	500	0	200	800	1200	1000	3700
100Ω	300	400	600	200	800	0	2300
1000Ω	200	600	200	600	0	1000	2600
Total	1000	1000	1000	1600	2000	2000	8600

Tabel 1-3

- Hændelser
  - A_i = vælg den i'te bin (i = 1,2,…,6), hvor vi har valgt $\Pr(A_i) = 1/6$.
  - B = træk en modstand på 10Ω
- Hvad er $\Pr(A_3|B)$?
- Bayes regel:

$$\Pr(A_3|B) = \frac{\Pr(B|A_3) \cdot \Pr(A_3)}{\Pr(B)} = \frac{\frac{200}{1000} \cdot \frac{1}{6}}{0{,}3833} = 0{,}0869$$

Cancer test

- Prior

  $\Pr(cancer) = 0.01 \quad \rightarrow \Pr(\overline{cancer}) = 0.99$

- Test

  $\Pr(positiv\ test|cancer) = 0.9$ (sensitivitet)

  $\Pr(negativ\ test|\overline{cancer}) = 0.9$ (specificitet)

  $\rightarrow$

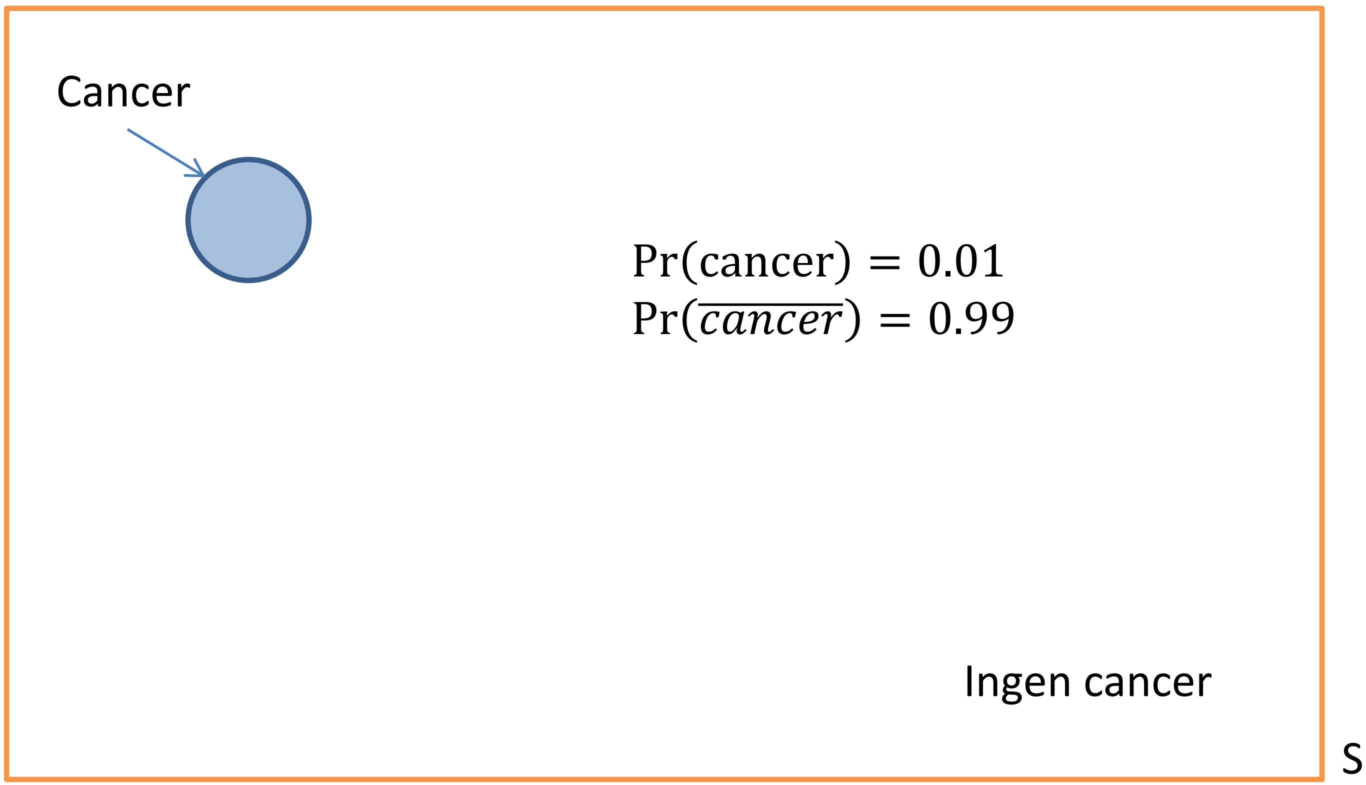
  $\Pr(negativ\ test|cancer) = 0.1$

  $\Pr(positiv\ test|\overline{cancer}) = 0.1$

- Det store spørgsmål

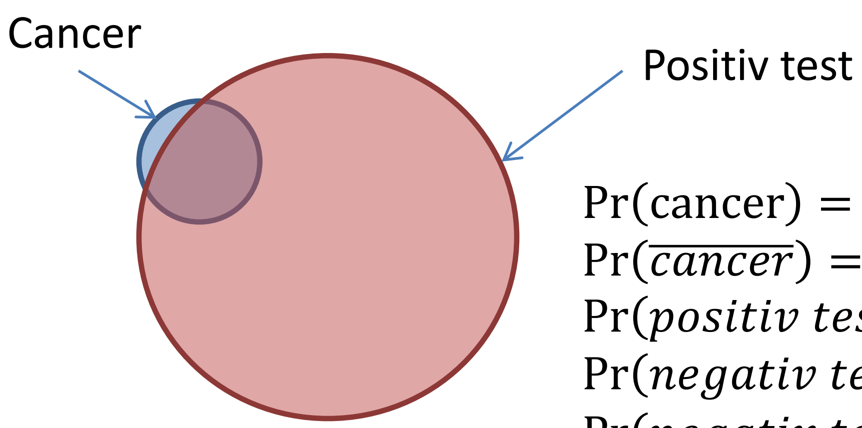
  Hvad er $\Pr(cancer|positiv\ test)$?

Venn diagram



$$\Pr(\text{cancer}) = 0.01$$
$$\Pr(\overline{cancer}) = 0.99$$

Venn diagram



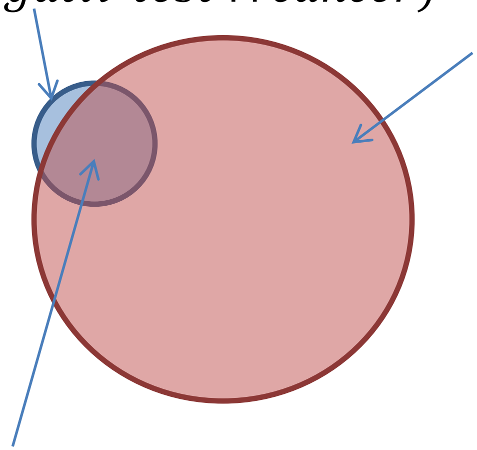
$$\Pr(\text{cancer}) = 0.01$$
$$\Pr(\overline{cancer}) = 0.99$$
$$\Pr(positiv\ test|cancer) = 0.9$$
$$\Pr(negativ\ test|\overline{cancer}) = 0.9$$
$$\Pr(negativ\ test|cancer) = 0.1$$
$$\Pr(positiv\ test|\overline{cancer}) = 0.1$$

Ingen cancer

S

Venn diagram

$\Pr(negativ\ test \cap cancer)$



$\Pr(positiv\ test \cap \overline{cancer})$

$\Pr(\text{cancer}) = 0.01$
$\Pr(\overline{cancer}) = 0.99$
$\Pr(positiv\ test|cancer) = 0.9$
$\Pr(negativ\ test|\overline{cancer}) = 0.9$
$\Pr(negativ\ test|cancer) = 0.1$
$\Pr(positiv\ test|\overline{cancer}) = 0.1$

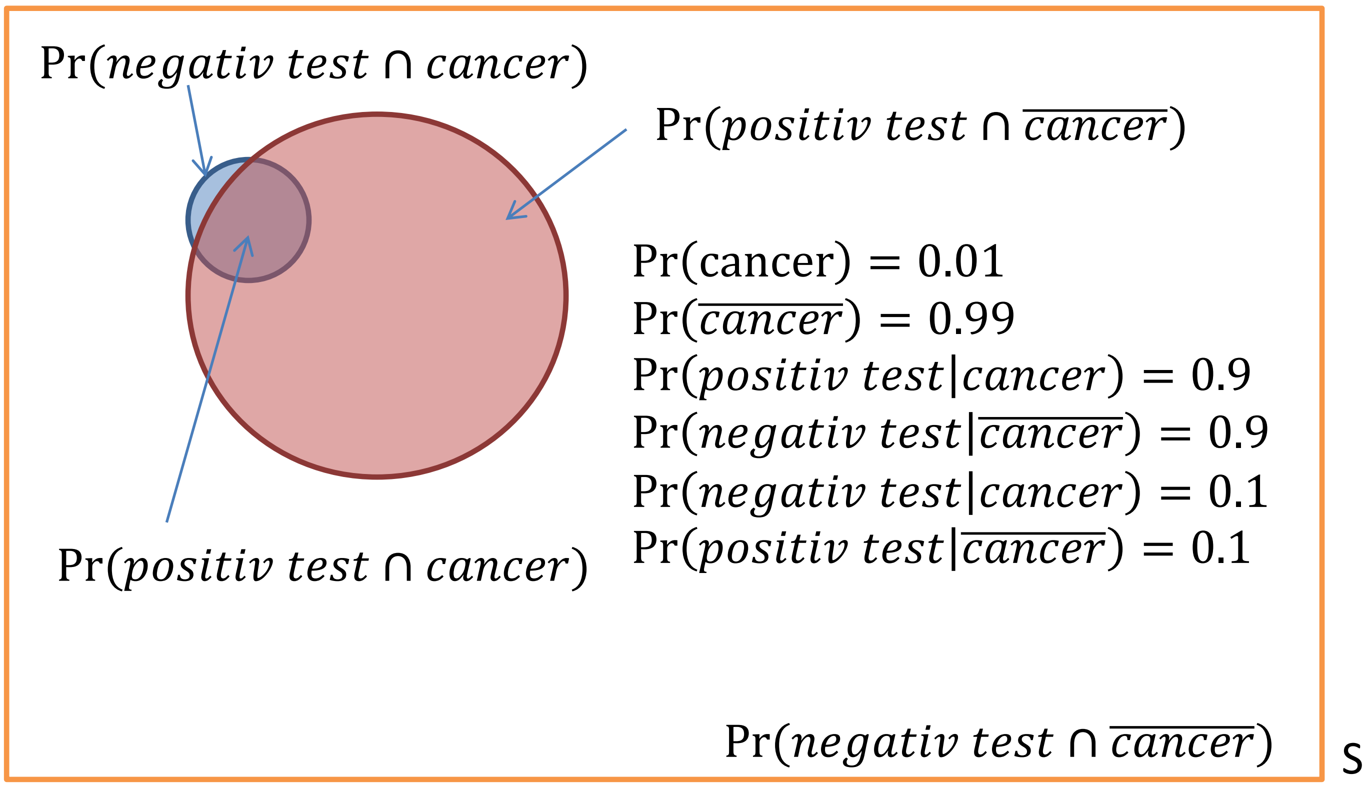
$\Pr(positiv\ test \cap cancer)$

$\Pr(negativ\ test \cap \overline{cancer})$

S

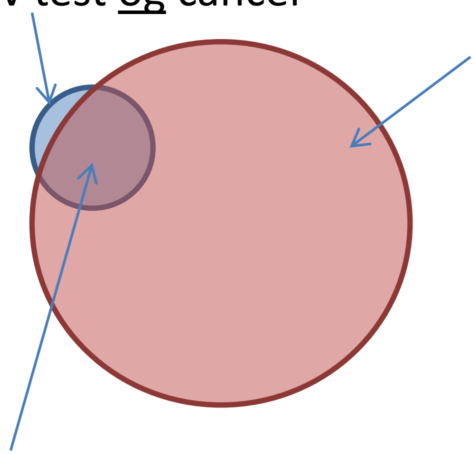
<u>**Ex:**</u> $\Pr(\text{neg. test} \cap cancer) = Pr(neg.test)Pr(cancer|neg.test)$
$= Pr(cancer)Pr(neg.test|cancer)$

Venn diagram



Det store spørgsmål: Hvad er $\Pr(cancer|positiv\ test)$?

Bayes regel



$$\Pr(\text{cancer}) = 0.01$$
$$\Pr(\overline{cancer}) = 0.99$$
$$\Pr(positiv\ test|cancer) = 0.9$$
$$\Pr(negativ\ test|\overline{cancer}) = 0.9$$
$$\Pr(negativ\ test|cancer) = 0.1$$
$$\Pr(positiv\ test|\overline{cancer}) = 0.1$$

$$\Pr(cancer|positiv\ test) = \frac{\Pr(positiv\ test|cancer)\Pr(cancer)}{\Pr(positiv\ test)}$$

Bayes regel

$$\Pr(cancer|positiv\ test) = \frac{\Pr(positiv\ test|cancer)\Pr(cancer)}{\Pr(positiv\ test)}$$

$$= \frac{\Pr(positiv\ test|cancer)\Pr(cancer)}{\Pr(positiv\ test \cap cancer) + \Pr(positiv\ test \cap \overline{cancer})}$$

$$= \frac{\Pr(positiv\ test|cancer)\Pr(cancer)}{\Pr(positiv\ test|cancer)\Pr(cancer) + \Pr(positiv\ test|\overline{cancer})\Pr(\overline{cancer})}$$

$$= \frac{0.9 \cdot 0.01}{0.9 \cdot 0.01 + 0.1 \cdot 0.99} = 0.0833 \approx 8\%$$

$\Pr(\text{cancer}) = 0.01$
$\Pr(\overline{cancer}) = 0.99$
$\Pr(positiv\ test|cancer) = 0.9$
$\Pr(negativ\ test|\overline{cancer}) = 0.9$
$\Pr(negativ\ test|cancer) = 0.1$
$\Pr(positiv\ test|\overline{cancer}) = 0.1$

Binomialfordelingen

- Vi laver n gentagne eksperimenter, hvor der i hvert trial er to mulige udfald:
  - Success med sandsynlighed p
  - Failure med sandsynlighed $q = 1 - p$

- $$\begin{aligned} \Pr_n(k) &= \Pr(k \text{ succeser ud af } n \text{ forsøg}) \\ &= \frac{n!}{k!(n-k)!} p^k q^{n-k} \\ &= \binom{n}{k} p^k q^{n-k} \end{aligned}$$