

## Projekt 9.8 Betingede sandsynligheder og paradokser i sandsynlighedsregningen

Et forløb om betingede sandsynligheder kan introduceres via et selvstændigt elevarbejde med materialet i projekt 9.7 – "Testet positiv?" sammen med eller i stedet for den indledende case om at "scanne for terrorisme".

### 1. Case: Potentielle terrorister og paradokset om de falsk positive

(Eleverne arbejder selv dette eksempel igennem som en opvarmning til emnet. Eksempel er lånt fra den canadiske forfatter Cory Doctorow, der diskuterer paradokset i sin bog *Little Brother*)

Antag man har opdaget en ny og meget sjælden sygdom, som får navnet Super-AIDS. Sygdommen optræder med en hyppighed på ca 1 tilfælde ud af en million. Der udvikles en test der 99% sikker, hvormed menes, at den giver det korrekte resultat 99 ud af 100 gange – sandt, hvis man faktisk er smittet og falsk hvis man ikke er smittet. Eller sagt omvendt: Den giver et falsk resultat i 1% af tilfældene.

Sygdommen anses for ekstremt farlig, så det besluttes at give testen til 1 million indbyggere.

- a) Udfyld en tabel som den følgende (med afrundede tal, det er jo hele mennesker):

test \ tilstand	har super-AIDS	har ikke super-AIDS	I alt
positivt udslag	1		
ikke positivt udslag	0		
I alt	1	999.999	1.000.000

Et rimeligt mål for, hvor præcis testen er, kunne være at angive hvor stor en procentandel af de positivt testede, der faktisk er syge.

- b) Hvor præcis er denne test?  
 c) Sammenhold dette tal med, at testen blev præsenteret som 99% sikker. Hvori ligger forklaringen på dette paradoks?

Når vi måler på meget små størrelser, skal vores måleudstyr også være meget fintmasket, eller være indrettet på at kunne registrere noget meget småt. Vil man pege på en enkelt pixel på sin skærm, kan en spids blyant godt bruges. Men blyanten er ikke anvendelig, hvis man skulle pege på et enkelt atom.

Lad os trække en parallel fra den sjældne sygdom "super-AIDS" til den aktuelle debat om overvågning af potentielle terrorister. Kan man ved at sammenkøre store datamængder fra mobiltelefoni, banktransaktioner, rejsemønstre, aktiviteter på sociale medier som fx Facebook mv finde potentielle terrorister?

Lad os antage, at efterretningsvæsenet har gode argumenter for at hævde, at hvis bestemte indikatorer er tilstede hos en person, så er der en vis sandsynlighed for, at vedkommende er potentiel terrorist. Og omvendt: Hvis en person faktisk er terrorist, så fanges de i filtret med 99% sikkerhed. Det afgørende i første omgang er ikke, om vi tror på sådanne indikatorer. Vi antager det er korrekt og analyserer metoden ud fra denne antagelse. Da vi ikke kan have sikker viden, men taler om *sandsynligheder*, og da vi taler om *potentielle* terrorister må man acceptere, at der også spottes nogen, der faktisk ikke er terrorister.

## Hvad er matematik? 1

ISBN 9788770668279

Projekter: Kapitel 9. Projekt 9.8 Betingede sandsynligheder og paradokser i sandsynlighedsregningen

Lad os antage, at en organisation som det amerikanske efterretningsvæsen, NSA har skaffet sig adgang til alle bankkonti, til overvågning af alle mobil samtaler i byen, til at kunne scanne alle facebook profiler mv. De har lagt filtre ind, der frikender 99,9%, mens 0,1% af befolkningen matcher NSA's definition af "potentielle terrorister".

Rigtige terrorister, der eksempelvis er villige til at optræde i selvmordsangreb, er sjældne. I en by som New York på 20 millioner indbyggere anslås der at være højst 10 sådanne terrorister. Det betyder mindre for det følgende, om dette tal er fx en faktor 10 større, men der er i vestlige lande trods alt kun set ganske få tilfælde af denne type, selv om det ville være en ret enkel sag at udføre.

- a) Anvend de givne oplysninger til at færdigudfylde følgende tabel over "scanningen" af New York for potentielle terrorister:

test \ tilstand	er terrorist	er ikke terrorist	I alt
positivt udslag	10		
ikke positivt udslag	0		
I alt	10		20.000.000

De potentielle terrorister opsamles på en liste, og en whistle-blower lækker listen til pressen. Det viser sig din nabo figurerer op listen.

- b) Diskuter i gruppen, hvordan I ville reagere på en sådan oplysning.  
c) Hvad er sandsynligheden for at en tilfældig person på listen faktisk er terrorist?  
d) Diskuter i gruppen den beskrevne metode til at spotte potentielle terrorister.

En test giver aldrig sikker viden. Derfor opererer man i statistik med følgende fire begreber:

*falsk positiv, falsk negativ, sand positiv, sand negativ*

- e) Placer de fire begreber i de fire rubrikker i tabellen og argumenter for hvordan du placerer dem.

## 2 Betingede sandsynligheder

Betingede sandsynligheder er sandsynligheder, der beregnes ud fra en eller anden given betingelse. Hvis man kaster med to terninger, så er sandsynligheden for at få en øje-sum på 10 lig med  $\frac{3}{36} = \frac{1}{12}$ , fordi der 36 kombinationer af to terninger og præcis kombinationerne (6,4), (5,5) og (4,6) giver summen 10. Hvis man nu ved, at en af terningerne er en 6'er, så er der i alt 11 kombinationer, hvoraf de to, nemlig (6,4) og (4,6) giver summen 10. Dvs med den nye viden er sandsynligheden ændret til  $\frac{2}{11}$ .

Dette kan vi udtrykke i et formelsprog, der viser sig nyttigt, på følgende måde:

### Definitioner og notation vedr sandsynlighedsfelter

1. Den samlede mængde af kombinationer kalder vi *udfaldsrummet* og betegner det  $U$ :

$$U = \{(1,1), (1,2), \dots, (1,6), (2,1), (2,2), \dots, (2,6), \dots, (6,1), (6,2), \dots, (6,6)\}$$

Generelt betegner  $u_i$  et *udfald* og mængden af alle udfald kaldes for *udfaldsrummet*:

$$U = \{u_1, u_2, \dots, u_n\}$$

2. *Hændelser* er delmængder af udfaldsrummet, og betegnes ofte med store bogstaver:

$A$  = alle kombinationer, der giver øje-summen 10:

$$A = \{(6,4), (5,5), (4,6)\}$$

$B$  = alle kombinationer, hvor en af terningerne viser 6:

$$B = \{(1,6), (2,6), \dots, (6,6), (6,1), (6,2), \dots, (6,5)\}$$

3. *Sandsynligheder* angives med en sandsynlighedsfunktion  $P$ , således:

$$P((5,5)) = \frac{1}{36}, \quad P(A) = \frac{3}{36}, \quad P(B) = \frac{11}{36}, \quad P(U) = \frac{36}{36} = 1, \quad P(\text{Ikke-}A) = \frac{33}{36}$$

Generelt gælder, at

$$0 \leq P(u_i) \leq 1 \text{ samt at } \sum_1^n u_i = u_1 + u_2 + \dots + u_n = 1$$

4. Hvis alle udfald har samme sandsynlighed, som det er tilfældet med terningekast med én terning eller med to terninger (sort og rød fx), så siger vi, at vi har et *symmetrisk sandsynlighedsfelt*. I et symmetrisk sandsynlighedsfelt kan vi udregne sandsynligheden af en hændelse, som fx  $A$ , ved

$$\text{formlen: } P(A) = \frac{\text{antal udfald i } A}{\text{samlede antal udfald i } U}$$

5. Symbolet  $A \cap B$  angiver, at to hændelser som  $A$  og  $B$  *begge indtræffer*. Det læses af og til: "både  $A$  og  $B$ ".  $A \cap B$  kaldes også fællesmængden af  $A$  og  $B$ . I vort eksempel er  $A \cap B = \{(6,4), (4,6)\}$ .

6. Når vi regner *ud fra en given betingelse*, som fx at vi ved at en af terningerne viser 6, dvs at  $B$  er indtruffet, så angiver vi det således:  $P(A|B)$ . Dette betyder: sandsynligheden for  $A$  når vi ved  $B$  er indtruffet. Det kaldes også *den betingede sandsynlighed for  $A$  givet  $B$*

Hvordan udregnes så  $P(A|B)$ ? Når vi ved, at  $B$  er indtruffet, så er de mulige udfald altså ikke hele  $U$ , men alene  $B$ . Og når vi ved, at  $B$  er indtruffet, så er den hændelse, vi spørger om, ikke hele  $A$ , men  $A \cap B$ .

Derfor kan vi i symmetriske sandsynlighedsfelter, som eksemplet med kast med to terninger, udregne:

$$P(A|B) = \frac{\text{antal udfald i } A \cap B}{\text{samlede antal udfald i } B} = \frac{2}{11}$$

Denne formel kan vi omskrive lidt (forkort brøken, dvs divider både tæller og nævner med samme tal):

$$P(A|B) = \frac{(\text{antal udfald i } A \cap B) / (\text{antal udfald i } U)}{(\text{samlede antal udfald i } B) / (\text{antal udfald i } U)} = \frac{P(A \cap B)}{P(B)}$$

Med de givne talværdier ville den sidste udregning give:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/36}{11/36} = \frac{2}{11}$$

altså naturligvis det samme som ovenfor.

Men årsagen til, at vi foretager denne omskrivning er, at i den sidste formel har vi sluppet optællingen af antal, som er knyttet til symmetriske sandsynlighedsfeter. Denne formel kan vi derfor anvende som den generelle definition af *den betingede sandsynlighed for A givet B*:

**Definition: Betinget sandsynlighed**

Den betingede sandsynlighed for A givet B betegnes  $P(A|B)$  og er givet ved:

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

Udtrykket fortolkes som: *Sandsynligheden for A når vi ved B er indtruffet.*

Argumentationen

ovenfor fortæller, at dette er en generalisering af tællemetoden fra symmetriske felter.

**Bemærkning om uafhængighed:**

Betingede sandsynligheder giver anledning til at give en formel definition på et centralt begreb i sandsynlighedsregning, nemlig begrebet *uafhængige hændelser*:

Hændelserne  $A$  og  $B$  kaldes uafhængige, hvis der gælder, at  $P(A|B) = P(A)$ . Dvs den ekstra oplysning om, at  $B$  er indtruffet påvirker ikke sandsynligheden for at  $A$  indtræffer.

Hvis  $A$  og  $B$  er uafhængige ser vi af formlen, at der gælder:  $P(A \cap B) = P(B) \cdot P(A)$

Nogle lærebøger bruger faktisk den sidste formel som en definition af uafhængighed. Men dermed mister man intuitionen om uafhængighed. Det er imidlertid vigtigt at holde fast i, at det er en formel definition. Begrebet indgår jo også i daglig sproget, og her skal man passe på ikke kun at forlade sig på sin intuition.

Begrebet uafhængige hændelser er helt central i behandlingen af de såkaldte binomialmodeller, der er behandlet i B-bogens kapitel 9. Vi vil ikke gå yderligere ind i dette her.

En af de stærke sider ved betingede sandsynligheder er, at man kan "regne baglæns"

**Sætning: At regne forlæns og baglæns med betingede sandsynligheder**

Hvis  $A$  og  $B$  er to hændelser i et udfaldsrum  $U$ , så gælder:

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

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

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

Beviset overlades til læseren.

Bemærk, at den tredje formel vender betingelsen om, og dermed giver os mulighed for at "regne baglæns".

I de følgende eksempler vil vi både arbejde med formlerne og med tabelopstillingerne, der ofte er et redskab til forholdsvis enkle løsninger.

## Hvad er matematik? 1

ISBN 9788770668279

Projekter: Kapitel 9. Projekt 9.8 Betingede sandsynligheder og paradokser i sandsynlighedsregningen

### Øvelse 1. Hvilket køn har det andet barn?

(Farvningen nedenfor kan du ophæve, og se svarene)

(Vi antager i dette eksempel, at der fødes lige mange piger og drenge i Danmark)

I et nabo hus flytter et par ind, som har to børn. Hvad er sandsynligheden for at begge er drenge? Hvad er sandsynligheden for, at mindst én er en dreng?

Udfaldsrummet her er  $U_1 = \{(P,P), (P,D), (D,P), (D,D)\}$ , hvor rækkefølgen i talparrene angiver i hvilken rækkefølge børnene blev født.  $P$  og  $D$  står for pige og dreng. Svarene er naturligvis:

$$P(\text{to drenge}) = \frac{1}{4} = 25\% \quad P(\text{mindst én dreng}) = \frac{3}{4} = 75\%$$

Du møder parret, der er ude at gå. De har det ene barn med sig. Det er en dreng.

Hvilket køn har det andet barn? Hvad er sandsynligheden for, at det andet barn også er en dreng?

25%?                      50%?                      Et helt andet tal?

Umiddelbart vil mange svare det første – er det andet barn også en dreng er der jo to, og vi har lige udregnet, at sandsynligheden for to drenge er 25%.

Andre hælder måske til 50%: Når vi ikke kender kønnet, så må det være fifty-fifty for dreng-pige.

Men begge svar er forkert. De tager ikke hensyn til at vi har fået en viden i og med vi nu ved, at det ene barn er en dreng.

Udregnet ved tællemetoden: Udfaldsrummet er nu  $U_2 = \{(P,D), (D,P), (D,D)\}$ . Så sandsynligheden for, at den anden også er en dreng er således:

$$P(\text{to drenge} | \text{en dreng}) = \frac{1}{3} = 33,3\%.$$

Da du kommer hjem fortæller du, at du mødte de nye naboer, og at du så, de har en dreng. Ja, og så har en mere, der ligger i barnevogn, lød svaret. Ændrer den nye oplysning på svaret på spørgsmålet: Hvad er sandsynligheden for, at det andet barn også er en dreng?

Mange vil nok svare, at det ikke ændrer noget. Men oplysningen rummer faktisk en ny information. Nu ved vi ikke blot, at den ene af de to børn er en dreng, men også at det er den ældste. Dvs udfaldsrummet er nu

$U_3 = \{(D,P), (D,D)\}$ . Så sandsynligheden for, at den anden også er en dreng er således:

$$P(\text{to drenge} | \text{den ældste er en dreng}) = \frac{1}{2} = 50\%$$

Læg mærke til, at sandsynligheden for at hændelsen indtræffer (her: to drenge) stiger med den information vi får.

### Øvelse 2. Testet positiv - er du syg?

Antag vi har en test for HIV, der er rimelig effektiv, idet den fanger 90% af alle der er smittede. Testen har således en falsk negativ rate på 10%. Men testen fanger ikke kun de syge, den har også en falsk positiv rate på 5%. Vi har en samlet population på 1000, hvoraf i alt 2% er smittede.

Du testes positiv. Hvad er sandsynligheden for at du faktisk er smittet?

Når man skal svare på sådanne spørgsmål, kan det være en fordel at stille oplysningerne op i en tabel som følger:

test \ tilstand	HIV smittet	ikke HIV smittet	I alt
positivt udslag	18	49	67
ikke positivt udslag	2	931	933
I alt	20	980	1000

## Hvad er matematik? 1

ISBN 9788770668279

Projekter: Kapitel 9. Projekt 9.8 Betingede sandsynligheder og paradokser i sandsynlighedsregningen

- Løs opgaven ved simpel optælling
- Hvordan vil du definere mængderne A og B, omtalt i definitionen på betingede sandsynligheder? Hvad udgør mængden  $A \cap B$ ?
- Hvad er  $P(A)$ ,  $P(B)$  og  $P(A \cap B)$ ?
- Udnyt nu formelen  $P(A|B) = \frac{P(A \cap B)}{P(B)}$  til at løse opgaven.

I et tilfælde hvor vi har givet absolutte tal, som i øvelsen ovenfor, er det klart lettere at optælle. Tabellen er et nyttigt redskab til at skabe overblik, så vi ser, at vi kun behøver at regne i den øverste række.

Hvis vi ikke har absolutte tal, kun %-tal., så kunne vi naturligvis gennemregne med et taleksempel, men kunne vi ikke klare os uden? Det handler næste eksempel om.

### Eksempel. Klassikeren fra Harvard Medical School

Mange amerikanske lærebøger om statistik indeholder følgende eksempel på, hvor let det er at slutte forkert i statistik:

Consider the following problem

*A particular heart disease has a prevalence of 1/1000 people. A test to detect this disease has a false positive rate of 5%. Assume that the test diagnoses correctly every person who has the disease. What is the chance that a randomly selected person found to have a positive result actually has the disease?*

This question was put to 60 students and staff at Harvard Medical School.

Almost half gave the response 95%.

The average answer was 56%.

The correct answer was given by just 11 participants.

### 1. Løsningsmetode

Sygdommen rammer en ud af 1000. Derfor opstiller vi en antalstabel baseret på en population på 1000. Det grønne felt (falsk negativ) rummer 0, da testet fanger alle der faktisk er syge.

Da der er en syg, er der 999 ikke-syge. Antal falsk positive er så 5% af 999. Dette er fundet lig med 50, så i alt testes 51 positivt.

Betragt nu tabellen. En person er testet positiv og er altså blandt de 51. I denne gruppe er der 1 syg.

Så sandsynligheden for at vedkommende har sygdommen er:

$$P(\text{syg} | \text{testet positiv}) = \frac{1}{51} = 1,96\% \approx 2\%$$

test \ tilstand	syg	ikke syg	I alt
positivt udslag	1	50	51
ikke positivt udslag	0	949	949
I alt	1	999	1000

2. Løsningsmetode

Vi formaliserer oplysningerne:

A: Mængden af syge	$P(A) = 0,001$	
not A: Mængden af raske	$P(\text{not } A) = 0,999$	
B: Mængden der testes positiv.	$P(B   A) = 1$	$P(B   \text{not } A) = 0.05$

En person er testet positiv og ønsker at kende sandsynligheden for at vedkommende faktisk er syg.

Dvs. i formelsproget ønsker vi at beregne  $P(A | B)$ .

Udnyt formlen: 
$$P(A | B) = \frac{P(A)}{P(B)} \cdot P(B | A) \quad (*)$$

Vi kan se, at vi her mangler kendskab til  $P(B) = P(\text{positiv test})$ .

Da alle individer enten er syg eller rask, dvs enten ligger i A eller i not A, så kan vi opdele B i dem der ligger i A, dvs  $B \cap A$  og dem der ligger i not A, dvs  $B \cap \text{not } A$ . Dermed kan vi udregne  $P(B)$  således:

$$P(B) = P(B \cap A) + P(B \cap \text{not } A)$$

Udnyt sætningen om at regne baglæns til omskrivningen:

$$\begin{aligned} P(B) &= P(B \cap A) + P(B \cap \text{not } A) \\ &= P(B | A) \cdot P(A) + P(B | \text{not } A) \cdot P(\text{not } A) \\ &= 1 \cdot 0,001 + 0,05 \cdot 0,999 \end{aligned}$$

Indsæt nu tallene i formlen (\*):

$$P(A | B) = \frac{P(A)}{P(B)} \cdot P(B | A) = \frac{0,001}{1 \cdot 0,001 + 0,05 \cdot 0,999} \cdot 1 = \frac{1}{1 + 49,5} \approx 2\%$$

Så sandsynligheden for at vedkommende er ca 2%.

Øvelse 3. Elisa testen for screening af blod

I en rapport om udviklingen i bestræbelserne på at bekæmpe AIDS-epidemien kan man læse følgende:

The ELISA test was introduced in the mid 1980's to screen donated blood for the presence of AIDS antibodies. When antibodies are present, ELISA is positive with a probability of about 0.98; when the blood tested is not contaminated with antibodies, the test gives a positive result with a probability of 0.07. These numbers are conditional probabilities. If one in a thousand of the units of blood screened by ELISA contain antibodies, then (??) of all positive responses will be false positive.

(kilde: Lynn Arthur Steen (red.), *New approaches to Numeracy*)

Hvad skal der stå på den tomme plads markeret med (??) (dvs. hvor stor en andel falsk positive er det?)

#### Øvelse 4. Medicinsk forsøg

Et medicinalfirma tilrettelægger en test af en ny allergimedisin. 1500 testpersoner deltager og opdeles i 3 grupper med 500 i hver. Den ene gruppe får det klassiske produkt, firmaet længe har haft på sin liste. Den anden får et placebo-præparat. Og endelig får den tredje gruppe det nye betydeligt stærkere præparat. De enkelte deltagere ved naturligvis ikke hvilke præparater de får.

	forbedring	ingen virkning	forværring	I alt
<b>Gruppe 1</b>	159	301	40	500
<b>Gruppe 2</b>	128	342	30	500
<b>Gruppe 3</b>	318	77	105	500
<b>I alt</b>	605	720	175	

Lad os antage, at stikprøven er repræsentativ for den relevante population.

- Hvilken konklusion – i form af anbefalinger til firmaet - vil du umiddelbart drage om det nye stærkere præparat?
- En tilfældig valgt person blandt de 1500 får det værre. Hvad er sandsynligheden for at han har fået det nye præparat?
- En tilfældig valgt person blandt de 1500 får det bedre. Hvad er sandsynligheden for at han har fået det nye præparat?
- Vil du ændre dine anbefalinger?

#### Øvelse 5. Hvad vej vender betingelsen?

Bayesianske metoder har i stor udstrækning fundet vej til især amerikanske retssale. Det gives vi en kort introduktion til i næste afsnit. I en af de artikler vi henviser til gives følgende eksempel:

Suppose a crime has been committed and that the criminal has left some physical evidence, such as some of their blood at the scene. Suppose the blood type is such that only 1 in every 1000 people has the matching type. A suspect, let's call him *Fred*, who matches the blood type is put on trial. The prosecutor claims that the probability that an innocent person has the matching blood type is 1 in a 1000 (that's a probability of 0.001). Fred has the matching blood type and therefore the probability that Fred is innocent is just 1 in a 1000.

Analyser anklagerens påstand ved hjælp af betingede sandsynligheder. De to centrale spørgsmål handler om blodtypen og om *Fred er uskyldig*, som han påstår. Indfør to hændelser A og B, og opstil et udtryk for, hvad anklageren beregner. Hvad er din konklusion? Du kan evt anvende næste øvelse, som et hint

#### Øvelse 6 $P(A|B)$ eller $P(B|A)$ ?

Du får et billede af en kvinde, der er attraktiv og smukt klædt, og bliver spurgt om sandsynligheden for at hun er fotomodel. Hvad er det egentlig for en betinget sandsynlighed vi spørger om?

Du kan evt løse opgaven ved først at definere følgende hændelser:

A: Kvinden er attraktiv og forstår at klæde sig smukt.

B: Kvinden er fotomodel.

Spørger vi om  $P(A|B)$  eller  $P(B|A)$ ?



### 3. Den generelle udgave af Bayes formel

I analysen af "klassikeren fra Harvard" anvendte vi følgende formel:

$$P(B) = P(B \cap A) + P(B \cap \text{not } A)$$

Formlen bygger på den enkle iagttagelse, at enten indtræffer hændelsen A eller også gør den ikke, dvs. de to situationer  $B \cap A$  og  $B \cap \text{not } A$  udtømmer alle muligheder. Men dette kan vi generalisere til situationer, hvor alle muligheder kan opdeles i adskilte mængder  $A_1, A_2, \dots, A_n$ . Det kan fx være situationen, hvor vi opdeler befolkningen i indkomstgrupper, svarende til  $A_1, A_2, \dots, A_n$ , og hvor hændelsen B kunne være holdningen til om vi i Danmark skal gå over til Euroen. Da er:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_n) \quad (*)$$

Hvis vi nu ønsker at beregne  $P(A_k | B)$ , så kan vi tage udgangspunkt i definitionen:

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

Heri indføres nu først (\*):

$$P(A_k | B) = \frac{P(A_k \cap B)}{P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_n)}$$

Ofte har vi situationer, hvor vi kender sandsynligheder for de omvendte betingelser, fx  $P(B | A_k)$ , og ved at anvende de tidligere formler, kan vi nu omskrive til følgende

$$P(A_k | B) = \frac{P(B | A_k) \cdot P(A_k)}{P(B | A_1) \cdot P(A_1) + P(B | A_2) \cdot P(A_2) + \dots + P(B | A_n) \cdot P(A_n)}$$

#### Sætning: Bayes formel

Hvis hændelserne A og B er to hændelser i et udfaldsrum U og hændelsen A kan opdeles i n adskilte hændelser / delmængder,  $A_1, A_2, \dots, A_n$ , så gælder:

$$P(A_k | B) = \frac{P(B | A_k) \cdot P(A_k)}{P(B | A_1) \cdot P(A_1) + P(B | A_2) \cdot P(A_2) + \dots + P(B | A_n) \cdot P(A_n)}$$

#### Øvelse 7. Bliver der regn på bryllupsdagen

En amerikansk kvinde Marie skal giftes ved en spektakulær udendørs ceremoni i et ørkenområde uden for Las Vegas. De senere år har det kun regnet 5 dage om året. Uheldigvis har en af TV stationernes meteorologer forudsagt, at det bliver regn, netop på bryllupsdagen. Når det faktisk regner, har meteorologen forudsagt dette i 90 % af tilfældene. Men også i 10% af de dage, hvor det ikke regner, har han forudsagt regn. Hvad er sandsynligheden for at det vil regne på bryllupsdagen?