Kapittel 2

Sannsynlighetsregning

2.1 Innledning

Sannsynlighetsbegrepet er et svært sentralt begrep innen statistikk og sannsynlighetsregning. Det er også et vanskelig begrep. Prøv selv, før du leser videre, å definere hva *du* mener med sannsynlighet.

Vi skal begrense oss til tilfeller der vi på en objektiv måte kan *tallfeste* sannsynligheten for at en hendelse skal inntreffe. En slik hendelse er for eksempel «å få kron i et tilfeldig myntkast med rettferdig kronestykke». De fleste vil være enige i at sannsynligheten for at dette skal skje er 50 %. Ved litt ettertanke vil de fleste også være enige i at det vi mener med 50 % i dette tilfellet er at i en lang rekke med myntkast vil ca. halvparten gi kron som utfall. Dette er tankegangen bak det **frekventistiske** sannsynlighetsbegrep, som vi skal begrense oss til. Ordet frekventistisk henspeiler her på at man tenker seg et forsøk gjentatt en rekke ganger under identiske betingelser, og at man beregner den relative frekvensen for en gitt hendelse.

Det engelske ordet probability betyr sannsynlighet, og vi bruker bokstaven P for å betegne sannsynlighet. $P = 0$ angir at hendelsen ikke kan inntreffe, for eksempel at Ola vil være i Norge og Malaysia samtidig. $P = 1$ (100 %) angir hendelser som *helt sikkert* inntreffer, for eksempel at dersom Kari går så beveger bena hennes seg.

Legg merke til at vi som regel må finne ganske sære eksempler på hendelser som enten ikke kan inntreffe ($P = 0$) eller er helt sikre ($P = 1$). For alle andre hendelser ligger P mellom 0 og 1. Dersom vi ønsker å angi sannsynligheter i %, multipliserer vi P med 100 %.

Mange studenter, kanskje de fleste, vil synes at sannsynlighetsregningen er noe av det vanskeligste i et kurs i grunnleggende statistikk og sannsynlighetsregning. Det er måten å tenke på som er spesiell. Trenger du gjennom denne «muren», åpner det seg en verden av artige problemstillinger å gyve løs på. For de som er glad i spill, er det nok å nevne eksempler som sannsynligheten for ulike gevinster i Lotto eller ulike korthender i bridge og andre kortspill. Du vil også bli i stand til å takle ganske så sammensatte problemstillinger.

De ulike avsnittene i dette kapitlet hører nøye sammen. Vær derfor grundig med å få med deg en mest mulig presis forståelse av begrepene undervegs. Og som ellers i matematisk statistikk: Nøl ikke før du gir deg i kast med oppgaveregning.

2.2 Utfallsrom, enkeltutfall og hendelse

Et (statistisk) *eksperiment* er prosessen med å samle data for et fenomen der utfallet varierer (statistisk variasjon). I tilknytning til eksperimenter er de viktige begrepene utfallsrom, enkeltutfall og hendelser definert som følger:

Utfallsrom, enkeltutfall og hendelse (definisjon)
Utfallsrommet er samlingen av alle mulige forskjellige (distinkte) utfall av et eksperiment. Hvert av disse utfallene skal vi kalle *enkeltutfall*. Vi benytter symbolet S for utfallsrommet, og symbolene $e_1,\ldots,e_n$ for enkeltutfallene i S.

En *hendelse* er en samling av enkeltutfall og utgjør et delområde av utfallsrommet S. Det er vanlig å benytte de første store bokstavene i alfabetet: A, B, C,… for å betegne hendelser. Vi sier at en hendelse A inntreffer når ett eller flere av enkeltutfallene i A inntreffer.

La oss belyse de begrepene vi har introdusert med noen eksempler. Vi betrakter følgende 4 eksperiment:

Eksperiment (a): Noter *kjønnet* til de 2 første spebarn som fødes imorgen.

Eksperiment (b): La hver av 10 personer smake et glass Cola og et glass Cola Light, og noter *hvor mange* som foretrekker Cola Light.

Eksperiment (c): Gi et antibiotikum til pasienter som lider av en virusinfeksjon inntil en pasient får en *antireaksjon*.

Eksperiment (d): Mål *luftfuktigheten i prosent* i ei badstu.

Vi får følgende utfallsrom:

S_a = { *GG, GJ, JG, JJ* }, *G* = Gutt, *J* = Jente, *GJ* betyr *først* gutt, *så* jente, osv. En slik kronologisk rekkefølge fra venstre mot høyre er ganske vanlig.

S_b = { 0, 1, 2,…,10 }

S_c = { *A, NA, NNA, NNNA*,… }, *A* = Antireaksjon, *N* = Ingen reaksjon

S_d = { x: 0 % $\leq x \leq$ 100 % }, som leses: «Mengden av alle reelle verdier av *x* fra og med 0 % til og med 100 %»

Legg merke til at kun ett av de angitte utfall kan opptre av gangen. Vi kan imidlertid ofte beskrive en *hendelse* der flere enkeltutfall kommer innenfor definisjonen av hendelsen:

Eks. 2.1 **Eksakt ei jente.** Betrakt eksperiment (a), og la *A* betegne hendelsen «eksakt ei jente». Vi får da at *A* = { *GJ, JG* }. ☺

Eks. 2.2 **Høyst 5 pasienter**. Betrakt eksperiment (c), og la *C* betegne hendelsen «høyst 5 pasienter får antibiotika før siste pasient får antireaksjon». Vi får da at *C* = { *A, NA, NNA, NNNA, NNNNA* }. ☺

Et utfallsrom er *diskret* når det består av et *endelig* eller *tellbart uendelig* antall enkeltutfall. Utfallsrommene for eksperimentene (a), (b) og (c) er diskrete. (a) og (b) er eksempler med et endelig antall enkeltutfall, mens eksperiment (c) er et eksempel med et tellbart uendelig antall enkeltutfall. Når et utfallsrom inneholder alle tall i et intervall på den reelle tallinja er utfallsrommet *kontinuerlig*. Eksperiment (d) er et eksempel med kontinuerlig utfallsrom.

2.3 Sannsynlighet for en hendelse

Vi skal bruke betegnelsen *P*(*A*) for å betegne sannsynligheten for at en hendelse A skal inntreffe (*P* er forkortelse for det engelske ordet «Probability» som betyr sannsynlighet). Et eksempel: La *A* betegne hendelsen «kronsida opp i tilfeldig kast med en rettferdig mynt». Intuitivt vil de aller fleste da være med på at det er «50 % sannsynlighet for å få kron», eller: Det er «50 % sannsynlighet for at hendelsen A skal inntreffe», $P(A) = 0.5$ (0,5 = 0.5·100 % = 50 %).

Tenker vil litt igjennom hva vi mener med «50 % sannsynlighet» i dette tilfellet, vil de fleste være enige i at det som ligger i underbevisstheten er følgende: Den relative andelen kron i en serie med myntkast vil mer og mer nærme seg 50 % ettersom antall kast øker.

Foreløpig skal vi holde oss til følgende intuitive oppfatning av begrepet *sannsynlighet*: Den andelen A inntreffer i en serie av forsøk under identiske forhold. (Sannsynlighet ≈ antall forsøk med «gunstige» utfall delt på antall forsøk totalt når antall forsøk er svært stor).

Vi skal først betrakte sannsynlighetsbegrepet i forbindelse med hva vi kaller en *uniform sannsynlighetsmodell*.

Uniform sannsynlighetsmodell (definisjon)
Hvis et utfallsrom består av k enkeltutfall $\{e_1,\ldots,e_k\}$, der hvert enkeltutfall inntrer med like stor sannsynlighet, så er sannsynligheten for hvert enkeltutfall lik $1/k$. Hvis en hendelse A består av m av disse k enkeltutfallene, består den uniforme sannsynligtesmodellen i at vi kan sette

(2.1) $P(A) = m/k$

Når et eksperiment innrettes slik at sannsynligheten for hvert enkeltutfall er like stor, kan vi anvende en *uniform sannsynlighetsmodell.*

I vårt myntkasteksempel har vi at $k = 2$ (2 like sannsynlige enkeltutfall: Kron og mynt), og $m = 1$ (kun ett enkeltutfall gir kron), dvs. $P(A) = P(«\text{kron}») = m/k = 1/2$.

Eks. 2.3 **Terningkast med rettferdig terning**.
Eksperiment: Antall øyne i ett tilfeldig kast.

Utfallsrom $S = \{ e_1, e_2, e_3, e_4, e_5, e_6 \}$

Enkeltutfall:

e_1: 1 øye e_2: 2 øyne e_3: 3 øyne e_4: 4 øyne e_5: 5 øyne e_6: 6 øyne

$P(e_1) = P(e_2) = \cdots = P(e_6) = 1/6$

Siden alle enkeltutfallene er like sannsynlige, har vi her et eksempel der en uniform sannsynlighetsmodell er riktig. Legg igjen merke til at kun ett av enkeltutfallene i utfallsrommet, S, kan opptre av gangen. Et terningkast med én terning kan for eksempel ikke gi både 2 og 5 i ett og samme kast. ☺

Eks. 2.4 ***A* = «antall øyne er et like tall»**,
dvs. $A = \{ e_2, e_4, e_6 \} \;\Rightarrow\; P(A) = P(e_2) + P(e_4) + P(e_6) = 3/6 = 1/2$

Siden vi her kan legge en uniform sannsynlighetsmodell til grunn, hadde det vært tilstrekkelig å telle opp det antall enkeltutfall som A består av (her: $m = 3$), og dele på totalt antall enkeltutfall i utfallsrommet (her: $k = 6$): $P(A) = m/k = 3/6 = 1/2$. ☺

Vi går nå løs på en mer generell definisjon av sannsynlighetsbegrepet knyttet til begrepet relativ frekvens, og vi skal sette opp betingelser for en sannsynlighets-modell basert på diskrete utfallsrom.

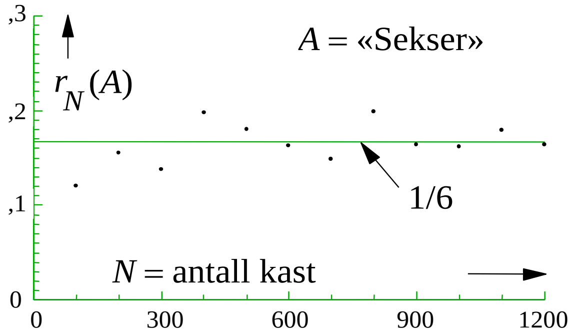
Relativ frekvens, $r_N(A)$ (definisjon)

Relativ frekvens av en hendelse A i N forsøk («trials»), $r_N(A)$, er definert som følger:

(2.2) $$r_N(A) = \frac{\text{Antall ganger } A \text{ inntreffer i løpet av } N \text{ forsøk}}{\text{Antall forsøk } N}$$

Her er det underforstått at forsøkene (eks: terningkast) er foretatt under identiske betingelser.

Tenk igjen på terningkasteksemplet. Si for eksempel at $N = 6$ og A = «sekser». Vi vil da kanskje få én sekser i løpet av de 6 forsøkene, dvs. $r_N = 1/6$. Det er imidlertid ikke usannsynlig at vi får 2 ($r_N = 2/6$) eller ingen ($r_N = 0$) seksere på 6 forsøk. Øker vi antall forsøk ($N = 12, 24, 100, 1000, \ldots$), forventer vi imidlertid at r_N (relativ andel seksere) vil stabilisere seg rundt 1/6. En slik betraktning legger vi til grunn for definisjonen av sannsynlighet $P(A)$ for en hendelse A (se neste ramme).



I figuren til venstre er vist et eksempel der relativ frekvens av en hendelse A stabiliserer seg når antall (N) forsøk øker. De svarte punktene er tilfeldige forsøk. Legg merke til at stabiliseringen går ganske sakte: Selv med over 1000 forsøk, er det en «synlig» usikkerhet rundt den teoretiske verdien 1/6.

Sannsynlighet, $P(A)$ (definisjon)

Vi bruker betegnelsen $P(A)$ for å betegne sannsynligheten for en hendelse A, og vi skal benytte følgende definisjon:

$$(2.3) \qquad P(A) = \lim_{N\to\infty} r_N(A)$$

dvs. $P(A)$ er lik den grenseverdien som relativ frekvens $r_N(A)$ går mot når antall eksperimenter N går mot uendelig. Det er her forutsatt at grenseverdien konvergerer mot en fast verdi.

Legg merke til den nære sammenhengen mellom begrepene *relativ frekvens* og *sannsynlighet*. Denne nære sammenhengen er svært sentral i statistikken, og vi skal kun se på det frekventistiske sannsynlighetsbegrep i denne boka.

I neste ramme har vi satt opp *betingelser* for den *sannsynlighetsmodell* vi skal anvende, basert på *diskrete* utfallsrom. Merk at vi definerer sannsynlighetene til å ligge mellom 0 og 1, hvilket tilsvarer 0 til 100 %.

Sannsynlighetsmodell, diskrete utfallsrom

Sannsynlighet er en funksjon, definert for hendelser, som tilfredsstiller følgende betingelser:

i) For alle hendelser, A, gjelder at $0 \le P(A) \le 1$
P(A) = 0 betyr at hendelsen A helt sikkert *ikke* kan inntreffe.
P(A) = 1 betyr at hendelsen A *helt sikkert* vil inntreffe.

ii) $P(A)$ er summen av sannsynlighetene for at hvert av enkeltutfallene som tilhører A skal inntreffe:

$$P(A) = \sum_{e_i \in A} P(e_i) \tag{2.4}$$

der summen skal tas over alle enkeltutfall e_i som er med i A.

iii) $P(S)$ er summen av sannsynlighetene til alle enkeltutfallene i hele utfallsrommet, S, og denne må være lik 1:

$$P(S) = \sum_{e_i \in S} P(e_i) = 1 \tag{2.5}$$

der summen skal tas over alle enkeltutfall e_i som er med i S.

2.4 Union, snitt og komplement

I dette avsnittet skal vi gjøre en del bruk av Venn-diagram. Vi kommer ikke til å gi noen formell definisjon av Venn-diagram, men det vil gjennom eksempler og oppgaver klart gå fram hva det dreier seg om. Enkelt sagt er et Venn-diagram et lukket område i planet som omslutter hele utfallsrommet, S, og der hendelser tegnes inn som lukkede delområder av utfallsrommet. Enkeltutfall markeres gjerne med punkter i diagrammet.

Eks. 2.5 **Surströmming 1**. En tilfeldig valgt student blir tvangsforet med surströmming. Studenten responderer på 2 måter. Enten liker han/hun surströmming (J), eller så liker han/hun ikke surströmming (N).

Eksperiment: Surströmmingen gis til 3 tilfeldig utvalgte studenter, og responsen til hver student registreres.

Oppgave

Lag et Venn-diagram for utfallsrommet, S, og angi følgende hendelser:
A: Bare én student liker surströmming
B: Første student liker surströmming
C: Både første og andre student liker ikke surströmming

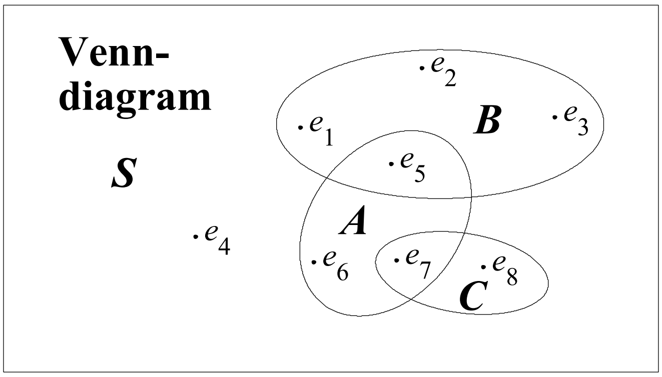
Løsningsforslag

La oss først finne alle enkeltutfallene i utfallsrommet for eksperimentet. For den første studenten er det 2 mulige resultater (J eller N), og hvert av disse resultatene blir etterfulgt av 1 av 2 muligheter (J eller N) for student nr.2. Hver av disse $2 \cdot 2 = 4$ kombinerte resultatene kan bli etterfulgt av en J eller en N for den tredje studenten. Totalt får vi derfor $n = 2 \cdot 2 \cdot 2 = 8$ mulige enkeltutfall, $e_1, \ldots, e_8$:

JJJ (e_1)	JJN (e_2)	JNJ (e_3)	NJJ (e_4)
JNN (e_5)	NJN (e_6)	NNJ (e_7)	NNN (e_8)

Hendelsene A, B og C består av enkeltutfallene:

$$A = \{ e_5, e_6, e_7 \}$$
$$B = \{ e_1, e_2, e_3, e_5 \}$$
$$C = \{ e_7, e_8 \}$$



Venn-diagrammet er vist i figuren til venstre, der rektangelet representerer utfallsrommet, S.

De ulike hendelsene omslutter akkurat de enkeltutfallene som er med i hendelsen. ☹

Vi skal nå definere 3 viktige operasjoner med hendelser, nemlig *union, snitt* og *komplement*:

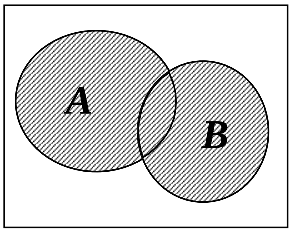
Union, snitt og komplement

Unionen av 2 hendelser A og B betegnes $A \cup B$ og er mengden av alle enkeltutfall som er med i A, *eller* i B, *eller* i både A og B.

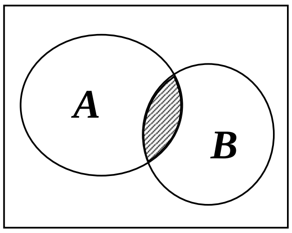
Snittet mellom 2 hendelser A og B betegnes $A \cap B$, eller kortere, AB, og er mengden av alle enkeltutfall som er med i *både* A og B.

Komplementet til en hendelse A betegnes med A^C og er mengden av alle enkeltutfall som *ikke* er med i A.

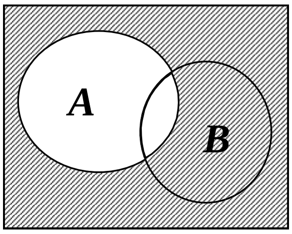
De definerte begreper er illustrert med skraverte områder nedenfor.



Union $A \cup B$



Snitt AB



Komplement A^C

Eks. 2.6 **Surströmming 2**. La hendelsene A, B og C være definert som i eks. 2.5. Angi hvilke enkeltutfall følgende hendelser består av:

$$A \cup B,\ AC,\ BC,\ B^C,\ (A \cup B)^C,\ \left(A^C\right) \cup \left(B^C\right)$$

Løsningsforslag

Fra før: $A = \{ e_5, e_6, e_7 \}$, $B = \{ e_1, e_2, e_3, e_5 \}$, $C = \{ e_7, e_8 \}$

Ved å bruke definisjonene for operasjonene union, snitt og komplement får vi:

$A \cup B = \{ e_1, e_2, e_3, e_5, e_6, e_7 \}$
$AC = \{ e_7 \}$
$BC = \emptyset$ (Tom mengde)
$B^C = \{ e_4, e_6, e_7, e_8 \}$
$(A \cup B)^C = \{ e_4, e_8 \}$

B^C er gitt ovenfor, og $A^C = \{ e_1, e_2, e_3, e_4, e_8 \}$

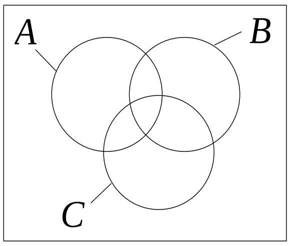
$\Rightarrow A^C \cup B^C = \{ e_1, e_2, e_3, e_4, e_6, e_7, e_8 \}$ ☹

Vi har ovenfor betraktet snitt og union av 2 hendelser. Reglene kan lett generaliseres til flere enn 2 hendelser, f.eks. $A \cup B \cup C$, $A \cap B \cap C$ og $(A \cup B) \cap C$:

(2.6) $$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C)$$

(2.7) $$A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C)$$

(2.8) $$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$



Figuren til venstre er til god hjelp for å overbevise seg om at lign.(2.6)-(2.8) er korrekte!

Et viktig spesialtilfelle når vi betrakter 2 hendelser A og B får vi når *ikke noe* enkeltutfall opptrer samtidig i de 2 mengdene. A og B sies da å være **gjensidig ekskluderende** hendelser, eller 2 hendelser som gjensidig utelukker hverandre. I dette tilfellet er $AB = \emptyset$, der $\emptyset$ er betegnelsen for en tom mengde.

Vi går nå over til å se på 2 viktige *sannsynlighetslover* for hendelser:

Sannsynlighetslover for hendelser

(2.9) $P(A^C) = 1 - P(A)$ (komplementærloven)

(2.10) $P(A \cup B) = P(A) + P(B) - P(AB)$ (addisjonsloven)

Som et spesialtilfelle av siste lov ovenfor får vi *addisjonsloven for unionen av gjensidig ekskluderende hendelser*:

(2.10b) $P(A \cup B) = P(A) + P(B)$

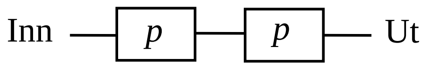
når A og B er gjensidig ekskluderende hendelser ($AB = \emptyset$).

Fra (2.10) kan vi utlede addisjonssetningen for 3 hendelser:

(2.10c) $$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

Eks. 2.7 Seriesystem

Figuren nedenfor viser et seriesystem med to like og uavhengige komponenter der p = sannsynligheten for at en komponent svikter.



Oppgave

a) Hva er sannsynligheten for at systemet funksjonerer?

b) I en kjettinglenke med n løkker er sannsynligheten pr. år for at en løkke ryker konstant lik p, uavhengig av hvilken løkke vi betrakter. Hva er sannsynligheten for at kjettingen holder ett år?

Løsningsforslag

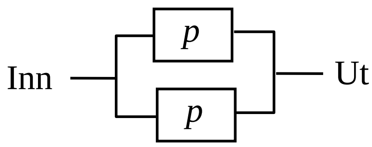
a) La A betegne hendelsen at komponent 1 er OK og B at komponent 2 er OK. Vi får da:

$$P(\text{system OK}) = P(AB) = P(A)\cdot P(B) = (1-p)\cdot(1-p) = (1-p)^2$$

b) Kjettinglenke med n løkker tilsvarer seriesystem med n komponenter. Generaliserer vi løsningen i a) får vi: $P(\text{kjetting OK}) = (1-p)^n$. ☺

Eks. 2.8 Parallellsystem

Figuren nedenfor viser et parallellsystem med to like og uavhengige komponenter der p = sannsynligheten for at en komponent svikter.



Oppgave

a) Hva er sannsynligheten for at systemet funksjonerer?

b) Hva er sannsynligheten for at et parallellsystem med n komponenter i parallell funksjonerer?

Løsningsforslag

a) La A betegne hendelsen at komponent 1 er OK og B at komponent 2 er OK. Vi får da:

$$P(\text{system OK}) = 1-P(\text{begge komponenter svikter}) = 1-P(A^C B^C) =$$
$$1-P(A^C)\cdot P(B^C) = 1-p^2$$

b) Med n komponenter i parallell får vi samme fremgangsmåte som i a), men må erstatte 2 med n: $P(\text{system OK}) = 1-p^n$ ☺

Eks. 2.9 **Eksamensresultater i matematikk og fysikk**, karakterisert i sannsynlighetstabellen nedenfor.

Oppgave
Hvis en student trekkes tilfeldig fra studentmassen som ligger bak sannsynlighetene i tabellen nedenfor, hva er da sannsynligheten for a) studenten har godt resultat i minst ett av fagene, b) studenten har ikke dårlig resultat i matematikk?

		Fysikkresultater		
		Godt	Middels	Dårlig
Matematikk-resultater	Godt	.06	.11	.18
	Middels	.12	.18	.10
	Dårlig	.16	.05	.04

Løsningsforslag

a) Definer hendelsene A og B som følger:
A: godt matematikkresultat, B: godt fysikkresultat

Vi får da:

P(godt resultat i minst ett fag)
$= P(A \cup B) = P(A) + P(B) - P(AB)$
$P(A) = .06 + .11 + .18 = .35$
$P(B) = .06 + .12 + .16 = .34$
$P(AB) = .06$

Følgelig: P(godt resultat i minst ett fag) $= .35 + .34 - .06 = \underline{.63}$

b) La D være hendelsen at studenten har dårlig matematikkresultat. Vi får da:

P(ikke dårlig resultat i matematikk) $= P(D^C) = 1 - P(D)$
$P(D) = .16 + .05 + .04 = .25$

Følgelig: P(ikke dårlig resultat i matematikk) $= 1 - .25 = \underline{.75}$ ☺

Urnemodell

Vi tenker oss en urne som inneholder n «merkete lapper». Ved utvalg på k tilfeldige trekninger fra urna skal vi skille mellom:

- Trekning **uten** tilbakelegging, der vi aldri kan trekke én og samme lapp mer enn én gang, og trekning **med** tilbakelegging, der vi kan trekke samme lapp flere ganger.
- **Ordnede** utvalg, der rekkefølgen av trekningene er vesentlig, og **uordnede** utvalg, der rekkefølgen er uvesentlig.

Vi skal benytte de forkortete betegnelsene (uten, uordnet), (uten, ordnet), (med, uordnet) og (med, ordnet) for de fire kombinerte tilfellene vi kan ha. For eksempel betyr (med, uordnet) utvalg ved trekning med tilbakelegging der rekkefølgen av trekningene er uvesentlig.

2.5 Kombinatorikk, telleregler

Dersom utfallsrommet består av i alt k like sannsynlige enkeltutfall, $P(e_1) = \cdots = P(e_k) = 1/k$, så kan vi anvende den uniforme sannsynlighetsmodellen. En hendelse, A, som består av i alt m av de k enkeltutfallene, vil da ha sannsynlighet $P(A) = m/k$ for å inntreffe. Med denne modellen i bakhodet, skal vi studere 3 viktige telleregler:

1) **Potensregelen**
2) **Ordningsregelen**
3) **Kombinasjonsregelen**

Ved løsning av oppgaver i sannsynlighetsregning, vil det ofte være til hjelp å tenke seg en urne som vi trekker tilfeldige lapper fra. Før vi går løs på tellereglene ovenfor, skal vi derfor se nærmere på den urnemodellen vi skal benytte.

La oss belyse urnemodellen ved følgende eksempler:

- tilfeldig utfylling av tipperekke med 12 kamper (med, ordnet)
- antall seksere etter 5 terningkast (med, uordnet)
- tilfeldig gjetting på de 3 første lagene i eliteserien i fotball (uten, ordnet)
- tilfeldig utfylling av Lotto-kupong (uten, uordnet)

Først betrakter vi tipperekka: Vi kan her tenke oss tre lapper som er merket henholdsvis *H* (hjemmeseier), *U* (uavgjort) og *B* (borteseier). Først trekker vi en lapp, og lar utfallet på lappen angi hva vi skal fylle ut på 1. kamp. Så legger vi lappen tilbake, og gjentar eksperimentet 12 ganger, slik at vi får fylt ut alle kampene. Rekkefølgen er her vesentlig (det er ikke nok å ha antall *H*, *U* og *B* rett, vi må også ha dem riktig plassert), så vi får et ordnet utfall. Ved å fylle ut rekke etter rekke på denne måten, vil alle kombinasjoner være like sannsynlige, og vi kan bruke vår uniforme sannsynlighetsmodell.

Terningkast kan vi overføre til vår urnemodell ved at vi tenker oss 6 lapper i urna. På hver lapp står det et heltall mellom 1 og 6, og alle lappene er forskjellige. Å kaste en terning tilsvarer da å trekke en tilfeldig lapp. Etter at vi har lest hva som står på lappen, må vi legge denne tilbake før vi trekker på nytt for å simulere neste terningkast. Dette gjentas 5 ganger, og vi teller opp hvor mange ganger vi har trukket en sekser. Rekkefølgen er her uvesentlig, det er kun antall seksere som teller, uansett når de ble trukket. Dette er derfor et eksempel på et *uordnet* utvalg *med* tilbakelegging.

La oss så betrakte gjetting på de 3 første lagene i eliteserien i fotball. Si at vi har 14 lag totalt. Vi tenker oss da at vi lager 14 lapper med forskjellige navn, der det på hver lapp står navnet på et av lagene. Deretter trekker vi tilfeldig 3 av lappene, *uten* å legge noen av de trukne lappene tilbake igjen. Også her får vi en *ordnet* rekkefølge – vi antar at det ikke er uvesentlig å skille mellom hvem som får gull, sølv og bronse. Ved en slik trekning vil enhver kombinasjon bli like sannsynlig og vi kan igjen bruke den uniforme sannsynlighetsmodellen.

Til slutt ser vi på Lotto-kupongen. Vi tenker oss her 34 nummererte lapper med nummer fra 1 til 34, der alle lappene har forskjellige nummer. Deretter trekker vi 7 tilfeldige lapper, *uten* å legge noen av lappene tilbake etter at de er trukket. Her er rekkefølgen vi trekker tallene i uvesentlig, så vi får en *uordnet* rekkefølge. Enhver tallkombinasjon vil på denne måten bli like sannsynlig som en hvilken som helst annen tallkombinasjon, og en uniform sannsynlighetsmodell kan anvendes.

Vi gjennomgår nå hver av reglene i den rekkefølge de er introdusert.

Potensregelen

Vi betrakter et sammensatt eksperiment som består av k «like» deler, der utfallsrommet for hver del av eksperimentet består av m enkeltutfall. Det totale utfallet av eksperimentet er da en samling av k enkeltutfall. At de ulike eksperimentdelene er «like», betyr at det er ett og samme utfallsrom for hvert deleksperiment. For et slikt eksperiment er det i alt

(2.11) m^k = totalt antall (ordnete) utfall

der

k = antall «like» deleksperimenter
m = antall enkeltutfall i utfallsrommet til hvert deleksperiment

Eks. 2.10 **Tippekupong**. En tippekupong består av 12 kamper, der hver kamp ender med hjemmeseier (H), uavgjort (U) eller borteseier (B). Ett utfall av eksperimentet (dvs. utfylling av en rekke på tippekupongen) kan da f.eks. være $\{HUHBHUHUBBHH\}$, som betyr hjemmeseier i 1. kamp, uavgjort i 2. kamp, hjemmeseier i 3. kamp, osv.

Oppgave
Hva er sannsynligheten for 12 rette dersom du fyller ut en rekke tilfeldig?

Løsningsforslag
Her er $k = 12$ og $m = 3$ og vi får totalt $m^k = 3^{12} = 531441$ rekker, som alle er like sannsynlige. Kun én rekke gir 12 rette, og sannsynligheten for 12 rette med en enkeltrekke (ett eksperiment) blir da $P(12 \text{ rette}) = 1/(3^{12}) = 1.9\cdot 10^{-6}$. ☺

Ordningsregelen (permutasjonsregelen)

Antall mulige måter vi kan ordne (stokke) r elementer av totalt n elementer på har betegnelsen P_r^n, som kan leses som «antall mulige ordninger (permutasjoner) av r av n elementer». Matematisk får vi:

(2.12) $P_r^n = n\cdot(n-1)\cdots(n-r+1)$

Eks. 2.11 **Resultatliste**. 15 syklister deltar i en motorsykkelkonkurranse. På hvor mange måter kan, teoretisk, øverste delen av resultatlista bestående av nr. 1, 2 og 3 se ut?

Løsningsforslag
Vi innser at n = 15, r = 3 og får P_3^{15} = 15·14·13 = 2730 muligheter. Resonnementet kan gå som følger: 15 forskjellige personer kan vinne. For hver av disse 15 personene er det 14 igjen som kan få 2.-plassen. Dette gir 15·14 = 210 kombinasjoner. For hver av disse 210 kombinasjonene er det 13 personer som kan komme på 3. plass. Dette gir totalt 210·13 = 2730 muligheter. ☺

Et spesialtilfelle av ordningsregelen får vi når r = n. I dette tilfellet får vi

(2.13) $$P_n^n = n\cdot(n-1)(n-n+1) = 1\cdot 2 \cdots n = n!$$

***n*!** er en spesiell og alternativ skrivemåte for P_n^n, og kalles ***n* fakultet**.

Eks. 2.12 **Kø**. 6 personer står etter hverandre i kø. I hvor mange forskjellige rekkefølger kan de 6 personene stå?

Løsningsforslag
Her er n = 6 og vi får 6! = 6·5·4·3·2·1 = 720 kombinasjoner. Resonnementet kan gå som følger: Tenk deg at de 6 personene er nummerert fra 1 til 6. Person 1 kan stå på 6 forskjellige plasser i køen (først, nestførst,…, bakerst). For en gitt plassering av person 1 har person 2 fem mulige plasseringer, hvilket gir 6·5 mulige kombinasjoner for person 1 og 2. For hver av disse 30 kombinasjonene, har person 3 fire mulige plasseringer, osv. ☺

Kombinasjonsregelen

Antall mulige måter å kombinere r av n objekter, der rekkefølgen av de r objektene innbyrdes er uvesentlig, har betegnelsen $\binom{n}{r}$ som leses: «antall kombinasjoner for r av n elementer». Matematisk får vi at

$$(2.14) \qquad \binom{n}{r} = \frac{P_r^n}{r!} = \frac{n\cdot(n-1)\cdots(n-r+1)}{1\cdot 2\cdots r} = \frac{n!}{r!\cdot\,(n-r)!}$$

Binomialformelen til venstre ovenfor har blant annet følgende egenskaper:

$$(2.15) \qquad \binom{n}{r} = \binom{n}{n-r}, \qquad \binom{n}{n} = \binom{n}{0} = 1, \quad \binom{n}{1} = n$$

Kombinasjonsregelen kan begrunnes som følger:

Antall måter å *ordne r* av *n* objekter	er det samme som	antall måter å *kombinere* (velge) r av n objekter	multiplisert med	antall måter å ordne de r valgte objektene innbyrdes
P_r^n	=	$\binom{n}{r}$	$\cdot$	$r!$

Eks. 2.13 **Bruk av kombinasjonsregelen**

Oppgave

Beregn alle mulige a) *kombinasjoner* og b) *ordninger* av 3 forskjellige bokstaver valgt fra de 4 bokstavene *A*, *B*, *C* og *D*.

(NB! Husk at forskjellen på kombinasjoner og ordninger er at rekkefølgen er uvesentlig når det gjelder kombinasjoner):

Løsningsforslag

Kombinasjoner:	Ordninger (permutasjoner):
{ *A*, *B*, *C* }	*ABC*, *ACB*, *BAC*, *BCA*, *CAB*, *CBA*
{ *A*, *B*, *D* }	*ABD*, *ADB*, *DAB*, *DBA*, *BAD*, *BDA*
{ *A*, *C*, *D* }	*ACD*, *ADC*, *CAD*, *CDA*, *DAC*, *DCA*
{ *B*, *C*, *D* }	*BCD*, *BDC*, *CBD*, *CDB*, *DBC*, *DCB*

Som vi ser har vi 4 forskjellige kombinasjoner. For hver av disse kombinasjonene, som består av 3 elementer, er det $3! = 1 \cdot 2 \cdot 3 = 6$ mulige måter å ordne (stokke) rekkefølgen av de 3 bokstavene på. Totalt har vi $P^{4,3} = 4 \cdot 3 \cdot 2 = 24$ mulige måter vi kan ordne $k = 3$ av totalt $n = 4$ bokstaver på. Ved litt refleksjon forstår vi at antall kombinasjoner av 3 bokstaver fremkommer ved å ta totalt antall ordninger, P^n_r og dividere på antall måter bokstavene i hver kombinasjon kan ordnes på, dvs $r!$. Vi får da den matematiske betgnelsen, $\binom{n}{r}$, for antall kombinasjoner av r av n. ☺

Eks. 2.14 **Komité**. En rådgivende komité for en fengselsreform består av 15 medlemmer. 9 er for reformen, 4 er imot reformen og 2 mener det er hipp som happ om reformen blir innført eller ikke. En reporter ønsker å velge 3 tilfeldige personer fra komiteen for å meddele deres syn i et TV-program.

Oppgave

a) Hva er sannsynligheten for at minst 2 av de utvalgte personene vil være for reformen?

b) Hva er sannsynligheten for at de 2 første personene vil være for reformen, og den tredje vil være imot?

NB! Legg merke til at i oppgave a) er rekkefølgen uvesentlig, mens dette ikke er tilfelle i oppgave b).

Løsningsforslag

a) 3 av 15 personer kan bli valgt ut på ialt $\binom{15}{3} = (15 \cdot 14 \cdot 13)/(1 \cdot 2 \cdot 3) = 455$ måter, som alle er like sannsynlige (tilfeldig utvalg). La A_2 og A_3 betegne hendelsene «eksakt 2 for» og «eksakt 3 for». Oppgave a) består i å finne $P(A_2 \cup A_3)$. Disse hendelsene er gjensidig ekskluderende, og vi har da at $P(A_2 \cup A_3) = P(A_2) + P(A_3)$.

For å beregne $P(A_2)$, må vi beregne antall måter vi kan trekke 3 av 15 personer, slik at 2 av dem blir trukket blant gruppa på 9 som er for reformen, og 1 blir trukket av gruppa på 6 som er enten imot eller likegyldige:

$$\binom{9}{2} \cdot \binom{6}{1} = \frac{9 \cdot 8}{1 \cdot 2} \cdot 6 = 216$$

Ved å anvende uniform sannsynlighetsmodell får vi da at hendelsen A_2 består av $m = 216$ like sannsynlige enkeltutfall, mens utfallsrommet består av $\binom{15}{3} = 455$ like sannsynlige enkeltutfall:

$$P(A_2) = \frac{\binom{9}{2}\binom{6}{1}}{\binom{15}{3}} = \frac{216}{455}$$

Ved tilsvarende resonnement for A_3 får vi:

$$P(A_3) = \frac{\binom{9}{3}\binom{6}{0}}{\binom{15}{3}} = \frac{84}{455}$$

og den søkte sannsynlighet blir:

P(minst 2 av 3 for) = P(A_2) + P(A_3) = 300/455 = <u>60/91</u>

b) Vi har totalt $P_3^{15} = 15 \cdot 14 \cdot 13 = 2730$ ordnede utvalg for 3 personer fra en gruppe på 15, og alle de ordnede utvalg er like sannsynlige (tilfeldig «trekning»). Antallet ordnede utvalg der de 2 første personene er for og den tredje er mot, er

$$P_2^9 \cdot P_1^4 = (9 \cdot 8) \cdot (4) = 288$$

$\Rightarrow$ P (første 2 for, tredje mot) = <u>288/2730</u> ☺

La oss oppsummere dette underkapitlet med å gjengi følgende hjelpetabell som viser sammenhengen mellom hvilken type trekning vi har (uten eller med tilbakelegging, uordnet eller ordnet rekkefølge), og hvilken telleregel som kommer til anvendelse:

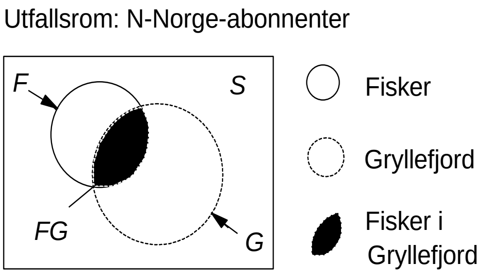
		Rekkefølge: uordnet	ordnet
Tilbake-	uten	Kombinasjonsregel	Ordningsregel
legging:	med	---------------	Potensregel

Gå nøye gjennom tabellen selv, og kontrollér at rubrikken (med,uordnet) ikke dekkes av noen av de tellereglene vi har behandlet. Vi kommer senere tilbake til dette tilfellet i forbindelse med binomisk fordeling.

2.6 Betinget sannsynlighet

Sannsynligheten for en hendelse, A, må ofte modifiseres dersom en får oppgitt informasjon(er) om en annen hendelse, B. Eks: Betrakt et tilfeldig telefonnummer i Nord-Norge-katalogen(e). La F og G betegne hendelsene:

F: Abonnenten er fisker G: Abonnenten bor i Gryllefjord



Situasjonen er illustrert i Venn-diagrammet til venstre. *Utfallsrommet S* består av alle telefonnumrene i N-Norge, og utgjør hele boksen. Den *heltrukne* sirkelen inneholder alle numrene der abonnenten er fisker, og den *stiplede* sirkelen inneholder alle numrene der abonnenten bor i Gryllefjord. Det *skraverte* området der de to sirklene overlapper hverandre, inneholder alle numrene der abonnenten *både* er fisker *og* bor i Gryllefjord (dvs. snitthendelsen FG). Området som ligger utenfor begge sirklene inneholder alle numrene der abonnenten *verken* er fisker *eller* bor i Gryllefjord (dvs. hendelsen $(F \cup G)^C$)

Tenk deg nå at alle Nord-Norge-numrene er lagt på data, og et regnemaskinprogram sørger for at du «trekker» et nummer tilfeldig. Sannsynligheten $P(G)$ for at nummeret skal tilhøre en abonnent i Gryllefjord vil da være antall abonnenter i Gryllefjord delt på antall abonnenter i Nord-Norge totalt.

Sett nå at programmet gir deg mulighet til først å bestemme retningsnummer (sted), og deretter å trekke tilfeldig blant alle numre med dette retningsnummeret. Si f.eks. at du ønsker å trekke blant Gryllefjord-numrene. Sannsynligheten for å trekke et nummer der abonnenten er fisker er da større enn om du trakk blant alle Nord-Norge-numrene (prosentvis flere fiskere i Gryllefjord enn i Nord-Norge totalt). Med henvisning til Venn-diagrammet ovenfor, blir sannsynligheten i dette tilfellet lik antall abonnenter i Gryllefjord som er fiskere delt på totalt antall abonnenter i Gryllefjord. Vi snakker nå om en *betinget* sannsynlighet for en hendelse F (fisker) fordi vi har *gitt* at en tilleggshendelse, G (Gryllefjord-abonnent), har inntruffet. Vi benytter i dette tilfellet en loddrett strek for å skille hendelsen vi ønsker å finne sannsynligheten for, fra de tilleggsbetingelser som er oppgitt:

$P(F|G) = P($«abonnenten er fisker *gitt* at abonnenten bor i Gryllefjord»$)$

NB! *Tilleggsbetingelsen(e)* er alltid oppgitt til *høyre* for den loddrette streken, mens den hendelse vi skal finne sannsynligheten til er oppgitt til *venstre* for streken.

Med henvisning til Venn-diagrammet, skjønner vi intuitivt at $P(F|G)$ = antall abonnenter i Gryllefjord som er fiskere delt på antall abonnenter i Gryllefjord totalt. Dette er ofte en praktisk og mulig måte å bestemme betinget sannsynlighet på. I mange tilfeller er det imidlertid nødvendig, og/eller langt mer lettvint, å bruke mer indirekte metoder, som vi skal se.

Generelt gjelder følgende:

Betinget sannsynlighet
Sannsynligheten for en hendelse A under betingelsen (tilleggsopplysningen) at en hendelse B har inntruffet, har følgende betegnelse:

(2.16) $P(A \mid B)$ som leses: «Sannsynligheten for A gitt B».

Legg merke til at betingelsen står på *høyre* side av den loddrette streken, mens den hendelsen vi skal finne sannsynligheten til, står på venstre side. Følgende viktige regler gjelder for betinget sannsynlighet:

(2.17) $P(AB) = P(A) \cdot P(B \mid A)$

(2.18) $P(A \mid B) = \dfrac{P(B \mid A) \cdot P(A)}{P(B)}$ (**Bayes' formel**)

La oss nå belyse formlene ovenfor ved å gå tilbake til telefonnummer-eksemplet ($A = G$, $B = F$). Ved hjelp av lign.(2.17) får vi:

$$P(F|G) = \frac{P(FG)}{P(G)}$$

som omsatt til ord kan formuleres som følger:

P (tilfeldig N-Norge-abonnent i Gryllefjord er fisker) =
P (tilfeldig N-Norge-abonnent er både fisker og bor i Gryllefjord) delt på
P (tilfeldig N-Norge-abonnent bor i Gryllefjord).

Sett nå at vi var interessert i å finne $P(G \mid F)$, mens vi kjente $P(G)$, $P(F)$ og $P(F \mid G)$. Vi kunne da brukt Bayes' formel (se lign.(2.18)):

$$P(G|F) = \frac{P(F|G) \cdot P(G)}{P(F)}$$

som omsatt til ord kan formuleres som følger:

P (tilfeldig N-Norge-abonnent som er fisker bor i Gryllefjord) =
P (tilfeldig N-Norge-abonnent i Gryllefjord er fisker) ganget med
P (tilfeldig N-Norge-abonnent bor i Gryllefjord) delt på
P (tilfeldig N-Norge-abonnent er fisker)

Eks. 2.15 **Avissalg**. En kiosk fører statistikk over hvor mange kunder som kjøper Dagbladet (*D*), hvor mange som kjøper Aftenposten (*A*), og hvor mange som kjøper begge aviser. Resultat:

40 % kjøper Dagbladet, og av disse kjøper 20 % Aftenposten. 30 % kjøper Aftenposten.

Oppgave
Hva er sannsynligheten for at en tilfeldig kunde

a) Kjøper både Dagbladet og Aftenposten ?
b) Kjøper verken Dagbladet eller Aftenposten ?
c) Kjøper Dagbladet dersom vedkommende kjøper Aftenposten ?

Løsningsforslag
Vi bruker forkortelsene *D* og *A* for henholdsvis hendelsene «tilfeldig kunde kjøper Dagbladet» og «tilfeldig kunde kjøper Aftenposten».

a) *P*(«både Dagbladet og Aftenposten»)

$= P(DA) = P(D) \cdot P(A \mid D) = 0.4 \cdot 0.2 = \underline{0{,}08}$

b) *P*(«Verken Dagbladet eller Aftenposten») $= P((D \cup A)^C) = 1 - P(D \cup A)$

$P(D \cup \mathrm{A}) = P(D) + P(A) - P(DA) = 0.04 + 0.3 - 0.08 = 0{,}62$

$\Rightarrow \ 1 - P(D \cup A) = \underline{0.38}$

c) Ved hjelp av Bayes' formel finner vi tilslutt:

$P(D ¤ A) = P(A \mid D) \cdot P(D)/P(A) = 0.2 \cdot 0.4/0.3 = \underline{0.267}$ ☺

Tilbake til det generelle: Lign.(2.17) kan generaliseres til flere enn 2 hendelser. For 3 hendelser A, B og C får vi:

(2.19) $$P(ABC) = P(A) \cdot P(B \mid A) \cdot P(C \mid BA)$$

Som huskeregel for høyre side av lign.(2.19) kan vi notere følgende: Vi begynner fra venstre i snitthendelsen ABC, dvs. vi begynner med $P(A)$. Så går vi til neste hendelse, nemlig B, og betinger med alt til venstre for B, dvs. A, og vi får $P(B \mid A)$. Slik fortsetter vi, og får C betinget med alt til venstre for C, nemlig AB. Den siste faktoren blir da $P(C \mid AB)$.

Merk at vi godt kunne snudd rekkefølgen i lign. (2.19), da faktorenes orden i snitthendelsen ABC er likegyldig. Eks:

(2.20) $$P(CBA) = P(C) \cdot P(B \mid C) \cdot P(A \mid CB)$$

I praksis vil en ofte bruke lign. (2.19) med A, B og C definert i kronologisk rekkefølge, dvs. A er den hendelsen som opptrer først, så kommer B og til slutt C.

Eks. 2.16 **Egg**. En gårdbruker har en eske som inneholder 30 egg, derav 5 med blodflekker. Han sjekker 3 egg ved å trekke dem tilfeldig en etter en fra eska.

Oppgave
Hva er sannsynligheten for at de første to eggene vil inneholde blodflekker og det tredje vil være uten blodflekker?

Løsningsforslag
La $R_1R_2H_3$ betegne snitthendelsen å trekke blodflekkete egg i de 2 første trekningene og klart egg i 3. trekning (R for rød, H for hvit). Vi bruker lign.(2.19) med $A = R_1$, $B = R_2$ og $C = H_3$:

$$P(R_1R_2H_3) = P(R_1) \cdot P(R_2 \mid R_1) \cdot P(H_3 \mid R_1R_2)$$

Fordi trekningene er tilfeldige, har vi at $P(R_1) = \binom{5}{1}/\binom{30}{1} = 5/30$. Vi ser så på $P(R_2 \mid R_1)$. Her er betingelsen R_1 gitt, dvs. vi skal finne sannsynligheten for å trekke et blodflekket egg i 2. trekning gitt at det er trukket et blodflekket egg i 1. trekning. Da er 29 egg igjen, og 4 av disse er med blodflekker: $P(R_2 \mid R_1) = 4/29$. Tilsvarende resonnement gir at $P(H_3 \mid R_1R_2) = 25/28$, og vi får:

$$P(R_1R_2H_3) = (5/30) \cdot (4/29) \cdot (25/28) = \underline{25/1218}$$

Her kunne vi også løst problemet ved å bruke reglene for ordnet utvalg uten tilbakelegging: Det er totalt P_3^{30} like sannsynlige måter å ordne $r = 3$ av $n = 30$ elementer uten tilbakelegging. Antallet ordnede utvalg der de 2 første eggene er flekkete og det tredje er uflekket er $P_2^5 \cdot P_1^{25}$. Vi får derfor:

$$P(R_1R_2H_3) = \left(P_2^5\right) \cdot \left(P_1^{25}\right) / \left(P_3^{30}\right) = (5 \cdot 4) \cdot (25) / (30 \cdot 29 \cdot 28) = \underline{25/1218}$$ ☺

2.7 Uavhengige hendelser

Begrepet *uavhengige* hendelser er et svært sentralt begrep nøye koplet til begrepet betinget sannsynlighet:

> **Uavhengige hendelser**
> To hendelser *A* og *B* er uavhengige hvis og bare hvis
>
> (2.21) $P\,(A \mid B) = P\,(A)$
>
> Dersom $P(B) \neq 0$, er betingelsen ovenfor ekvivalent med at
>
> (2.22) $P\,(AB) = P(A) \cdot P(B)$

NB! To hendelser, *A* og *B*, som er **gjensidig ekskluderende** ($AB = \emptyset$), er *aldri* uavhengige, og omvendt.

Eks. 2.17 **To kast med en rettferdig mynt**. $S = \{ KK, KM, MK, MM \}$, der K = kron, M = mynt og alle enkeltutfallene er like sannsynlige. Vi definerer følgende 3 hendelser:

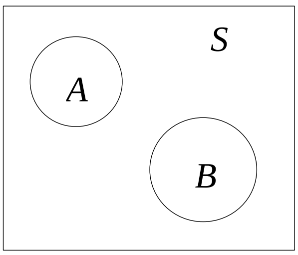
A: Kron (*K*) i første kast
B: Kron i andre kast
C: Kron i begge kast eller mynt i begge kast

Oppgave
Finn ut hvilke par av hendelsene ovenfor som er uavhengige.

Løsningsforslag
Intuitivt burde A og B være uavhengige hendelser. Det er imidlertid ikke like opplagt om A og C er uavhengige hendelser, eller om B og C er uavhengige. Ved regning finner vi (siden hvert enkeltutfall har sannsynlighet 1/4):

$$P(A) = 1/2, \quad P(AB) = 1/4 = P(A)\cdot P(B)$$
$$P(B) = 1/2, \quad P(AC) = 1/4 = P(A)\cdot P(C)$$
$$P(C) = 1/2, \quad P(BC) = 1/4 = P(B)\cdot P(C)$$

Følgelig, alle par av hendelser er uavhengige. Legg også merke til at ingen av hendelsene er gjensidig ekskluderende, både A, B og C inneholder jo enkeltutfallet KK. Vi vil forøvrig alltid ha at gjensidig ekskluderende hendelser er avhengige. Dette er intuitivt svært logisk: Dersom en av to gjensidig ekskluderende hendelser har inntruffet, vet vi jo med 100% sikkerhet at den andre ikke kan ha inntruffet. ☺



At to hendelser er gjensidig ekskluderende fremkommer i et Venn-diagram ved at de ikke har noe overlapp. I figuren til venstre er hendelsene A og B gjensidig ekskluderende. I dette tilfellet skjønner vi at $P(AB) = 0$.

På tampen minner vi om 2 viktige formler:

1) $P(A) = 1 - P(A^c)$
2) $P(A \cup B) = P(A) + P(B) - P(AB)$

Dersom A og B er gjensidig ekskluderende, forenkles 2) til $P(A \cup B) = P(A) + P(B)$.

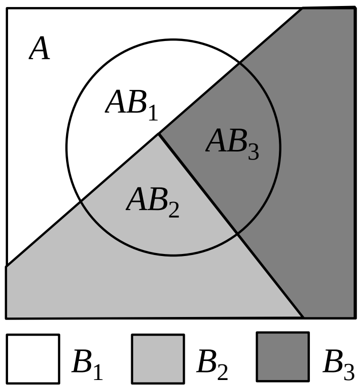
Gjensidig ekskluderende hendelser og betinget sannsynlighet

Vi skal ta utgangspunkt i at en hendelse B utgjør hele utfallsrommet ($P(B) = 1$), og at B er inndelt i n gjensidig ekskluderende delhendelser: $B_1,\ldots,B_n$. Dette kalles også en **partisjon** av S, og medfører at $P(B_1 \cup B_2 \cup \cdots \cup B_n) = P(B_1) + \cdots + P(B_n) = 1$. La videre A være en hendelse vi ønsker å bestemme sannsynligheten til, der de betingete sannsynlighetene $P(A \mid B_1),\ldots,P(A \mid B_n)$ er kjente. Anta også at sannsynlighetene $P(B_1),\ldots, P(B_n)$ er kjente størrelser. Da får vi følgende (husk at $P(B) = 1$):

(2.21) $P(A) = P(AB) = P(A(B_1 \cup \cdots \cup B_n)) = P(AB_1) + \cdots + P(AB_n)$

Å innse at lign.(2.21) er riktig, er trolig enklest ved hjelp av Venn-diagram, der et eksempel med $n = 3$ er vist i neste figur. A er hendelsen inneholdt i sirkelen, mens B utgjør hele utfallsrommet (boksen).

Bruker vi nå lign.(2.17) på hvert av deluttrykkene til høyre i lign.(2.21) får vi:



(2.22) $P(A) = P(AB) = P(AB_1) + P(AB_2) + \cdots + P(AB_n)$

$$= P(A|B_1) \cdot P(B_1) + \cdots + P(A|B_n) \cdot P(B_n)$$

$$= \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

2.8 Oppgaver

2.1 Du trekker ett lodd fra en urne med 100 lodd. 3 lodd gir kakegevinst (K), 1 lodd gir bokgevinst (B) og 96 lodd gir ingen gevinst (I).

Skriv opp et passende utfallsrom, S, for det trukne loddet.

2.2 Som i 2.1, men du trekker nå to lodd.

2.3 En urne inneholder en rød (R), en blå (B) og en gul (G) kule. Hvor mange enkeltutfall består utfallsrommet S av i følgende tilfeller:

a) 2 trekninger uten tilbakelegging, rekkefølgen vesentlig.
b) 2 trekninger med tilbakelegging, rekkefølgen uvesentlig.
c) 3 trekninger uten tilbakelegging, rekkefølgen vesentlig.

2.4 Et ruteark med $3 \cdot 3 = 9$ ruter skal fylles ut med 5 like kryss og 4 like ringer. Hvor mange kombinasjoner får vi?

2.5 Et brett med $3 \cdot 3$ ruter skal dekkes med 9 forskjellige brikker. Hvor mange kombinasjoner får vi?

2.6 En terning kastes en gang. Hendelse A betyr at terningen viser en sekser, mens hendelse B betyr at terningen viser høyst 4 øyne. Forklar hva følgende hendelser betyr, og beregn sannsynligheten for hver av dem:

a) $(A \cup B)^C$
b) $A^C \cap B^C$
c) $(A \cap B)^C$
d) $A^C \cup B^C$

2.7 Vis matematisk at

$$P(A \cup B \cup C) = P(A)+P(B)+P(C) -P(AB)-P(AC)-P(BC)+P(ABC)$$

2.8 Gitt utfallsrommet $S = \{e_1, e_2, e_3, e_4, e_5\}$ der alle enkeltutfallene er like sannsynlige. La videre hendelsene A, B og C være definert som følger: $A = \{e_1, e_2, e_4\}$, $B = \{e_2, e_3\}$ og $C = \{e_2, e_4, e_5\}$.

a) Tegn Venn-diagram og tegn inn hendelsene A, B og C.
b) Beregn $P(A)$, $P(A \cup B)$, $P(AC)$, $P(B^C)$, $P[(A \cup B)^C]$ og $P(A^C \cup B^C)$

2.9 Angi telleregel og finn hvor mange enkeltutfall følgende eksperiment består av:

a) Ett myntkast med 5 pengestykker
b) Ett terningkast med 3 terninger
c) Trekning av 1., 2. og 3. premie blant 7 personer med like store vinnersjanser (trekning uten tilbakelegging).
d) Trekning uten tilbakelegging av 3 kuler fra urne med 2 gule og 4 røde kuler.
e) Antall kombinasjoner av 13 kort fra kortstokk med 52 forskjellige kort.

2.10 En oversikt over hvilken utdanning ansatte ved en høyteknologisk bedrift har, er satt opp i følgende tabell:

	ingeniør-utdannet	annen utdanning
kvinner	10	30
menn	20	40

a) Hva er sannsynligheten for at en tilfeldig ansatt er kvinne?
b) Hva er sannsynligheten for at vedkommende er ingeniørutdannet og kvinne?
c) Hva er sannsynligheten for at vedkommende er ingeniørutdannet når vi vet at vedkommende er kvinne?
d) To tilfeldige ansatte snakker sammen. Hva er sannsynligheten for at ingen av dem er ingeniørutdannet?

2.11 I en klasse på 30 elever kunne 12 danse sving (S), 4 kunne spille gitar (G) og det var 17 jenter (J) i klassen. 2 av jentene spilte gitar. En av guttene og en av jentene i klassen kunne både danse sving og spille gitar. 7 av guttene kunne verken spille gitar eller danse sving.

a) Hvor mange jenter kunne verken danse sving eller spille gitar?
b) Hvor mange gutter kunne danse sving?

(hint: Tegn Venn-diagram over alle elevene i klassen).

2.12 Anta at du har 20 ingredienser tilgjengelig for å lage pizzafyll.

a) Hvor mange forskjellige pizzaer kan bli laget av
   i) eksakt en ingrediens?
   ii) eksakt to ingredienser?
   iii) eksakt tre ingredienser?
b) Svar på ii) dersom to halvparter av en pizza kan være forskjellige.

2.13 Fra en kortstokk på 52 kort gjør vi to trekk etter hverandre uten tilbakelegging. Hva er sannsynligheten for å trekke

a) et svart og et rødt kort?
b) ikke mer enn én honnør (ess, konge, dame, knekt)?

2.14 På en skole med 120 elever er det 70 jenter. 20 gutter og 10 jenter synes matematikk er gøy.

a) Hva er sannsynligheten for at en tilfeldig elev ikke syntes matematikk er gøy, når det viste seg at eleven var en gutt?
b) Hva er sannsynligheten for at to tilfeldige elever har forskjellig kjønn og synes matematikk er gøy?

2.15 Du trekker 13 kort fra en kortstokk på 52 kort der 16 er honnørkort. Beregn sannsynligheten for at trekningen gir:

a) 13 kort med samme farge (hjerter, ruter, kløver eller spar).
b) Ingen honnørkort.
c) Bare honnørkort.

2.16 En 6-kantet kubisk terning har to røde sider (R) og 4 blå sider (B). Terningen kastes 3 ganger.

a) Bruk den generelle addisjonssetningen til å beregne sannsynligheten for minst ett rødt kast.
b) Beregn også sannsynligheten på enklest mulig måte.

***2.17** En tipperekke inneholder 4 kamper, hver med 3 utfall (hjemmeseier (H), uavgjort (U) eller borteseier (B)). Hvor mange rekker må fylles ut for å være sikret minst 3 av 4 rette?

2.18 Sannsynligheten for at en mann er fargeblind er 0.08. Sannsynligheten for at en kvinne er fargeblind er 0.004.

a) Hva er sannsynligheten for at en tilfeldig utvalgt person er fargeblind?

Forutsett at halvparten av befolkningen er kvinner.

b) Det er begått en forbrytelse. Det er åpenbart at forbryteren er fargeblind. Hva er sannsynligheten for at forbryteren er en mann?

2.19 Bladet Motor slår fast i en omfattende artikkel 2/85 at bilparken i Norge er beheftet med betydelige feil. Anta at du besøker en bruktbilforretning som har 25 biler å tilby av en bestemt årsmodell. Av disse har to biler ingen feil, 5 har en feil, 10 har to feil, syv har tre feil og en har flere enn tre feil.

Du kjøper to biler. Hva er sannsynligheten for at du får:

a) begge biler uten feil,
b) én og bare én med høyst to feil,
c) to med minst tre feil hver,
d) to med minst tre feil tilsammen?

2.20 Riktig eller galt?

a) Mulige utfall når vi kaster en rettferdig mynt to ganger er: ingen «kron», en «kron» og to «kron».
b) Hvis en rettferdig mynt har vært kastet 5 ganger med «kron» som utfall i alle 5 kast, er sannsynligheten for å få «mynt» i det 6. kastet større enn 0.5.
c) Det er mulig at vi for to hendelser, der $P(A) = 0.5$ og $P(\mathrm{B}) = 0.7$, kan ha $P(A \cup B) = 1.2$.
d) $P(A|B)$ er alltid mindre enn $P(B)$
e) To ikke-tomme og gjensidig ekskluderende (disjunkte) hendelser er aldri uavhengige.
f) Hvis $P(A) = P(A|B)$, så er A og B gjensidig ekskluderende hendelser.
g) Hvis $P(A) = 0.4$ og $P(B) = 0.3$ og $P(A|B) = 0.8$, så er $P(AB) = 0.2$.

2.21 En student skal opp til eksamen i 4 fag (A, B, C og D), og anslår de respektive sannsynligheter for å bestå eksamen til 0.8, 0.9, 0.7 og 0.5. Hvis vi forutsetter uavhengighet, hva er sannsynligheten for at studenten:

a) består alle 4 eksamener,
b) stryker i minst ett fag,
c) hvis studenten stryker i ett (og bare ett) fag, at han/hun stryker i fag D?

2.22 I en klubb med 10 medlemmer skal det utpekes et styre som består av i alt 3 medlemmer.

a) Hvor mange styrer har en å velge mellom?

Anta at et styre skal bestå av en formann, en nestformann og en sekretær.

b) Hvor mange styrer har en nå å velge mellom?

2.23 (E) En medisin har 2 forskjellige «bivirkninger»: Kvalme (K) og hodepine (H). 20 % av brukerne blir kvalme ($P(K) = 0.2$), og blant de som blir kvalme er det 30 % som også får hodepine ($P(H|K) = 0.3$). 10 % av brukerne får hodepine. De som verken blir kvalme eller får hodepine får ingen bivirkninger (I).

a) Hva er sannsynligheten for at en tilfeldig bruker både blir kvalm og får hodepine?

b) Er hendelsen bivirkning (B) det samme som unionen ($K \cup H$) eller snittet (KH) av K og H? Begrunn svaret og finn deretter $P(B)$ for en tilfeldig bruker.

Det viser seg at røykere (R) er betydelig mer utsatt for bivirkninger enn ikke-røykere (R^C). Av brukerne er det 75 % som ikke røyker, og blant disse er det 4 % som får bivirkninger.

c) Vis at sannsynligheten for å få bivirkning for en røyker er 84%.
d) Finn sannsynligheten for at en tilfeldig bruker uten bivirkninger er ikke-røyker.

2.24 (E) Sannsynligheten for at et tilfeldig besøk hos tannlegen vil resultere i tann-uttrekning er 0.06, sannsynligheten for at det vil resultere i plombering er 0.23 og sannsynligheten for at det vil resultere i både tannuttrekning og plombering er 0.02.

a) Hva er sannsynligheten for at et tilfeldig besøk hos tannlegen vil resultere i tannuttrekning, men ikke plombering?

b) Hva er sannsynligheten for at besøket vil resultere i tannuttrekning, plombering eller begge deler?

c) Hva er sannsynligheten for at besøket vil resultere i tannuttrekning eller plombering, men ikke begge deler?

d) Hva er sannsynligheten for at besøket verken vil resultere i tannuttrekning eller plombering?

2.25 (E) A og B er to kjennetegn slik at $P(A) = 0.25$, $P(B) = p$ og $P(A \cup B) = 0.5$.

a) Finn sannsynligheten $P(A \cap B)$ uttrykt ved p.

b) Hvilken verdi har p når A og B er uavhengige kjennetegn?

2.26 (E) En ny medisin mot en bestemt sykdom blir innført. Forsøk har vist at 70 % av pasientene som får medisinen blir helbredet. Dessverre kan medisinen ha bivirkninger. For en pasient som blir helbredet er sannsynligheten for bivirkninger 0.25. For en pasient som ikke blir helbredet er sannsynligheten for bivirkninger 0.1.

a) Hva er sannsynligheten for at en tilfeldig pasient skal bli helbredet og dessuten ikke skal få bivirkninger?

b) Hva er sannsynligheten for bivirkninger hos en tilfeldig pasient?

c) Hva er sannsynligheten for at en pasient blir helbredet når vi vet at han ikke får bivirkninger av medisinen?

2.27 (E) I en klubb er 60 % av medlemmene kvinner. 50 % av kvinnene røyker og 30 % av mennene røyker. Et medlem trekkes ut på måfå.

a) Hva er sannsynligheten for at medlemmet røyker?

b) Hva er sannsynligheten for at medlemmet er en kvinne, forutsatt at vedkommende røyker?

c) Hvor mange prosent av ikke-røykerne er menn?

2.28 (E) FBI bruker en løgndetektor som viser «skyldig» med sannsynlighet 0,92 for personer som virkelig er skyldige i en bestemt forbrytelse. Dersom løgndetektoren brukes på en uskyldig person vil den vise «uskyldig» med sannsynlighet 0.98.

Vi definerer følgende hendelser:

A: Personen er «skyldig».
B: Løgndetektoren viser «skyldig».

Det velges en tilfeldig person fra en (mistenkelig) gruppe der 10 % er «skyldige».

a) Skriv opp sannsynlighetene $P(A)$, $P(B \mid A)$, $P(B^C \mid A^C)$ og beregn

   $P(A^C)$, $P(B^C \mid A)$, $P(B \mid A^C)$ og $P(AB)$.

   Forklar hva den siste sannsynligheten står for.

b) Finn sannsynligheten for at løgndetektoren skal vise skyldig.

c) Anta at detektoren viser «uskyldig» for personenen. Hva er sannsynligheten for at personen i virkeligheten er «uskyldig»?

2.29 (E) I en urne er det 4 røde og 2 grønne kuler. Bortsett fra farge er de like. I denne oppgaven skal vi finne sannsynligheten for å trekke 2 røde og 1 grønn kule ved forskjellige trekningsprosedyrer.

a) Bestem sannsynligheten for å trekke 2 røde og 1 grønn kule dersom trekningen er *uten* tilbakelegging.

b) Hva blir sannsynligheten for å trekke 2 røde og 1 grønn kule dersom trekningen er *med* tilbakelegging?

c) Trekk 3 kuler på følgende måte: Trekk først 2 *uten* tilbakelegging, registrer farge og legg dem så oppi urna igjen før 3. trekk gjøres. Hva blir nå sannsynligheten for å få 2 røde og 1 grønn kule?

2.30 (E) I et bygg på 6 etasjer kommer det 5 personer inn i heisen i 1. etasje. Personene har ingenting med hverandre å gjøre, og det er like stor sannsynlighet for at hver enkelt skal gå av i etasje 2-6 (5 stopp).

a) Hvor stor sannsynlighet er det for at alle går av i 2. etasje?

b) Hvor stor sannsynlighet er det for at alle går av i samme etasje?

c) Hvor stor sannsynlighet er det for at ingen skal til 6. etasje?

d) Hva er sannsynligheten for at det går av én i hver etasje?

e) Anta at det i starten bare var 4 personer i heisen. Hva er da sannsynligheten for at det går av høyst én i hver etasje?

2.9 Formelsamling

Relativ frekvens, $r_N(A)$

$$r_N(A) = \frac{\text{Antall forsøk med } A\text{-utfall}}{\text{Totalt antall forsøk } N}$$

Sannsynlighet, $P(A)$

$$P(A) = \lim_{N\to\infty} r_N(A)$$

Sannsynlighetslover

$$P(A^C) = 1 - P(A)$$

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

Potensregel

Gitt et forsøk som består av k like del-eksperiment, der alle deleksperiment har samme utfallsrom bestående av m like sannsynlige enkeltutfall, og der utfallet av forsøket er en **ordnet** rekkefølge av tilfeldige utfall av de k deleksperimentene. Vi har da totalt:

$\boldsymbol{m^k}$ like sannsynlige utfall.

Ordningsregel

r «nummererte» av n «nummererte» elementer kan ordnes på ialt

$$P_r^n = n\cdot(n-1)\cdots(n-r+1)$$

forskjellige og like sannsynlige måter (**ordnet** rekkefølge).

$P_n^n = 1\cdot 2\cdots n = n!$ som kalles «n **fakultet**». 0! = 1 pr. definisjon

Kombinasjonsregel

r «nummererte» av n «nummererte» elementer, der rekkefølgen er uvesentlig (uordnet rekkefølge), kan kombineres på

$$\binom{n}{r} = \frac{P_r^n}{n!} = \frac{n!}{r!\cdot(n-r)!}$$

forskjellige og like sannsynlige måter.

$$\binom{n}{k} = \binom{n}{n-k},\ \binom{n}{0} = \binom{n}{n} = 1$$

Betinget sannsynlighet

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

der $P(A|B)$ leses «sannsynligheten for A gitt B», og likheten mellom første og siste uttrykk ovenfor er **Bayes**' formel.

La B utgjøre hele utfallsrommet ($P(B)$ = 1) bestående av n gjensidig ekskluderende delhendelser (partisjon) $B_1,\ldots,B_n$. Da gjelder:

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)\cdot P(B_i)$$

For en sammensatt snitthendelse ABC gjelder:

$$P(ABC) = P(A)\cdot P(B|A)\cdot P(C|AB)$$

Uavhengige hendelser

To hendelser, A og B, er uavhengige hvis og bare hvis

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

Dersom $P(B) \neq 0$ er ligninga over ekvivalent med at $P(AB) = P(A)\cdot P(B)$.

Union

$A\cup B$ = mengden som består av alle elementer som er med i A eller i B eller i både A og B.

Snitt

$A\cap B = AB =$ Mengden som består av alle elementer som er med både i A og i B.

Komplement

A^C = Mengden av alle elementer som ikke er med i A.