ECON2130 - Statistikk 1
Forelesning 3: Sannsynlighet

Jo Thori Lind
j.t.lind@econ.uio.no

Oversikt

1. Hva er sannsynlighet?
2. Grunnleggende regler for sannsynlighetsregning
3. Tilfeldighet i datamaskinen
4. Trekning fra et univers
5. Permutasjoner og kombinasjoner
6. Betinget sannsynlighet

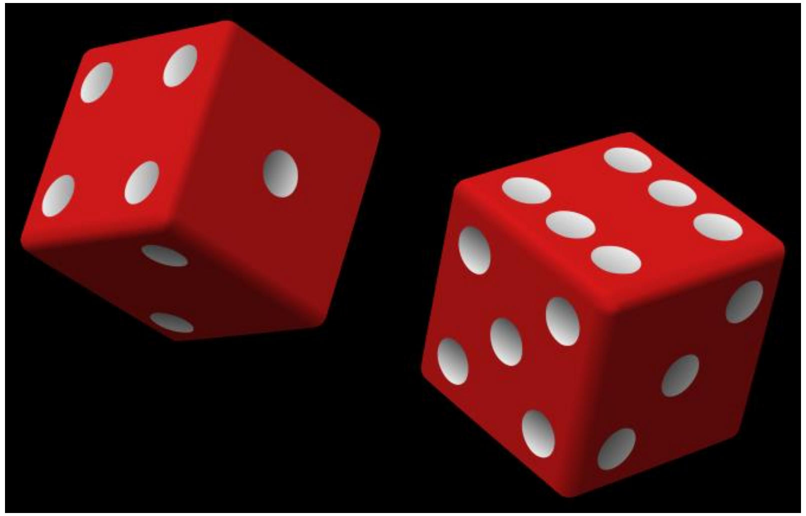
1. Hva er sannsynlighet

Hva er sannsynligheten for at kron-siden havner opp hvis vi kaster kron og mynt?

- Hvis mynten ikke er fikset og den ikke havner på siden er begge utfallene like sannsynlige
- Så da er sannsynligheten ½
- Kan også si 50 % men en dårlig vane

Hva er sannsynligheten for at begge viser samme tall hvis vi kaster to terninger?

- Kan like godt kaste en først og så den andre
- Si vi får 5 på den første
  - Hva er sannsynligheten for 5 på den andre også?
  - Det er 1/6
- Det samme gjelder uansett hva den første terningen viste
- Så svaret er 1/6

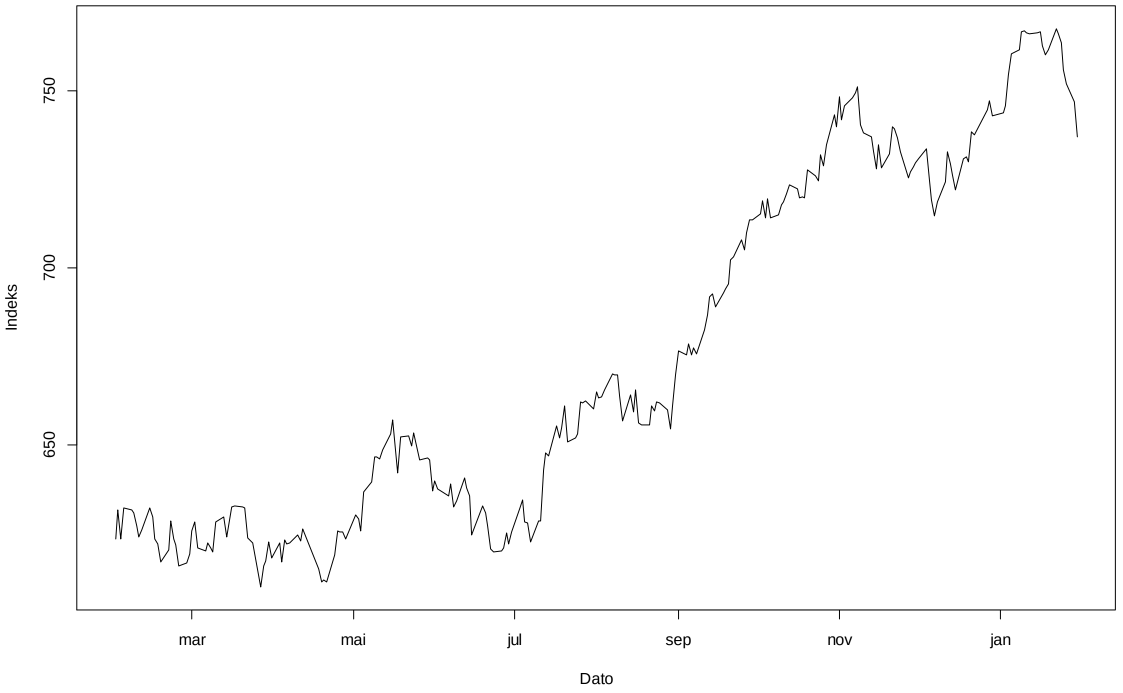


Hva er sannsynligheten for at det snør i morgen?

- Ikke så lett å si for mange av oss
- Men meteorologer kan sette opp modeller som kan si noe om det

Hva er sannsynligheten for at hovedindeksen på Oslo Børs går over 800 i løpet av året?

- Framtidig utvikling i aksjemarkedet viktig for investeringsbeslutninger
- Fornuftig anslag vil avhenge av
  - Hva skjer med økonomien?
  - Utviklingen for store selskap
- Ofte vil ikke alle ha samme anslag
  - Subjektive sannsynligheter



Hva er egentlig tilfeldig?

- Kron og mynt eller terningkast nesten tilfeldig
  - Men vet du nøyaktig hvordan mynten eller terningen kastes kan du forutse hvordan den lander
  - Men avhenger av veldig mange faktorer → sommerfugleffekt
    - Kaos versus tilfeldighet
- Finnes det noe helt tilfeldig?
  - Kvanteeffekter
- Økonomisk utvikling er i stor grad bestemt av mennesker
  - Ikke egentlig tilfeldig
  - Men vanskelig å forutse
    - Ser på det som tilfeldig

2. Grunnleggende regler for sannsynlighetsregning

Noen begrep

- **Eksperiment**
  Det vi gjør som skaper forskjellige utfall
  - Kaste en mynt i været og se hvordan den lander
  - Handle aksjer og se på prisen 2. januar 2019
- **Utfallsrom Ω**
  Alt som kan skje
  - Kron, mynt
  - Indeks på alle positive tall med to desimaler
- **Utfall A**
  En hendelse (eller en gruppe av hendelser)
  - Mynten lander på kron
  - Indeksen er 800.00
  - Indeksen er 800.00 eller høyere

Like sannsynlige utfall

- Anta alle utfallene i utfallsrommet er like sannsynlige
- Da er sannsynligheten for at utfall A skal inntreffe

$$P(A) = \frac{\text{Utfall i } A}{\text{Utfall i } \Omega}$$

Like sannsynlige utfall – kaste en terning

- Hvis vi kaster en terning er $\Omega = \{1,2,3,4,5,6\}$, dvs. 6 elementer
- Hva er sannsynligheten for *A*=å trille en sekser (dvs. et element)

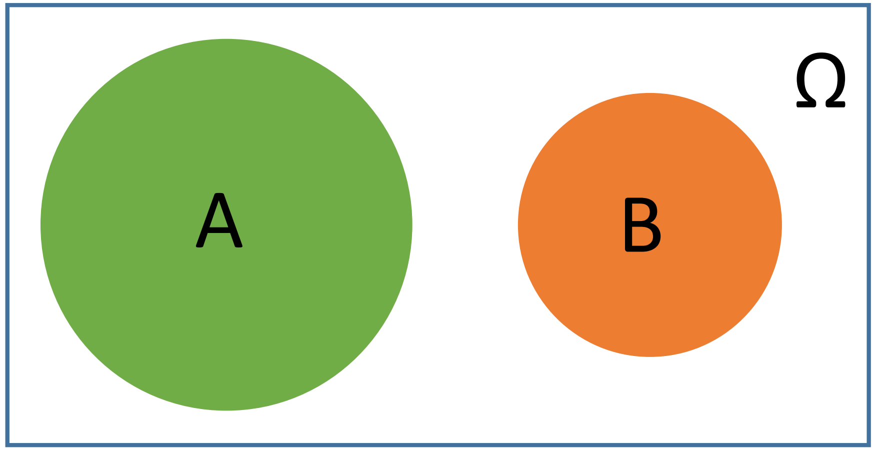
$$P(\text{Trille } 6) = \frac{1}{6}$$

- Hva er sannsynligheten for A=et partall
  - Da er A=trille 2, 4, eller 6 dvs. 3 element

$$P(\text{Partall}) = \frac{3}{6} = \frac{1}{2}$$

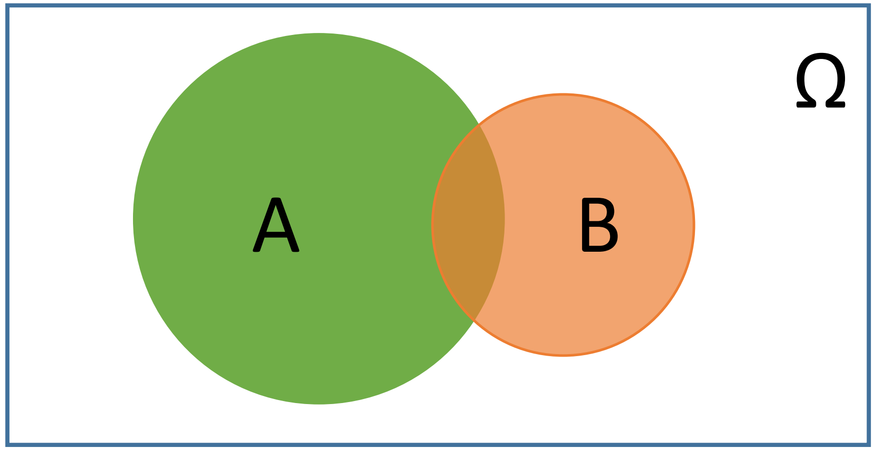
Regler for sannsynlighetsregning

1. Sannsynligheten for ethvert utfall er mellom 0 og 1
$$0 \leq P(A) \leq 1$$
2. Sannsynligheten for et av utfallene i utfallsrommet er 1
$$P(\Omega) = 1$$
3. Hvis *A* og *B* er gjensidig utelukkende (ikke kan inntreffe samtidig) er
$$P(A \text{ eller } B) = P(A) + P(B)$$



Hvis *A* og *B* ikke er gjensidig utelukkende

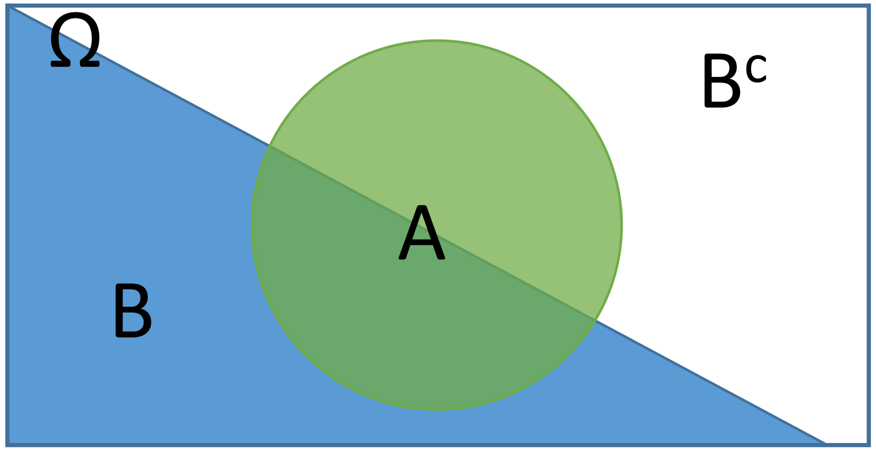
- Da er sannsynligheten
$$P(A \text{ eller } B) = P(A) + P(B) - P(A \text{ og } B)$$
- Hvis ikke får vi dobbelt-telling av snittet av *A* og *B*



Loven om total sannsynlighet

- For et utfall B skriver vi komplementet til B som B^c
- Dette er utfallet «B skjer ikke»
  - Merk at B og B^c er gjensidig utelukkende
- For et hvert utfall A har vi

$$P(A) = P(A \text{ og } B) + P(A \text{ og } B^c)$$



3. Tilfeldighet i datamaskinen

Hvordan regne på tilfeldighet med datamaskin

- Ofte vel så enkelt å få datamaskinen til å regne ut sannsynligheten for et utfall
- Hva er sannsynligheten for å slå 6 på en terning?
    - Slå terning mange ganger, se på andelen seksere
- Kan vi få en datamaskin til å slå terning mange ganger?
- Nesten: Den kan trekke tilfeldige tall

Tilfeldige tall i datamaskin

- Hvor får datamaskinen tilfeldige tall fra?
- Se på hva klokka er
  - Hvor mange hundredeler er det siden 1/1 1960 kl 00:00
  - Helt greit for å få et tilfeldig tall
  - Men skal vi ha mange vil de være veldig like
- Bruke en ekstern kilde
  - Atmosfærisk støy: random.org
  - Halvledere
  - Kvante-effekter

Pseudo-tilfeldige tall

- Ofte ikke så farlig om tallene er virkelig tilfeldige
- Kan lage en rekke med tall som virker tilfeldige
  - En kaotisk prosess
- Enkelt tilnærming: von Neumans midtkvadrat-metode
  1. Velg et tall med fire siffer: 1111
  2. Opphøy i annen: 01234321
  3. Ta ut de midterste fire siffrene
  4. Opphøy i annen: 05489649
  5. Osv.
- I dag bruker vi mer sofistikerte metoder som lager lengre rekker av tall

4. Trekning fra et univers

Hvordan trekke fra et univers

- Si vi har et univers

$$\Omega = \{1,2,3,4,5,6\}$$

- I R lager vi dette universet med `c(1,2,3,4,5,6)` eller `1:6`
- For å trekke et tilfeldig tall fra dette universet bruker vi

```
sample(1:6,1)
```

Hva vi trekker fra

Hvor mange vi trekker

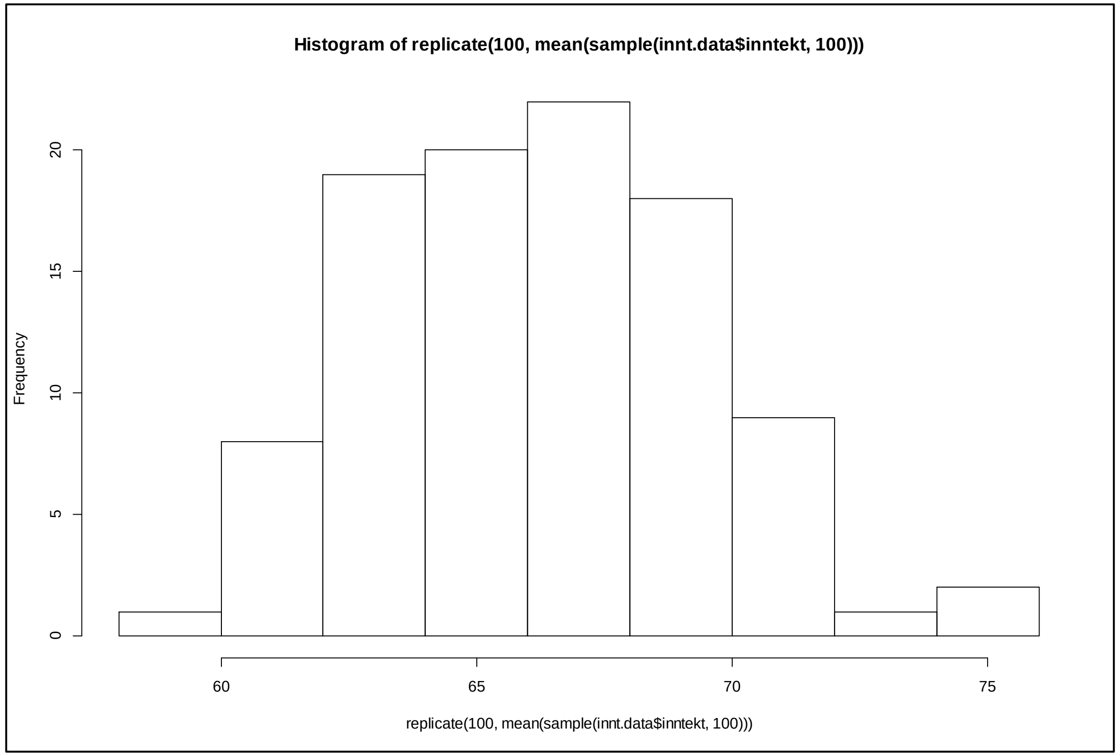
Trekke flere tall

- Ofte vil vi trekke flere tall
  - Kast terning *mange* ganger
- Da må vi si ifra at vi «legger tilbake» så samme tall kan trekkes flere ganger
- Bruker

```
sample(1:6,100,replace = TRUE)
```

Trekke et utvalg

- Si kjenner hele universet
  Ligger i `innt.data`
- Kan trekke et utvalg
  `sample(innt.data$inntekt,100)`
- Finne gjennomsnitt i utvalget
  `mean(sample(innt.data$inntekt,100))`
- Hvordan kan vi gjøre dette mange ganger?
- Bruker
  `replicate(100,mean(sample(innt.data$inntekt,100)))`



5. Permutasjoner og kombinasjoner

Permutasjoner

- Ofte er det en utfordring å telle hvor mange elementer det er i universet
  - Hvor mange mulige utfall er det av en Lotto-trekning?
- Vi har en pose med tre kuler merket A, B og C
  Trekker ut kulene etter tur
- Hva er sannsynligheten for å trekke ABC?
- Må vite hvor mange mulige kombinasjoner vi kan få ut:
  {ABC, ACB, BAC, BCA, CAB, CBA} dvs. 6 kombinasjoner
- Da er sannsynligheten for ABC lik 1/6

Permutasjoner (forts.)

- Må vi skrive opp alle mulighetene?

A			B			C	
B	C		A	C		A	B
C	B		C	A		B	A

- Tre muligheter for første kule
- For hver første kule, to muligheter i andre trekk
- For hver kombinasjon av de to første kulene er det en mulighet i siste
- Totalt antall blir $3 \times 2 \times 1 = 6$

Fakultet

- Funksjonen fakultet skriver vi !
- For et heltall n er $n! = 1 \times 2 \times \cdots \times n$
- Da er $3! = 1 \times 2 \times 3 = 6$
- Hvis vi har kuler med A, B, C, D, E, F, hvor mange rekkefølger kan vi trekke ut?
  - Kan trekke ut $6! = 720$ rekkefølger

Permutasjoner uten å trekke ut alle

- Hvis jeg trekker tre kuler fra posen med A, B, C, D, E, F
  hvor mange rekker kan jeg få?
- I første trekk 6 muligheter
- For hvert første trekk er det fem mulige trekk i andre trekk
- For hvert par i første og andre er det fire muligheter i tredje trekk
- Totalt er det $6 \times 5 \times 4$ muligheter
- Merk at dette kan skrives som $\frac{6!}{3!} = \frac{6\times5\times4\times3\times2\times1}{3\times2\times1}$

Trekke *k* permutasjoner fra *n* mulige

- Generelt, hvis vi skal trekke et antall *k* utvalg fra *n* mulige trekk er antallet mulige kombinasjoner

$$ {}_{n}P_{k} = \frac{n!}{(n-k)!} $$

Permutasjoner med tilbakelegging

- Så langt har vi antatt at når en ball er trukket er den ute
- Hva om vi skriver ned verdien og legger den tilbake?
- Eksempel:
  Hvis et passord består av 5 små bokstaver (ikke æøå), hvor mange passord kan vi lage?
- Det er 26 bokstaver å ta fra
  - Så hver av de 5 bokstavene i passordet er en av de 26 bokstavene a-z
  - I valg av første er det 26
  - For hvert valg i første er det 26 nye valg i andre osv.
  - Totalt $26 \times 26 \times 26 \times 26 \times 26 = 26^5 =$ 11 881 376

Bursdag på samme dag

- Hvis det er 20 personer i et rom, hva er sannsynligheten for at minst to av dem har bursdag på samme dag?
- Tilsvarende, hva er sannsynligheten for at 20 personer alle har forskjellig bursdag?
  - Anta alle fødselsdatoer er like sannsynlige og glem skuddår
- Hvor mange bursdagsdatoer er mulige
  - $365 \times 365 \times \cdots \times 365 = 365^{20}$
- Hvor mange rekkefølger av bursdagsdatoer er mulig slik at alle er ulike
  - $365 \times 364 \times 363 \times \cdots \times 346 = {}_{365}P_{20}$
- Da er sannsynligheten for bare ulike bursdager $\frac{{}_{365}P_{20}}{365^{20}} = \frac{365!}{365^{20} \times 345!} \approx 0.59$

Kombinasjoner

- Med permutasjoner er rekkefølgen viktig
  - AB er ikke det samme som BA
- Men i mange sammenhenger er ikke rekkefølgen relevant
  - Velge noen individer fra en gruppe
  - Trekke Lotto-kuler
- Dette kalles *kombinasjoner*
  - Hvis vi trekker *k* objekter ut fra *n* er det ${}_nC_k = \binom{n}{k}$ muligheter
  - Skal vi trekke to bokstaver fra A, B, C er mulighetene AB, AC og BC så ${}_3C_2 = 3$

Et uttrykk for ${}_nC_k$

- Vi vet at skal vi trekke *k* objekter fra *n* er det ${}_nC_k$ muligheter
- Hver av disse mulighetene kan rangeres på $k!$ Måter
- Derfor har vi

$$ {}_nC_k \times k! = {}_nP_k $$

- Bruker vi uttrykket for ${}_nP_k$ får vi

$$ \boxed{{}_nC_k = \binom{n}{k} = \frac{{}_nP_k}{k!} = \frac{n!}{k!\,(n-k)!}} $$

- I R kan vi bruke `choose(3,2)`

Eksempel på bruk av kombinasjoner

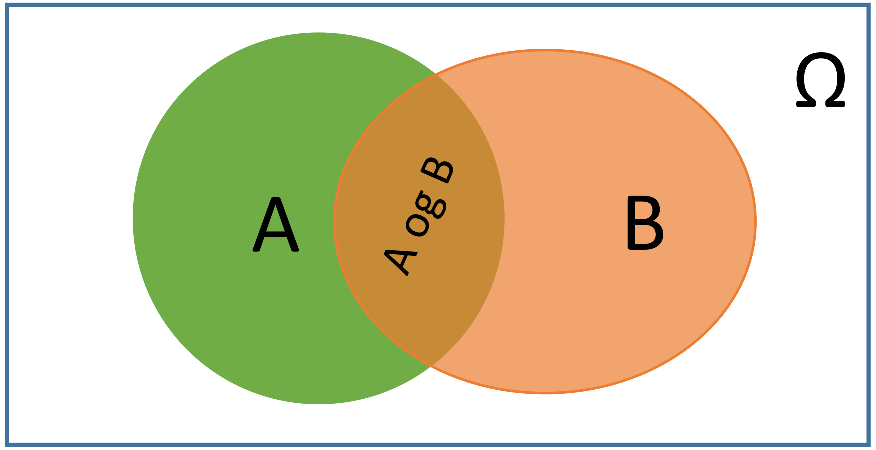
- Instituttet skal ansette 2 nye vit.asser
  - Blant søkerne er det 4 jenter og 6 gutter
  - Hvis søkerne velges tilfeldig, hva er sannsynligheten for at ingen jenter blir ansatt?
- Vi skal trekke 2 personer fra en mengde på 10
  - Det er ${}_{10}C_2 = 45$ muligheter
- Hvor mange muligheter er det for å trekke bare gutter?
  - Hvor mange varianter av 2 gutter kan vi trekke fra 6 mulige
  - Det er ${}_{6}C_2 = 15$ muligheter
- Da er sannsynligheten $P(\text{Bare gutter}) = \frac{15}{45} = 1/3$

6. Betinget sannsynlighet

Betinget sannsynlighet

- Hvis vi vet at B har inntruffet, hvor sannsynlig er det da et også A inntreffer?
- Sannsynligheten for A gitt B er

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



Hva skal vi med betinget sannsynlighet

- Ofte det vi bruker for å sette sammen informasjon
- Hva er sannsynligheten for at hovedindeksen på Oslo børs er over 800 ved utgangen av året
  - Gitt at den på 1. desember er nede i 720
- Hva er sannsynligheten for at en arbeidstaker tjener mer enn 500 000
  - Gitt at hun jobber som barnehageassistent

Uavhengighet

- Vi sier at A og B er uavhengige hvis sannsynligheten for A ikke avhenger av om B har inntruffet
- Det vil si hvis

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

- Hvis A og B er uavhengige har vi

$$P(A\ og\ B) = P(A) \times P(B)$$

  - Eksempel: Sannsynligheten for at to terninger begge viser 6 er

$$\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$