Introduksjonsforelesning

Hendelser og sannsynlighetsmål

D | Sannsynlighetsrom | $(\Omega, \mathcal{E}, \mathbb{P})$ definerer et (stokastisk) eksperiment:

- Ω er mengden av alle enkeltutfall i eksperimentet.
- $\mathcal{E}$ er mengden av hendelser. Inntil videre holder det å tenke på hendelser som delmengder av Ω, altså null eller flere av enkeltutfallene.[1]
- D | Sannsynlighetsmål | $\mathbb{P}$ gitt ved *sannsynlighetsaksiomene*:
  1. $\mathbb{P} : \mathcal{E} \to [0, 1]$
     Sannsynlighetsmålet forteller oss sannsynligheten for en hendelse, og denne må være minst null (inntreffer aldri) og maksimalt én (inntreffer alltid).
  2. $\mathbb{P}(\Omega) = 1$
     Det skjer alltid *noe*.
  3. $\mathbb{P}(E_1 \cup \cdots \cup E_n) = \mathbb{P}(E_1) + \cdots + \mathbb{P}(E_n)$ for alle hendelser $E_1, \ldots, E_n$ som ikke har felles elementer (er *disjunkte*).
     Om *enten* E_1 *eller* E_2 kan inntreffe, er sannsynligheten for at én av dem inntreffer lik summen av de to sannsynlighetene.

D | Komplement | $A^\complement$ til en hendelse A oppfyller $\mathbb{P}(A^\complement) = 1 - \mathbb{P}(A)$

Uavhengige hendelser

D | Uavhengighet | $A \perp B$ hvis $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

Tilsvarende er flere mengder uavhengige hvis $\mathbb{P}(E_1 \cap \cdots \cap E_n) = \mathbb{P}(E_1) \cdot \cdots \cdot \mathbb{P}(E_n)$.

Z | Det er **ikke** tilstrekkelig at mengdene er *parvis* uavhengige. | Z

E | To mynter

- $\Omega = \{\mathrm{KK}, \mathrm{KM}, \mathrm{MK}, \mathrm{MM}\}$
- $\mathcal{E} = \mathcal{P}(\Omega)$[2]. Vi gir tre mengder spesielle navn:
  - $A = \{\mathrm{KK}, \mathrm{KM}\}$, mengden av hendelser der første kast er kron
  - $B = \{\mathrm{KK}, \mathrm{MK}\}$, mengden av hendelser der andre kast er kron
  - $C = \{\mathrm{KK}, \mathrm{MM}\}$, mengden av hendelser der de to kastene er like

Er A, B og C uavhengige? Parvis uavhengige?

Betinget sannsynlighet

D | Betinget sannsynlighet | $\mathbb{P}(A \mid B)$ $= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

Med andre ord er $\mathbb{P}(A \mid B)$ sannsynligheten for at A inntreffer, gitt at B inntreffer.

Stokastiske variabler

Ofte bryr vi oss mer om funksjoner av utfall enn utfallene i seg selv. For eksempel kan et utfall $\omega \in \Omega$ være gitt ved en fullstendig beskrivelse av alle luftmolekylene i atmosfæren over Kjeller en gitt dag. Dette er litt vanskelig å forholde seg til, så vi lager oss en "gjennomsnittlig bevegelsesenergi-funksjon" T og kaller det den spytter ut for temperatur $T(\omega)$ i $^\circ\mathrm{C}$ (ish). Når vi skriver $\mathbb{P}(T \leq 20)$, snakker vi egentlig om $\mathbb{P}(\{\omega : T(\omega) \leq 20\})$[3].

Det er viktig å merke seg at det ikke er noe stokastisk med variabelen (funksjonen) *i seg selv*. All tilfeldigheten "bor" i det underliggende utfallsrommet, og stokastisiteten i verdiene variabelen tar er simpelthen et resultat av dette.

Den formelle definisjonen er kronglete, så vi nøyer oss med å si at verdiene som funksjonen tar, ofte er intervaller av reelle tall, mengder av heltall, eller kategorier som ikke lar seg ordne (f.eks. rød/gul/grønn eller ja/nei).

Om vi kan definere en fornuftig ordning på verdiene, er det ofte nyttig å snakke om sannsynligheten for at variabelen er mindre enn et visst tall:

D | Kumulativ fordelingsfunksjon | $F_X(x)$ $= \mathbb{P}(X \leq x)$

E | Egenskapene til F_X Hva skjer med F_X når $X \to \pm\infty$? Hva kan vi si om $F_X(x)$ og $F_Y(y)$ når $x < y$?

I de to siste tilfellene sier vi at variabelen er *diskret*, og vi kan definere en *massefunksjon*:

D | Sannsynlighetsmassefunksjon | $p_X(x)$ $= \mathbb{P}(X = x)$

Om du husker at $\mathbb{P}(X = x) = \mathbb{P}(\{\omega : X(\omega) = x\})$ og tar en titt på sannsynlighetsaksiomene over, er det lett og se at p_X har følgende egenskaper:

- $p_X(x) \in [0, 1]$
- $\sum\limits_{x \in A} p_X(x) = \mathbb{P}(X \in A)$, der A er en mengde av verdier X kan ta
  Ulike verdier representerer disjunkte hendelser - $T(\omega)$ kan ikke ha flere verdier samtidig (temperaturen kan ikke være både 20 og 25 grader på samme tid!)
- $\sum\limits_{x \in \mathrm{ran(X)}} p_X(x) = 1$, der $ran(X)$ er alle mulige verdier X kan ta

Hvis X tar verdier i $\mathbb{R}$ og F_X er kontinuerlig, sier vi at X er kontinuerlig. Da gir det ikke lenger mening å snakke om sannsynlighetsmasse (hva er sannsynligheten for at temperaturen er *nøyaktig* 20 grader?)

Vi introduserer derfor konseptet *tetthet*:

D | Sannsynlighetstetthetsfunksjon (pdf) | $f_x(x)$ slik at $F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$

[4]

Det følger at $\mathbb{P}(a < X \leq b) = \int_{a}^{b} f_X(u)\, du$.

E | Egenskapene til f_X Hvilke verdier kan f_X ta? Kan f_x være negativ (på et intervall)? Kan f_x større enn 1?

Simultan sannsynlighetsfordeling og uavhengige stokastiske variabler

D | Simultan massefunksjon | $p_{X,Y}(x,y)$ $= \mathbb{P}(X = x \text{ og } Y = y) = \mathbb{P}(\{\omega : X(\omega) = x \text{ og } Y(\omega) = y\})$

Definisjonen av uavhengighet for diskrete stokastiske variabler følger direkte av den for hendelser over.

D | Simultan tetthetsfunksjon | $f_{X,Y}(x,y)$ slik at $F_{X,Y}(x,y) = \mathbb{P}(X \leq x \text{ og } Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\, du\, dv$

På samme måte har vi at $X \perp Y \iff f_{X,Y}(x,y) = f_X(x) f_Y(y)$.

Betinget tetthet

D | Betinget massefunksjon | $p_{Y|X}(y)$ $= \mathbb{P}(Y = y \mid X = x)$[5]

D | Betinget tetthet | $f_{Y|X}(y)$ $= \frac{f_{X,Y}(x,y)}{f_X(x)}$

For uavhengige variabler er den betingede massefunksjonen (tetthetsfunksjonen) lik den marginale massefunksjonen (tetthetsfunksjonen): $p_{y|X} = p_Y$, $f_{Y|X} = f_Y$.

Bayes' teorem

T | Bayes (hendelser) $\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}$

T | Bayes (tettheter) $f_{Y|X=x}(y) = \frac{f_{X|Y=y}(x) f_Y(y)}{f_Y(y)}$

E | Bevis Bayes' teorem (følger fra definisjonen av betinget sannsynlighet)

Momenter

D | Forventning (diskret X) | $\mu = \mathbb{E}[X]$ $= \sum_{\text{alle } x} x \cdot p_X(x)$

D | Forventning (kontinuerlig X) | $\mu = \mathbb{E}[X]$ $= \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$

Vi kan tolke forventningen som gjennomsnittet av verdiene X tar, vektet etter hvor ofte de inntreffer.

D | Varians | $\mathbb{V}[X]$ $= \mathbb{E}[(X - \mu)^2]$

Tolkning: Vi finner avstanden fra hver av verdiene X kan ta til forventningen (gjennomsnittet). Så tar vi forventningen igjen for å finne den gjennomsnittlige avstanden (igjen vektet etter hvor ofte verdiene inntreffer). Vi kvadrerer slik at positive og negative avstander ikke skal nulle hverandre ut.

E | Vis at $\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

E | Vis at $\mathbb{E}[a \cdot X + b \cdot Y] = a \cdot \mathbb{E}[X] + b \cdot \mathbb{E}[Y]$ der X og Y er stokastiske variabler og a og b er reelle tall

E | Vis at $\mathbb{V}[a \cdot X + b \cdot Y] = a^2 \cdot \mathbb{V}[X] + b^2 \cdot \mathbb{V}[Y]$ der X og Y er **uavhengige** stokastiske variabler og a og b er reelle tall. Hvorfor krever vi $X \perp Y$?

E | $X \sim \text{Beta}(\alpha, \beta)$. Finn $\mathbb{E}[X]$ og $\mathbb{V}[X]$. [6]

> D | $\text{Beta}(\alpha, \beta)$ har tetthet $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$[7]
>
> Hint: Flytt det som ikke avhenger av x utenfor integralet. Skriv $a = a + 1 - 1$.

Transformasjoner av stokastiske variabler

La den stokastiske variabelen Y være gitt ved en funksjon g[8] av en annen stokastisk variabel X, altså $Y = g(X)$.

T | Transformasjon av kontinuerlig stokastisk variabel $f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dx}{dy} \right|$

E | $Y = e^X$

> $g^{-1}(y) = \ln(y)$
>
> $\left| \frac{dx}{dy} \right| = \left| \frac{dg^{-1}(y)}{dy} \right| = \left| \frac{1}{y} \right| = \frac{1}{y}$
>
> $f_Y(y) = f_X(\ln(y)) \cdot \frac{1}{y}$

E | La $X \sim \mathcal{N}(0,1)$. Vis at $f_Y(y) = \frac{1}{y\sqrt{2\pi}} \exp\left(-\frac{\ln(y)^2}{2}\right)$.

---

1. Dette er ikke helt korrekt, og fungerer ikke når vi møter på mindre "hyggelige" mengder enn de vi tar for oss nå. Da må vi introdusere konseptet *målbarhet*, som er til å få vondt i hodet av. ↩
2. $\mathcal{P}(\Omega)$ er mengden av alle delmengder av Ω, og inneholder $2^4 = 16$ mengder (sjekk selv!). ↩
3. Kolonet betyr "slik at". ↩
4. Lenge siden du har sett et integral? Frykt ikke! Tenkt på dette som at vi begynner med den laveste mulige verdien variabelen kan ta (f.eks. $T = -273.15$) og ser på hvilken sannsynlighets*tetthet* vi har ved denne verdien. Så øker vi gradvis verdien av variabelen til vi når, og "summerer opp" alle tetthetene. Til slutt ender vi opp med en *sannsynlighet* (f.eks. for hendelsen $T < 20$, med $\mathbb{P}(T < 20) = F_T(20) = \int\limits_{-\infty}^{20} f_T(u)\, du$). For ordens skyld er det verdt å nevne at vi også krever at F er deriverbar. ↩
5. Husk at $X = x$ viser til hendelsen $\{\, \omega : X(\omega) = x \,\}$ ↩
6. $\sim$ betyr "følger" (som i "X følger en betafordeling") ↩
7. Gammafunksjonen Γ dukker opp nærmest overalt i matematikken. Siden vi kun integrerer over x i forventningen, trenger du ikke å forholde deg til hvordan den er definert. ↩
8. g må være deriverbar og strengt stigende eller strengt synkende. Ser du hvorfor? ↩