Mengdelære

Uttrykket
$$A = \{a_1, a_2, \cdots, a_n\}$$
kalles en mengde, og $a_1, a_2, \cdots, a_n$ kalles mengdens elementer. To eksempler på mengder er
$$A = \{1, 2\}$$
og
$$B = \{1, 3, 4\}.$$
Vi kan kombinere mengder til nye mengder. De viktigste operasjonene er union, snitt og komplement.

- $A \cup B$: Alt som er i enten A eller B eller begge.
- $A \cap B$: Alt som er i både A og B.
- $\overline{A}$: Alt som ikke er i A.

Eksempel. Hvis $A = \{1, 2\}$ og $B = \{1, 3, 4\}$, er
$$A \cup B = \{1, 2, 3, 4\},$$
$$A \cap B = \{1\}$$
og
$$\overline{B} = \{2, 5, 6\}.$$
△

Mengdeoperasjonene kan igjen kombineres, slik som
$$\overline{A \cup B} = \{5, 6\},$$
og det finnes en haug regneregler, for eksempel
$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$
og
$$\overline{A \cap B} = \overline{A} \cup \overline{B}.$$
Jeg vet ikke om vi får bruk for dem, men skal fylle ut litt her om det trengs. Union og snitt av n utfall A_i, skriver vi
$$\bigcup_{i=1}^{n} A_i \qquad \text{og} \qquad \bigcap_{i=1}^{n} A_i.$$
Tegnet $\emptyset$ betyr 'den tomme mengden'. Dette er en mengde uten elementer.

Utfall

Vi bruker mengder til å definere utfallene i et tilfeldig forsøk. Vi triller en terning, og definerer utfallene
$$A = \{1, 2\}$$
og
$$B = \{1, 3, 4\}.$$
Vi sier at A inntreffer dersom terningen lander på 1 eller 2, og at B inntreffer dersom terningen lander på 1, 3 eller 4. Utfall sies å være disjunkte dersom de ikke kan kan inntreffe samtidig. La
$$C = \{3, 4\}.$$
Utfallene $A = \{1, 2\}$ og $B = \{1, 3, 4\}$ er ikke disjunkte, for hvis terningen lander på 1, inntreffer både A og B. Utfallene A og C er disjunkte. Vi skriver
$$A \cap C = \emptyset.$$

Utfallsrom

En mengde der alt som kan skje i det tilfeldige forsøket er representert ved disjunkte utfall, kalles et *utfallsrom*. Hvis vi triller en terning, er to eksempler på utfallsrom
$$\{1, 2, 3, 4, 5, 6\}$$
og
$$\{oddetall, partall\}.$$
Mengden
$$\{1, 2, 3, 4, 5\}$$
er ikke et utfallsrom, for terningen kan lande på 6, som ikke er tatt med, mens
$$\{oddetall, 1, 2, 3, 4, 5, 6\}.$$
er heller ikke et utfallsrom, for terningkast 1, 3 og 5 er representert ved to elementer i mengden.

Sannsynlighetsfunksjonen

Hvert utfall i vårt tilfeldige forsøk har en sannsynlighet for å inntreffe. *Sannsynlighetsfunksjonen* er en benevningsløs funksjon P som tilordner en sannsynlighet p til hvert utfall A
$$P(A) = p.$$
For en sannsynlighetsfunksjon setter vi opp tre krav. La S være et utfallsrom.

- $0 \leq P(A) \leq 1$ for alle $A \in S$
- $P(A \cap B) = P(A) + P(B)$ for alle $A, B \in S$
- $P(S) = P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) = 1$

Disse tre kravene sier at sannsynligheter må være mellom null og en, at de kan legges sammen, og at alle sannsynlighetene må summere til en (noe må skje).

Eksempel. Vi triller terningen. La $S = \{1, 2, 3, 4, 5, 6\}$. Da er
$$P(i) = 1/6$$
for alle i i S. For $A = \{1, 2\}$ og $B = \{1, 3, 4\}$ får vi
$$P(A) = P(1 \cup 2) = P(1) + P(2) = 2/6$$
og
$$P(B) = P(1 \cup 3 \cup 4) = P(1) + P(3) + P(4) = 3/6$$
Vi ser også at
$$P(A \cup B) = 4/6,$$
$$P(A \cap B) = 1/6$$
og
$$P(\overline{A}) = 4/6.$$
△

To regneregler for sannsynlighetsfunksjonen

Den første regneregelen er lett å bevise. La A være en hendelse. Det er temmelig innlysende at

$$S = A \cup \overline{A}$$

er et utfallsrom, for enten A eller $\overline{A}$ må inntreffe. Hvis P er en sannsynlighetsfunksjon, kan vi beregne

$$1 = P(S) = P(A \cup \overline{A}) = P(A) + P(\overline{A}),$$

og den første regneregelen er på plass,

$$P(\overline{A}) = 1 - P(A).$$

Dersom A og B ikke er disjunkte, finnes en variant av addisjonsegenskapen i definisjonen av sannsynlighetsfunksjonen, nemlig

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

Beviset for denne er litt mer komplisert. Det kan være jeg skriver det ut senere.

Eksempel. La igjen

$$A = \{1, 2\}$$

og

$$B = \{1, 3, 4\}.$$

Siden

$$A \cup B = \{1, 2, 3, 4\},$$

må $P(A \cup B) = 4/6$. Men hvis vi tar

$$P(A) + P(B) = 2/6 + 3/6 = 5/6.$$

får vi litt for høy sannsynlighet. Det er fordi utfallet $A \cap B = \{1\}$ telles to ganger, og vi må altså trekke det fra igjen for å få rett sannsynlighet,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 4/6.$$

△

Betinget sannsynlighet

I noen tilfeller vil sannsynligheten for en hendelse avhenge av hva som er skjedd tidligere. Vi definerer et nytt type utfall

$$A|B$$

som leses 'A gitt B'.

Eksempel. Hvis du trekker to kort fra en stokk, vil hvorvidt du trekker ruter på andre kort avhenge om du trakk ruter på første kort. Vi setter opp utfallene A = 'andre kort ruter', B = 'første kort ruter', og beregner

$$P(A|B) = \frac{12}{51}$$

og

$$P(A|\overline{B}) = \frac{13}{51}$$

△

I eksemplet over, fortalte sunt bondevett oss hva sannsynlighetene var. Men vi må nesten sette opp en presis definisjon for betinget sannsynlighet. Vi definerer

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

Av denne følger en ny regneregel som kalles Bayes lov. Siden $A \cap B = B \cap A$, følger at

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

eller

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

Eksempel. Her er en gammel eksamensoppgave fra Høgskolen i Bergen.

Oppgave 1

Vi skal se på forretninger med tyverialarm fra et bestemt alarmselskap. Følgende feil kan forekomme:

– Alarmen går selv om det ikke er innbrudd.

– Alarmen går ikke selv om det er innbrudd.

Erfaring over tid i selskapet viser at:

i) Det er 0,1 % sannsynlighet for at alarmen går selv om det ikke er innbrudd.

ii) I 1 % av alle innbrudd går ikke alarmen.

iii) Sannsynligheten for innbrudd en natt er 2 %.

La I være hendelsen at det er innbrudd i en butikk og A være hendelsen at alarmen går.

a) Finn $P(A\,|\,I)$, $P(\overline{I}\,|\,A)$ og $P(I)$ ut fra opplysningene gitt i oppgaveteksten.

b) Bestem sannsynligheten for at alarmen går.

c) Bestem sannsynligheten for at det ikke er innbrudd selv om alarmen går.

For å løse oppgaven må man være nøyaktig med å definere utfallene, og så bruke regnereglene. Vi definerer

$$A : \text{Alarmen går}$$

og

$$I : \text{Innbrudd}$$

Oppgaveteksten gir at $P(A|\overline{I}) = 0.001$, $P(\overline{A}|I) = 0.01$ og $P(I) = 0.02$. Nå er det et par andre sannsynligheter vi kan regne ut uten videre. Enten er det innbrudd, eller så er det ikke innbrudd:

$$P(\overline{I}) = 1 - P(I) = 0.98.$$

Hvis innbrudd vil alarmen enten gåa av eller ikke gå av:

$$P(A|I) = 1 - P(\overline{A}|I) = 0.99.$$

Når det ikke er innbrudd, vil alarmen enten gå av eller ikke gå av:

$$P(\overline{A}|\overline{I}) = 1 - P(A|\overline{I}) = 0.999.$$

Nå har vi funnet to av sannsynlighetene som etterspørres i oppgave a. Den tredje, $P(\overline{I}|A)$, kan vi ikke regne ut helt ennå, men den etterspørres også i oppgave c. (Denne eksamenen er nok laget i en fei. Det var mye greier på Høgskolen i Bergen.) Nå bruker vi et klassisk triks, nemlig å splitte utfallet A i to

$$A = (A \cap I) \cup (A \cap \overline{I})$$

Denne ligningen sier at dersom alarmen går (A), er det enten innbrudd ($A \cap I$), eller så er det ikke innbrudd ($A \cap \overline{I}$). Siden utfallene $A \cap \overline{I}$ og $A \cap I$ åpenbart er disjunkte (enten er det innbrudd, eller så er det ikke innbrudd), kan vi skrive

$$\begin{aligned} P(A) &= P(A \cap I) + P(A \cap \bar{I}) \\ &= P(I)P(A|I) + P(\bar{I})P(A|\bar{I}) \\ &= 0.02 \cdot 0.99 + 0.98 \cdot 0.001 \\ &= 0.02078. \end{aligned}$$

I oppgave c) skal vi ha tak i sannsynligheten $P(\bar{I}|A)$. Den kan vi nå regne ut ved Bayes lov

$$P(\bar{I}|A) = \frac{P(\bar{I})P(A|\bar{I})}{P(A)} = \frac{0.98 \cdot 0.001}{.02078} = 0.04716.$$

△

Uavhengighet

To utfall A og B sies å være uavhengige dersom

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

Sammenligner vi med definisjonen på betinget sannsynighet, ser vi at for uavhengige hendelser er

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

Dette forklarer hvorfor det kalles uavhengighet; sannsynligheten for at A inntreffer er den samme uavhengig av hvorvidt B har inntruffet.

Eksempel. Et enkelt eksempel på uavhengige utfall er å kaste en terning to ganger. Vi definerer

$$A_i = i \text{ øyne i kast nr. } 1$$

og

$$B_i = i \text{ øyne i kast nr. } 2$$

er utfallene A_i uavhengige utfallene B_i. Vi beregner sannsynligheten for først 2 og så 3

$$P(A_2 \cap B_3) = P(A_2)P(B_3) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

△

Litt kombinatorikk

Ordnet utvalg uten tilbakelegg: Først kaster vi en terning mange ganger. Når du kaster første gang, har du seks utfall. Kaster du engang til, har du seks utfall for hvert utfall i forrige kast, altså trettiseks. Kaster du enda en gang, har du seks nye utfall for hvert av de trettiseks forrige, altså 216. Hvis terningen har n sider, og du kaster r ganger, får du n^r utfall.
Ordnet utvalg med tilbakelegg: Sett at man skal velge et styre på tre mennesker fra en forsamling med 30 medlemmer. Når man først velger lederen, kan man velge blant alle de 30 medlemmene i forsamlingen. Nestleder kan så velges blant de 29 resterende medlemmene. For hvert valg av leder har man altså 29 muligheter for å velge nestleder. Det blir $30 \cdot 29$ mulige kombinasjoner. For hver av disse $30 \cdot 29$ kombinasjonene har man så 28 muligheter til å velge sekretær. Så da er vi oppe i $30 \cdot 29 \cdot 28$ mulige kombinasjoner. Dersom man skal velge et styre på 6 mennesker av en forsamling på 34, har man

$$34 \cdot 33 \cdot 32 \cdot 31 \cdot 30 \cdot 29 = \frac{34!}{28!} = \frac{34!}{(34-6)!}$$

forskjellige kombinasjoner. Husk at

$$n! = n \cdot (n-1) \cdots 4 \cdot 3 \cdot 2 \cdot 1.$$

Uttales 'n fakultet'.
Uordnet utvalg uten tilbakelegg: I forrige avsnitt hadde rekkefølgen på uttrekket noe å si, siden alle styrer med de samme menneskene i forskjellige verv ble telt med. Vi omdanner nå de seks personene i avsnittet til lottokuler, slik at rekkefølgen på uttrekket ikke har noe å si; alle vet at lottokuler trekkes ut og organiseres i stigende rekkefølge til slutt. Seks nummererte kuler kan ordnes i rekkefølge på $6! = 720$ forskjellige måter. Dersom vi ikke er interessert i rekkefølgen, må denne interne organiseringen deles ut, og det finnes derfor

$$\frac{30 \cdot 29 \cdot 28}{3 \cdot 2 \cdot 1} = \frac{30!}{27! \cdot 3!}$$

forskjellige utvalg. Dette tallet kalles **binomialkoeffisient**, og har en spesiell notasjon. Vi definerer

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

som altså er antall måter å plukke ut r elementer av en samling på n elementer når rekkefølgen ikke spiller noen rolle.