STK1100 våren 2023

Betinget sannsynlighet og uavhengighet

Svarer til avsnittene 2.4 og 2.5

Matematisk institutt

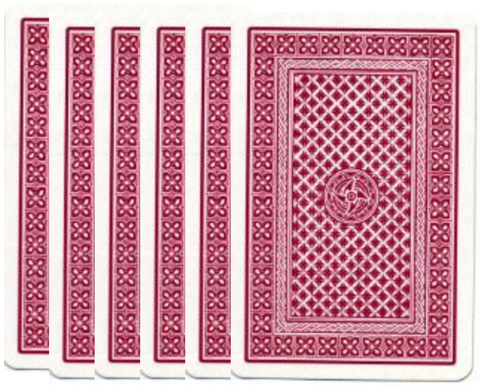
Universitetet i Oslo

Betinget sannsynlighet

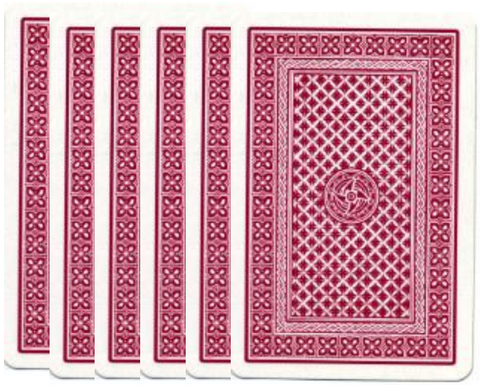
Eksempel 1

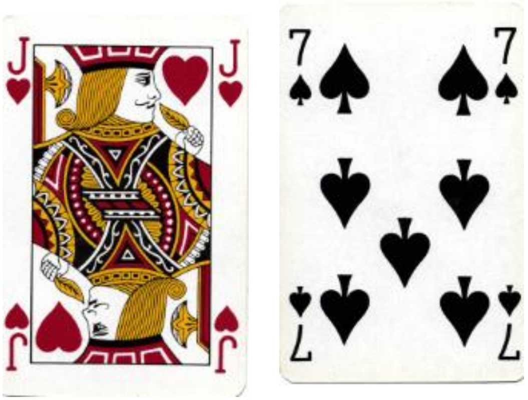
Vi vil først ved hjelp av et eksempel se intuitivt på hva betinget sannsynlighet betyr.

Vi legger **fire røde** kort og **to svarte** kort i en bunke.



Vi trekker tilfeldig ett kort og så ett kort til:





Se på begivenhetene:

A = «første kort er rødt»
B = «andre kort er svart»

Vi har at $P(A) = 4/6 = 2/3$.

Hvis A har inntruffet er sannsynligheten for B lik 2/5.

Dette er den *betingede sannsynligheten* for B gitt A.

Vi skriver $P(B \mid A) = 2/5$.

I eksempel 1 er det *intuitivt* klart hva betinget sannsynlighet er.

Det er ikke alltid like enkelt:

- Hva er den betingede sannsynligheten for at begge kortene er røde gitt at minst ett av dem er rødt?
- Hva er den betingede sannsynligheten for at det første kortet er rødt gitt at det andre er svart?

Vi trenger en definisjon av betinget sannsynlighet!

Vi vil bruke et eksempel til å motivere definisjonen.

Eksempel 2

Norske barn delt inn etter kjønn og fargeblindhet:

	Normal	Fargeblind	Totalt
Gutt	**47.3 %**	**4.1 %**	**51.4 %**
Jente	**48.3 %**	**0.3 %**	**48.6 %**
Totalt	**95.6 %**	**4.4 %**	**100 %**

Vi velger tilfeldig ett barn og ser på begivenhetene:

F = «barnet er fargeblindt»

G = «barnet er en gutt»

Vi har $P(G) = 0.514$ og $P(F \cap G) = 0.041$

Det er «opplagt» at

$$P(F \mid G) = \frac{0.041}{0.514} = \frac{P(F \cap G)}{P(G)}$$

Eksemplet motiverer *definisjonen*:

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

Ved å bytte om «rollene» til *A* og *B*

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

Eksempel 2.26 fra boken – bruk av formel

Av alle som kjøper ny iPhone, kjøper 60% samtidig nytt deksel, 40% kjøper powerbank, og 30% begge deler. For en tilfeldig kunde, la *A*=«kjøper deksel» og *B*=«kjøper powerbank»

Da har vi $P(A) = 0.60$, $P(B) = 0.40$ og $P(A \cap B) = 0.30$

For en tilfeldig valgt kunde, hvis kunden kjøper powerbank, er sannsynligheten for at vedkommende også kjøper deksel

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.30}{0.40} = 0.75$$

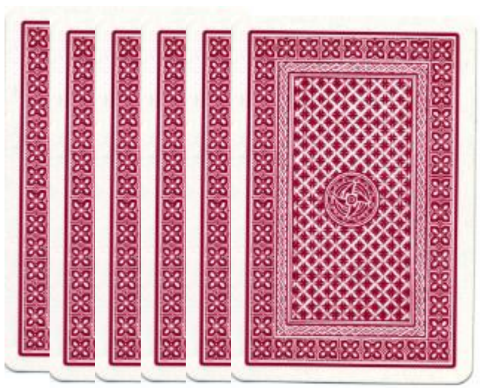
Tilsvarende, hvis kunden kjøper deksel, er sannsynligheten for at vedkommende også kjøper powerbank

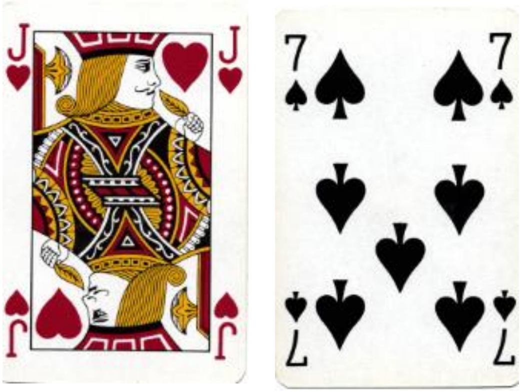
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.30}{0.60} = 0.50$$

Eksempel 3

Vi ser igjen på korteksemplet.

Vi trekker tilfeldig ett kort og så ett kort til.





Se på begivenhetene:

A = «første kort er rødt»

B = «andre kort er svart»

Vi vil bestemme $P(B \mid A)$ ut fra definisjonen.
Det gir en sjekk på at definisjonen er rimelig.

Vi kan trekke to kort på $6 \cdot 5 = 30$ måter.

Vi kan trekke først et rødt og så et svart kort på $4 \cdot 2 = 8$ måter.

Vi kan trekke det første kort rødt på $4 \cdot 5 = 20$ måter.

Det gir

$$P(A \cap B) = \frac{8}{30} \quad \text{og} \quad P(A) = \frac{20}{30}$$

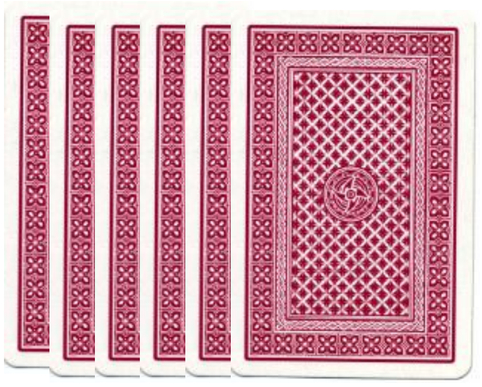
Dermed er

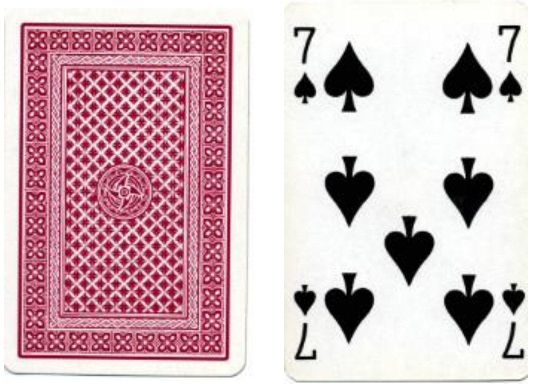
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{8/30}{20/30} = \frac{8}{20} = \frac{2}{5}$$

(selvfølgelig!)

Eksempel 4

Hva er den betingede sannsynligheten for at det første kortet er rødt, gitt at det andre er svart? (Jf. det andre spørsmålet ovenfor.)





Se på begivenhetene:

A = «første kort er rødt»
B = «andre kort er svart»

Vi vil bestemme $P(A\,|\,B)$.

Vi har funnet at $P(A \cap B) = \dfrac{8}{30}$.

Vi kan få et svart kort andre gang på to måter:

- først rødt, så svart kort, dvs $A \cap B$
- to svarte kort, dvs $A' \cap B$

Derfor kan vi skrive *B* som en disjunkt union:

$$B = (A \cap B) \cup (A' \cap B)$$

Det gir at

$$P(B) = P(A \cap B) + P(A' \cap B) \;=\; \frac{4 \cdot 2}{6 \cdot 5} + \frac{2 \cdot 1}{6 \cdot 5} \;=\; \frac{10}{30} = \frac{1}{3}$$

Altså har vi at

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \;=\; \frac{8/30}{10/30} \;=\; \frac{8}{10} = \frac{4}{5}$$

Hva betyr det egentlig at den betingede sannsynligheten er 4/5 = 80% for at det første kortet er rødt gitt at det andre er svart?

Husk at *sannsynlighet er relativ frekvens «i det lange løp».*

Vi tenker oss at vi trekker to kort mange ganger.

- At $P(A) = 2/3$ betyr at det første kortet vil være rødt ca 2/3 av gangene.

- At $P(A \mid B) = 4/5$ betyr at *hvis vi bare teller med de gangene der det andre kortet er svart,* så vil det første kortet være rødt ca 4/5 *av disse gangene.*

Produktsetningen

Definisjon av betinget sannsynlighet:

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

Denne gir *produktsetningen:*

$$\boxed{P(A \cap B) = P(A \mid B) \cdot P(B)}$$

Tilsvarende:

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

Eksempel 5

Ovenfor fant vi $P(A \cap B)$ i korteksemplet som antall gunstige utfall delt på antall mulige utfall (når vi trekker to kort).

Vi kan også finne denne sannsynligheten ved produktsetningen.

Vi har $P(A) = \dfrac{4}{6}$ og $P(B \mid A) = \dfrac{2}{5}$

Dermed gir produktsetningen:

$$P(A \cap B) = P(A) \cdot P(B \mid A) = \frac{4}{6} \cdot \frac{2}{5} = \frac{8}{30}$$

Produktsetningen for tre begivenheter:

$$P(A_1 \cap A_2 \cap A_3)$$

$$= P(A_1 \cap A_2) \cdot P(A_3 \mid A_1 \cap A_2)$$

$$= P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2)$$

Produktsetningen gjelder på tilsvarende måte for fire og flere begivenheter.

Eksempel 6

Etter offentlig statistikk er sannsynligheten

- 96% for at 65 år gammel kvinne skal bli minst 70 år
- 93% for at 70 år gammel kvinne skal bli minst 75 år
- 88% for at 75 år gammel kvinne skal bli minst 80 år

Hva er sannsynligheten for at 65 år gammel kvinne skal bli minst 80 år?

Vi tar for oss 65 år gammel kvinne og ser på begivenhetene

A_1 = «kvinnen blir minst 70 år»

A_2 = «kvinnen blir minst 75 år»

A_3 = «kvinnen blir minst 80 år»

Opplysningene gir:

$$P(A_1) = 0.96$$

$$P(A_2 \mid A_1) = 0.93$$

$$P(A_3 \mid A_1 \cap A_2) = 0.88$$

Hvis kvinnen blir minst 80 år, blir hun også minst 70 år og minst 75 år

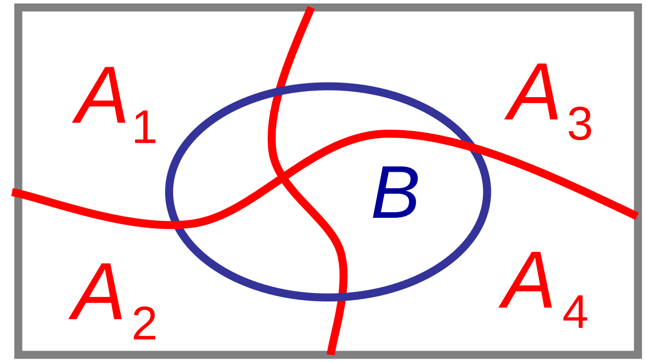
Derfor er $A_1 \cap A_2 \cap A_3 = A_3$

$$P(\text{minst 80 år}) = P(A_3) = P(A_1 \cap A_2 \cap A_3)$$

$$= P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2)$$

$$= 0.96 \cdot 0.93 \cdot 0.88 = 0.79$$

Total sannsynlighet



Anta at $A_1, A_2, ..., A_k$
er disjunkte og at

$$A_1 \cup A_2 \cup .... \cup A_k = S$$

Vi kan da skrive en begivenhet B som en union av disjunkte begivenheter:

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup .... \cup (A_k \cap B)$$

Dette og produktsetningen gir setningen om *total sannsynlighet*

$$\boxed{P(B) = \sum_{i=1}^{k} P(A_i \cap B) = \sum_{i=1}^{k} P(B \mid A_i) P(A_i)}$$

Eksempel 7

Vi legger

- **fire røde** kort og **to svarte** kort i bunke I
- **to røde** kort og **fire svarte** kort i bunke II

Vi velger tilfeldig en bunke og trekker to kort fra denne.

Hva er sannsynligheten for at vi får to røde kort?

A_1 = «vi trekker fra bunke I»

A_2 = «vi trekker fra bunke II»

B = «vi trekker to røde kort»

Opplysningene gir:

$$P(A_1) = P(A_2) = \frac{1}{2}$$

$$P(B \mid A_1) = \frac{4 \cdot 3}{6 \cdot 5} = \frac{6}{15} \qquad P(B \mid A_2) = \frac{2 \cdot 1}{6 \cdot 5} = \frac{1}{15}$$

Setningen om total sannsynlighet gir

$$P(B) = \frac{6}{15} \cdot \frac{1}{2} + \frac{1}{15} \cdot \frac{1}{2} = \frac{7}{30} = 0.233$$

Bayes' setning

Anta at $A_1, A_2, \ldots, A_k$ er disjunkte og at

$A_1 \cup A_2 \cup \ldots \cup A_k = S$

Definisjonen av betinget sannsynlighet gir at

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

Vi bruker produktsetningen for telleren og total sannsynlighet for nevneren og får *Bayes' setning:*

$$\boxed{P(A_j \mid B) = \frac{P(B \mid A_j) \cdot P(A_j)}{\sum_{i=1}^{k} P(B \mid A_i) P(A_i)}}$$

Eksempel 8

Vi ser på eksempel 7

Hvis begge kortene er røde, hva er sannsynligheten for at vi trakk fra bunke I ?

A_1 = «vi trekker fra bunke I»
A_2 = «vi trekker fra bunke II»
B = «vi trekker to røde kort»

$$P(A_1) = P(A_2) = \frac{1}{2} \qquad P(B \mid A_1) = \frac{6}{15} \qquad P(B \mid A_2) = \frac{1}{15}$$

Bayes setning gir:

$$P(A_1 \mid B) = \frac{P(B \mid A_1)P(A_1)}{P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2)} = \frac{\frac{6}{15}\cdot\frac{1}{2}}{\frac{6}{15}\cdot\frac{1}{2} + \frac{1}{15}\cdot\frac{1}{2}} = \frac{6}{7}$$

Eksempel 9

En kvinne tar en mammografiundersøkelse.

Se på begivenhetene:

S = «kvinnen har brystkreft»
M = «undersøkelsen viser tegn på kreft»

Fra erfaringer med mammografi har vi

$$P(M \mid S) = 0.95 \qquad P(M \mid S') = 0.035$$

Vi antar at $P(S) = 0.007$

Anta at undersøkelsen viser tegn på kreft.

Hva er da sannsynligheten for at kvinnen virkelig har kreft?

Bayes setning gir:

$$P(S \mid M) = \frac{P(M \mid S) \cdot P(S)}{P(M \mid S) \cdot P(S) + P(M \mid S') \cdot P(S')}$$

$$= \frac{0.95 \cdot 0.007}{0.95 \cdot 0.007 + 0.035 \cdot 0.993}$$

$$= 0.16$$

Selv om undersøkelsen viser tegn på kreft, er det bare 16% sannsynlig at hun virkelig har det

Avhengige og uavhengige begivenheter

Husk definisjonen av betinget sannsynlighet

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

Hvis $P(A \mid B) = P(A)$ er A og B *uavhengige* begivenheter

Hvis $P(A \mid B) \neq P(A)$ er A og B *avhengige* begivenheter

Hvis $P(A \mid B) = P(A)$, så er $P(B \mid A) = P(B)$ (jf. forelesningen)

Eksempel 10: Kast én terning to ganger

A = «minst fem øyne i første kast» $\qquad P(A) = \frac{12}{36}$

B = «sum øyne lik sju» $\qquad P(B) = \frac{6}{36}$

C = «minst én sekser» $\qquad P(C) = \frac{11}{36}$

(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)

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

$$P(A \cap C) = \frac{7}{36}$$

$$P(B \mid A) = \frac{\frac{2}{36}}{\frac{12}{36}} = \frac{2}{12} = P(B)$$

$$P(C \mid A) = \frac{\frac{7}{36}}{\frac{12}{36}} = \frac{7}{12} \neq P(C)$$

Hvis $P(A \mid B) = P(A)$, dvs A og B er uavhengige begivenheter, så er (jf. forelesningen)

- $P(A' \mid B) = P(A')$, dvs A' og B er uavhengige
- $P(A \mid B') = P(A)$, dvs A og B' er uavhengige
- $P(A' \mid B') = P(A')$, dvs A' og B' er uavhengige

Viktig resultat (bevis på forelesningen):

A og B er uavhengige hvis og bare hvis

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

Resultatet gjelder også når A og/eller B har sannsynlighet null

Eksempel 11: Kast kronestykke og femkrone

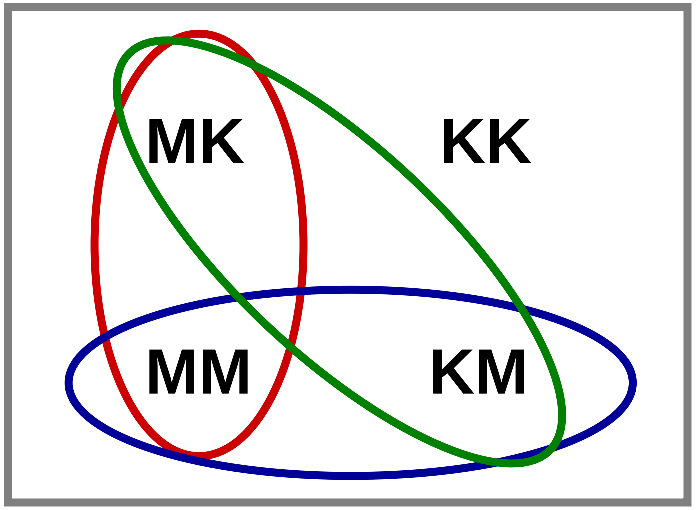
La for eksempel KM bety krone på kronstykket og mynt på femkronen

Utfallsrom U = {MM, MK, KM, KK}

A = {MM, MK}

B = {MM, KM}

C = {MK, KM}



$$P(A)\cdot P(B)=\frac{2}{4}\cdot\frac{2}{4}=\frac{1}{4}=P(A\cap B)$$

A og B er uavhengige

A og C er uavhengige

B og C er uavhengige

Men $P(A\cap B\cap C)=0$

Uavhengighet for flere begivenheter

Begivenhetene $A_1,\ A_2\ ,....,A_n$ er uavhengige hvis vi for enhver *k* og enhver delmengde av indekser $i_1,\ i_2\ ,....,i_k$ har at

$$P(A_{i_1} \cap A_{i_2} \cap .... \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot ... \cdot P(A_{i_k})$$

Ovenfor har vi sett på situasjoner der vi kan vise at begivenheter er uavhengige

Men vanligvis er uavhengighet en forutsetning for sannsynlighetsmodellen vår

Eksempel 12

Kast fem terninger. Hva er sannsynligheten for at vi får minst én sekser?

Vi finner først

$$P(\text{ingen seksere}) = \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \left(\frac{5}{6}\right)^5 = 0.402$$

Dermed er

$$P(\text{minst én sekser}) = 1 - P(\text{ingen seksere})$$

$$= 1 - 0.402 = 0.598$$

Eksempel 13

Kast fem terninger. Hva er sannsynligheten for at alle terningene viser samme antall øyne (yatzy).

Vi har at

$$
\begin{aligned}
P(\text{yatzy}) &= P(\text{fem enere}) + P(\text{fem toere}) + P(\text{fem treere}) \\
&\quad + P(\text{fem firere}) + P(\text{fem femere}) + P(\text{fem seksere}) \\
&= \left(\frac{1}{6}\right)^5 + \left(\frac{1}{6}\right)^5 + \left(\frac{1}{6}\right)^5 + \left(\frac{1}{6}\right)^5 + \left(\frac{1}{6}\right)^5 + \left(\frac{1}{6}\right)^5 \\
&= 6 \cdot \left(\frac{1}{6}\right)^5 = 0.00077
\end{aligned}
$$

Eksempel 14
Chevalier de Mérés problem

Antoine Gombaud
«Chevalier de Méré»
(1607-84)

Hva er mest sannsynlig:

- å få minst én sekser i fire kast med én terning
- å få minst en «dobbelsekser» i 24 kast med to terninger

$$P(\text{minst én sekser i fire kast})$$

$$P(\text{minst én dobbelsekser i 24 kast})$$

Eksempel 15

Kast en terning til vi får sekser og registrer hvor mange kast vi må gjøre

Sett $P(k) = P(\text{første sekser i } k\text{-te kast})$

For å få første sekser i *k*-te kast må vi ikke få sekser i de *k* – 1 første kastene og så få sekser

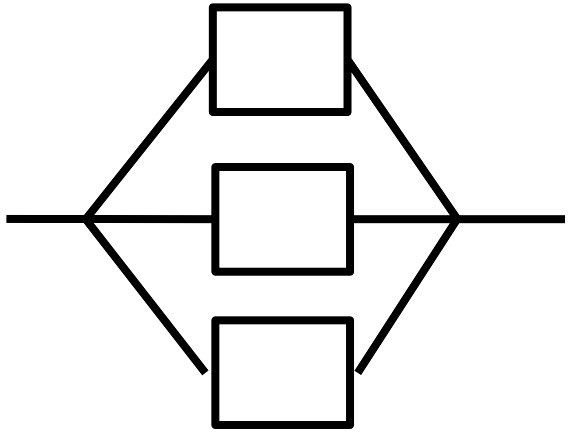
Siden resultatene av kastene er uavhengige får vi

$$P(k) = \underbrace{\frac{5}{6} \cdot \frac{5}{6} \cdot \ldots \frac{5}{6}}_{k-1 \text{ ganger}} \cdot \frac{1}{6} \quad = \left(\frac{5}{6}\right)^{k-1} \cdot \frac{1}{6}$$

Eksempel 16: Serie- og parallellsystemer

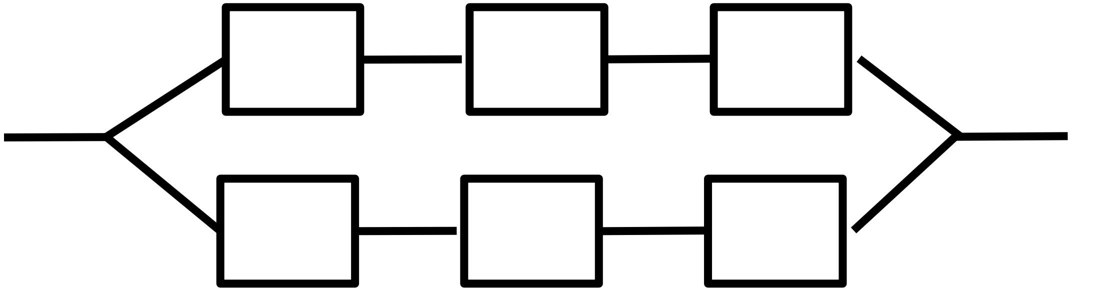
Et seriesystem fungerer hvis alle komponentene fungerer

Et paralellsystem fungerer hvis minst én av komponentene fungerer



Eksempel 2.37 i boken - sammensatte systemer

Serie-parallell system



Krysskoblet system