FORELESNINGSNOTATER OG OPPGAVER STK1100 – SANNSYNLIGHETSREGNING OG STATISTISK MODELLERING, VÅREN 2026

27. mars 2026

EMIL A. STOLTENBERG

Sammendrag. Dette er forelesningsnotater og oppgaver til emnet STK1100 gitt ved Matematisk institutt, Universitetet i Oslo. Notatene følger boken Devore et al. (2021) ganske tett, og vil bli oppdatert etterhvert som vi skrider fremover. Takk til Welid Alebady og Penelope Sørbø for å gjøre meg oppmerksom på feil og mangler. Send meg gjerne en epost når du oppdager feil, mangler, skrivefeil, etc.

Innhold

1. Eksperimenter, utfallsrom og hendelser 2
2. Sannsynlighetsmål 5
3. Betinget sannsynlighet 7
4. Uavhengige hendelser 9
5. Tilfeldige variable og fordelinger 12
6. Forventing og varians 16
7. Konvergens i sannsynlighet 20
8. Den bivariate normalfordeligen (og tilfeldige vektorer) 23
9. Oppgaver 24
Referanser 25

1. Eksperimenter, utfallsrom og hendelser

Dette er notater delvis basert på Kapittel. 2.1 i Devore et al. (2021). Du bør lese dette kapittelet før du leser videre. I det minste, gjør deg kjent med definisjonene av union, snitt, og komplement.

1.1. **Eksperimenter.** Vi kan være usikre på hva som skal skje, på hva som er tilfellet akkurat nå, og på hva som har skjedd. Tenk på et *eksperiment* om enhver situasjon der vi er usikre på noe. Spinner vi et kronestykke opp i lufta er vi usikre på om det vil vise kron eller mynt. En lege gir en pasient en ny type medisin og er usikker på om den vil være lindrende, ikke ha noen effekt, eller gjøre ting værre. Teller vi antall kast med en terning til vi får vår første sekser er vi usikre på hvor mange ganger vi må kaste terningen. En meteorolog er usikker på hvor mye regn det vil falle i morgen. Meterologen er også usikker på differansen mellom regnfallet i morgen og dagen etter i morgen. En kvinner lurer på om hun er gravid. En historiker lurer på hvor mange av de som levde i Norge i 1348 fortsatt var i live i 1351. En dommer er usikker på om den tiltalte er skyldig eller ikke.

1.2. **Utfallsrom.** Et utfallsrom er en mengde som inneholder alle mulige utfall av et eksperiment. Vi bruker typisk symbolet Ω ('store omega') for utfallsrommet. For myntkastingen er utfallene kron eller mynt, så utfallsrommet er

$$\Omega = \{\text{kron}, \text{mynt}\}.$$

For testingen av den nye medisinen er utfallsrommet

$$\Omega = \{\text{lindrende}, \text{ingen effekt}, \text{værre}\}.$$

Utfallsrommet for antallet ganger vi må kaste en terning til vi får vår første sekser er

$$\Omega = \{1, 2, \ldots\}.$$

For mengden regn i morgen, (gitt av vi kan måle det helt presist), kan et naturlig utfallsrom være

$$\Omega = [0, \infty) = \{\omega \in \mathbb{R} \colon \omega \geq 0\},$$

og for differansen mellom regnet i morgen og i overimorgen er et naturlig utfallsrom

$$\Omega = \mathbb{R}$$

Et mulig utfall er et element i utfallsrommet. Dersom vi kaster en terning gjennomfører vi et eksperiment med seks mulige utfall, utfallsrommet er $\Omega = \{1, 2, 3, 4, 5, 6\}$, og $1, 2, 3, 4, 5, 6$ er utfall. Dersom vi kaster to terninger, en blå og en rød, har utfallsrommet $6 \times 6 = 36$ mulige utfall på formen $\omega = (\text{blå terning}, \text{rød terning})$. Utfallsrommet er

$$\Omega = \{(1, 1), (1, 2), \ldots, (6, 5), (6, 6)\}.$$

og $(6, 6)$ er et mulig utfall.

1.3. **Hendelser.** La Ω være et utfallsrom. En *hendelse* er en delmengde $A \subset \Omega$, slik at vi under betingelsene til eksperimentet, kan avgjøre om A intraff eller ikke. At en delmengde A inntreffer betyr at utfallet $\omega \in A$. Tenk på eksperimentet der vennen din kaster to kronerstykker, et farget blått og et farget rødt, slik at utfallsrommet er

$$\Omega = \{(\text{kron}, \text{kron}), (\text{kron}, \text{mynt}), (\text{mynt}, \text{kron}), (\text{mynt}, \text{mynt})\},$$

men der vennen din kun forteller deg utfallet til det blå kronestykket. Dersom $\omega = (\omega_1, \omega_2)$ er et utfall betyr dette at vennen din kun forteller deg ω_1, og holder ω_2 skjult. På bakgrunn av et slikt eksperiment vil vi kunne vite om

$$A = \{\text{blå kron}\} = \{(\text{kron}, \text{kron}), (\text{kron}, \text{mynt})\}$$

intraff, og om

$$A^c = \{\text{blå mynt}\} = \{(\text{mynt}, \text{kron}), (\text{mynt}, \text{mynt})\},$$

intraff, og her er dette all interessant informasjon eksperimentet gir. Altså er delmengden A en hendelse, og A^c også en hendelse. Den ikke så interessante informasjon eksperimentet gir, er at Ω intraff (det gjør det alltid) og at $\emptyset$ ikke intraff (det gjør det aldri). Så Ω og $\emptyset$ er også hendelser. Samler vi sammen alle hendelsene i eksperimentet der vennen din kaster to kronestykker men skjuler utfallet til det rød kronestykket, får vi familien av hendelser gitt ved

$$\mathcal{A} = \{\emptyset, \Omega, A, A^c\}.$$

Nå, se på delmengden

$$B = \{\text{blå og rød kron}\} = \{(\text{kron}, \text{kron})\}$$

Siden vi kun har informasjon om det blå kronestykket, kan vi ikke vite om B intraff, uansett utfall. Dette betyr at i eksperimentet der vi kun har informasjon om det blå kronestykket er ikke B en hendelse. At B ikke er en hendelse ser vi også fra at B ikke er i $\mathcal{A}$, som jo er familien av hendelser.

La nå A og B være to delmengder av et eller annet utfallsrom Ω. Hvis vi kan avgjøre om A intreffer, og vi kan avgjøre om B intreffer, så kan vi avgjøre om A *eller* B intreffer, altså om $A \cup B$ intreffer. Sagt på en annen måte, dersom A og B er hendelser, så er $A \cup B$ en hendelse. Og slik kan vi fortsette.

σ-algebra

Motivert av dette skal vi samle alle delmengdene av Ω som vi kaller hendelser i en familie $\mathcal{A}$ (som ikke er tom), og kreve at denne familien av delmengder er slik at

(i) hvis $A \in \mathcal{A}$, så er $A^c \in \mathcal{A}$; og
(ii) hvis $A_1, A_2, \ldots \in \mathcal{A}$, så er $\cup_{n\geq 1} A_n \in \mathcal{A}$.

Her er $\cup_{n\geq 1} A_n = A_1 \cup A_2 \cup \cdots$. Vi sier at $\mathcal{A}$ er lukket under komplementer og tellbare unioner. En slik familie av delmengder kalles en σ-algebra. Merk at fra (i) og (ii) følger det at $\Omega \in \mathcal{A}$ og at den tomme mengden $\emptyset = \Omega^c \in \mathcal{A}$ (vis det). For et gitt utfallsrom Ω utstyrt med en σ-algebra $\mathcal{A}$, definerer vi delmengdene inneholdt i $\mathcal{A}$ som *hendelser*: En delmengde $A \in \mathcal{A}$ er en hendelse.

Def. hendelse

For myntkastet har vi et utfallsrom $\Omega = \{\text{kron}, \text{mynt}\}$. Da kan vi sjekke at familien av hendelser

$$\mathcal{A} = \{\emptyset, \Omega, \{\text{kron}\}, \{\text{mynt}\}\},$$

er en σ-algebra. For en fotballkamp er utfallsrommet $\Omega = \{h, u, b\}$, og en σ-algebra på Ω er

$$\mathcal{B} = \{\emptyset, \Omega, \{h\}, \{u\}, \{b\}, \{h, u\}, \{h, b\}, \{u, b\}\}.$$

Familien $\mathcal{F}$ av delmengder av $\Omega = \{h, u, b\}$ gitt ved

$$\mathcal{C} = \{\emptyset, \Omega, \{h\}, \{u, b\}\},$$

er også en σ-algebra på Ω, men det er σ-algebra som kun gir oss informasjon om hjemmelaget vant eller ikke (ikke om det ble uavgjort, f.eks.). Dette betyr at $\{b\}$ er en hendelse når vi arbeider med $\mathcal{B}$, men ikke dersom vi arbeider med $\mathcal{C}$.

Fra de tre σ-algebraene $\mathcal{A}, \mathcal{B}, \mathcal{C}$ ser vi at kravet om at de er lukket under uendelig tellbare unioner er å ta litt hardt i, fordi $\mathcal{A}$ bare har fire medlemmer, $\mathcal{B}$ har åtte, og $\mathcal{C}$ har fire. Generelt, for et endelig utfallsrom Ω med k elementer, kan ingen σ-algebra på Ω inneholde mer enn 2^k delmengder av Ω. Potensmengden

$$2^\Omega = \{\text{alle mulige delmengder av } \Omega\},$$

inneholder $2^{|\Omega|}$ elementer, der $|\Omega| =$ antall elementer i Ω. På et endelig utfallsrom holder det derfor å kreve at $\mathcal{A}$ er en *algebra* av hendelser. Familien $\mathcal{A}$ av delmengder av Ω er en algebra på Ω dersom den har medlemmer og er slik at

algebra av hendelser

(i)′ hvis $A \in \mathcal{A}$, så er $A^c \in \mathcal{A}$; og
(ii)′ hvis $A, B \in \mathcal{A}$, så er $A \cup B \in \mathcal{A}$.

Igjen, fra (i)′ og (ii)′ følger det at Ω og $\emptyset$ er i $\mathcal{A}$. Du kan sjekke at en σ-algebra alltid er en algebra, og at på endelige utfallsrom er en algebra en σ-algebra, så derfor holder det å snakke om σ-algebraer.

1.4. **Oppgaver.**

(a) Overbevis deg selv om at for alle mengder A, B, C, har vi

(i) $A \cup B = B \cup A$, og $A \cap B = B \cap A$;
(ii) $A \cup (B \cup C) = (A \cup B) \cup C$, og $A \cap (B \cap C) = (A \cap B) \cap C$;
(iii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, og $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Dette er henholdvis den kommutative, den assosiative, og den distributive egenskapen til union og snitt.

Oppgave 1.1 La Ω være en mengde, og la A og B være delmengder av Ω.

(a) Tegn Venn-diagrammer eller finn andre argumenter som overbeviser deg om det følgende

(i) $\Omega^c = \emptyset$;
(ii) $\Omega \cap A = A$;
(iii) $A \cup \emptyset = A$;
(iv) $A \cap \emptyset = \emptyset$;
(v) $A \cap A^c = \emptyset$;
(vi) $A = (A \cap B) \cup (A \cap B^c)$;
(vii) $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$;
(viii) $A \subset (A \cup B)$, og $B \subset (A \cup B)$;
(ix) $(A \cap B) \subset A$, og $(A \cap B) \subset B$;
(x) Hvis $A \subset B$ så er $B^c \subset A^c$;
(xi) $\emptyset \subset A$.

(b) De Morgans lover sier at

De Morgans

$$(A \cup B)^c = A^c \cap B^c;$$
$$(A \cap B)^c = A^c \cup B^c.$$

Bevis dem. Gjerne ved hjelp av Venn-diagrammer.

Oppgave 1.2 La $B, A_1, A_2, \ldots$ være delmengder av en mengde Ω. Husk definisjonene

$$\cup_{n\geq 1} A_n = \cup_{n=1}^\infty A_n = A_1 \cup A_2 \cup \cdots = \{\omega \in \Omega \colon \omega \in A_n \text{ for minst en } n \geq 1\},$$
$$\cap_{n\geq 1} A_n = \cap_{n=1}^\infty A_n = A_1 \cap A_2 \cap \cdots = \{\omega \in \Omega \colon \omega \in A_n \text{ for alle } n \geq 1\},$$

med den endelige unionen $\cup_{n=1}^{m} A_n = A_1 \cup \cdots \cup A_m$ og det endelige snittet $\cap_{n=1}^{m} A_n = A_1 \cap \cdots \cap A_m$ definert tilsvarende.

(a) Vis at

$$\cup_{i\geq 1} A_i = A_1 \cup (\cup_{i\geq 2} A_i), \quad \text{og} \quad \cap_{i\geq 1} A_i = A_1 \cap (\cap_{i\geq 2} A_i),$$

og at det her ikke er noe spesielt med A_1.

(b) Vis at

$$B \cap (\cup_{n\geq 1} A_n) = \cup_{n\geq 1}(B \cap A_n), \quad \text{og} \quad B \cup (\cap_{n\geq 1} A_n) = \cap_{n\geq 1}(B \cup A_n).$$

(c) Anta at $\cup_{n\geq 1} A_n = \Omega$, la $B \subset \Omega$, og vis at

$$B = \cup_{n\geq 1}(B \cap A_n).$$

Anta nå at $A_1, \ldots, A_n$ er disjunkte, det vil si $A_i \cap A_j = \emptyset$ for alle $i \neq j$. Vis at da er $(B \cap A_i)$ og $(B \cap A_j)$ også disjunkte for alle $i \neq j$. En disjunkt samling mengder $A_1, A_2, \ldots$ hvis union utgjør Ω, kalles en partisjon av Ω.

(d) Utvid De Morgans lover fra to til uendelige mange mengder: Vis at

$$(\cup_{n\geq 1} A_n)^c = \cap_{n\geq 1} A_n^c, \quad \text{og} \quad (\cap_{n\geq 1} A_n)^c = \cup_{n\geq 1} A_n^c.$$

(e) Forklar at de egenskapene du nå har utledet for uendelige unioner og snitt, også holder for endelige unioner og snitt.

Oppgave 1.3 Vi slår kron-eller-mynt med tre kronestykker samtidig.

(a) Skriv ned utfallsrommet Ω.

(b) Hvor mange delmengder er det i potensmengden 2^Ω på Ω? Du skal slippe å skrive dem ned.

(c) Skriv ut delmengden $A = \{\text{tre like}\}$.

(d) Skriv ned den minste algebraen som inneholder A.

Oppgave 1.4 La $A_1, \ldots, A_n, B$ være mengder i en algebra $\mathcal{A}$ på Ω.

(a) Vis at $\cup_{i=1}^{n} A_i \in \mathcal{A}$.

(b) Vis også mengder som $\cap_{i=1}^{n} A_i^c$, og $B \setminus (\cup_{i=1}^{n} A_i)$, og så videre, er i $\mathcal{A}$.

Oppgave 1.5 La Ω være et utfallsrom.

(a) La $\mathcal{A}_0 = \{A_1, \ldots, A_n\}$ være en partisjon av Ω. Betrakt familien av mengder definert ved

$$\mathcal{A} = \{\emptyset, \text{alle unioner av mengder in } \mathcal{A}_0\}.$$

Vis at $\mathcal{A}$ er en algebra. Vi sier at $\mathcal{A}$ er algebraen generert av $\mathcal{A}_0$.

(b) La $\mathcal{B}$ være en algebra på Ω slik at $\mathcal{A}_0 \subset \mathcal{B}$. Vis at $\mathcal{A} \subset \mathcal{B}$, altså at algebraen generert av $\mathcal{A}_0$ den minste algebraen som inneholder $\mathcal{A}_0$.

2. Sannsynlighetsmål

Vi har et ufallsrom Ω og en σ-algebra av hendelser $\mathcal{F}$ på dette utfallsrommet. Sagt på en tørr(?) måte, så handler den matematiske teorien om sannsynlighet om å måle den relative størrelsen til hendelser.

2.1. **Kolmogorovs aksiomer.** Et sannsynlighetsmål er en funksjon $\Pr(\cdot)$ med domene $\mathcal{F}$ som antar reelle verdier,

$$\Pr\colon \mathcal{F} \to \mathbb{R},$$

Kolmogorovs aksiomer

og tilfredsstiller Kolmogorovs aksiomer, det vil si de tre følgende egenskapene

(i) $\Pr(A) \geq 0$ for alle hendelser A;
(ii) $\Pr(\Omega) = 1$;
(iii) $\Pr(\cup_{n\geq 1} A_n) = \sum_{n\geq 1} \Pr(A_n)$ for disjunkte hendelser $A_1, A_2, \ldots$, det vil si $A_i \cap A_j = \emptyset$ når $i \neq j$.

Den tredje egenskapen kalles tellbar additivitet. Setter vi samme utfallsrommet Ω, σ-algebraen $\mathcal{F}$ av hendelser, og et sannsynlighetsmål Pr definert på $\mathcal{F}$, får vi sannsynlighetsrommet $(\Omega, \mathcal{A}, \Pr)$. Et slikt sannsynlighetsrom er en matematisk modell for å si noe om tilfeldighet eller usikkerhet. Det er den eneste slike modellen vi forholder oss til i dette emnet (og for den saks skyld, i alle andre emner i sannsynlighet og statistikk jeg vet om på Blindern). Dersom utfallsrommet vårt er endelig, $\Omega = \{\omega_1, \ldots, \omega_k\}$, så σ-algebraen $\mathcal{F}$ kun innholder endelig mange delmengder, blir kravet i (iii) til et krav om endelig additivitet

sannsynlighetsrom

(iii)′ $\Pr(A \cup B) = \Pr(A) + \Pr(B)$ for alle disjunkte $A, B \in \mathcal{F}$.

Så dersom vi jobber på et endelig utfallsrom med en algebra $\mathcal{A}$ av hendelser, så er en funksjon $Q\colon \mathcal{A} \to \mathbb{R}$ som tilfredsstiller (i), (ii), og (iii)′ et sannsynlighetsmål.

At sannsynlighetsmål er tellbart additive gir dem visse kontinuitetsegenskaper, kalt kontinuitet av sannsynlighetsmål: La $A_1, A_2, \ldots$ være en voksende følge av hendelser, det vil si $A_1 \subset A_2 \subset \cdots$. Da er

kontinuitet av sannsynlighetsmål

$$\Pr(\cup_{n\geq 1} A_n) = \lim_n \Pr(A_n),$$

Tilsvarende, dersom $B_1, B_2, \ldots$ er en avtagende følge av hendelser, $B_1 \supset B_2 \supset \cdots$, så er

$$\Pr(\cap_{n\geq 1} B_n) = \lim_n \Pr(B_n).$$

Vi beviser den første av disse for å se hvordan tellbar additivitet brukes. La $A_1, A_2, \ldots$ være en voksende følge mengder. Sett $A_0 = \emptyset$ og definer

$$D_n = A_n \setminus A_{n-1}, \quad \text{for } n \geq 1.$$

Som en oppgave, forklar hvorfor D_n er en hendelse; at D_n og D_m er disjunkte for alle $n \neq m$; og at $\cup_{n\geq 1} A_n = \cup_{n\geq 1} D_n$. Bruker vi disse tingene kombinert med tellbar additivitet finner vi at

$$\begin{aligned}\Pr(\cup_{n\geq} A_n) = \Pr(\cup_{n\geq} D_n) &= \sum_{n\geq 1} \Pr(D_n) \\ &= \lim_n \sum_{j=1}^{n} \Pr(D_j) = \lim_n \sum_{j=1}^{n} \{\Pr(A_j) - \Pr(A_{j-1})\} = \lim_n \Pr(A_n).\end{aligned}$$

Den andre likhetene bruker tellbar additivitet, den fjerde bruker en egenskap du skal bevise i Oppgave 2.1(**e**) (og at $A_{j-1} \subset A_j$), mens den siste likheten bruker at til venstre for denne har vi en teleskoperende sum.

2.2. **Oppgaver.**

Oppgave 2.1 La $A, A_1, \ldots, A_n$ og B være hendelser, og Pr et sannsynlighetsmål. Vis at

(a) $\Pr(\emptyset) = 0$;

(b) $\Pr(\cup_{i=1}^n A_i) = \sum_{i=1}^n \Pr(A_i)$, for disjunkte $A_1, \ldots, A_n$;

(c) $\Pr(A) = 1 - \Pr(A^c)$;

(d) $\Pr(A) \leq 1$;

(e) $\Pr(B \setminus A) = \Pr(B) - \Pr(A \cap B)$;

(f) $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$, så $\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$;

(g) Hvis $A \subset B$ så er $\Pr(A) \leq \Pr(B)$.

Oppgave 2.2 I denne oppgaven tar vi egenskapen fra Oppgave 2.1**(f)** litt videre.

(a) Vis at for alle hendelser $A_1, \ldots, A_m$ er $\Pr(\cup_{j=1}^m A_j) \leq \sum_{j=1}^m \Pr(A_j)$.

(b) Vis at dette også holder for uendelig mange hendelser $A_1, A_2, \ldots$, det vil si

$$\Pr(\cup_{n\geq 1} A_n) \leq \sum_{n\geq 1} \Pr(A_n),$$

som betyr at sannsynlighetsmål er tellbart subadditive.

Oppgave 2.3 La $\Omega = \{1, 2, 3, 4, 5, 6\}$ være utfallsrommet forbundet med et terningkast.

(a) Definer funksjonen

$$Q_1(A) = \frac{|A|}{6}, \quad \text{for alle } A \in 2^\Omega,$$

der $|A|$ = antall elementer i A. Vis at Q_1 er et sannsynlighetsmål på 2^Ω. Sannsynlighetsrommet $(\Omega, 2^\Omega, Q_1)$ er det forbundet med å kaste en rettferdig terning.

(b) Definer funksjonen

$$Q_2(A) = \frac{1}{21} \sum_{\omega \in A} \omega, \quad \text{for alle } A \in 2^\Omega,$$

Vis at Q_2 er et sannsynlighetsmål.

(c) Forsøk å konstruer en funksjon W på 2^Ω slik at $W(A) \geq 0$ for alle $A \in 2^\Omega$ og $W(\Omega) = 1$, men som ikke er et sannsynlighetsmål.

(d) La $\Omega = \{\omega_1, \ldots, \omega_m\}$ være et utfallsrom, og $\mathcal{A}$ en algebra av hendelser på Ω. La $p_1, \ldots, p_m$ være tall in $[0, 1]$ slik at $\sum_{j=1}^m p_j = 1$. Vis at $Q_3(A) = \sum_{j:\omega_j \in A} p_j$ for $A \in \mathcal{A}$ er et sannsynlighhetsmål.

3. Betinget sannsynlighet

Dette er notater knyttet til Kap. 2.4 i Devore et al. (2021). Sannsynligheten vi gir en hendelse kan være avhengig av hva vi vet om andre hendelser.

3.1. **Definisjonen.** Vi har et sannsynlighetsrom $(\Omega, \mathcal{F}, \Pr)$ (vi har alltid et sannsynlighetsrom, og som regel vil vi ikke nevne det i det hele tatt). La A og B være to hendelser, med $\Pr(B) > 0$. Da er den betingende sannsynligheten for A gitt B definert ved

$$\Pr(A \,|\, B) = \frac{\Pr(A \cap B)}{\Pr(B)}. \tag{1}$$

I tilfeller der $\Pr(B) = 0$, er $\Pr(A \mid B)$ ikke definert. Hvis du tegner noen Venn-diagrammer er denne definisjonen ganske intuitiv: Sannsynligheten for A gitt B er størrelsen (målt ved Pr) til den delen av A som overlapper B, normalisert ved størrelsen til B.

Definisjonen i (1) fungerer for enhver hendelse A, så vi kan definere en funksjon

$$\Pr(\cdot \mid B) \colon \mathcal{F} \to \mathbb{R}.$$

Det er en fin oppgave å sjekke at $\Pr(A \mid B)$ definert ved (1) for alle $A \in \mathcal{F}$, er et sannsynlighetsmål. Notasjonen '$A \mid B$' leses som 'A gitt B', og det er viktig å forstå at $A \mid B$ *ikke* er en hendelse, '$\mid$' er *ikke* en mengdeoperasjon. Kanskje notasjon som $Q_B(A) = \Pr(A \cap B)/\Pr(B)$, eller lignende, for den betingede sannsynligheten til A gitt B er tydeligere, men $\Pr(A \mid B)$ er så vanlig at vi vil stort sett bruke den.

3.2. **Oppgaver.**

Oppgave 3.1 La B være en hendelse med $\Pr(B) > 0$. Definer $A \mapsto \Pr(A \mid B)$ for alle hendelser A, ved (1).

(a) Vis at $\Pr(\cdot \mid B)$ er et sannsynlighetsmål.

(b) Hold hendelsen A konstant, og sjekk at $B \mapsto \Pr(A \mid B)$ *ikke* er et sannsynlighetsmål.

(c) Vis *multiplikasjonsregelen* multiplikasjonsregelen

$$\Pr(A \cap B) = \Pr(A \mid B)\Pr(B) = \Pr(B \mid A)\Pr(A),$$

såfremt de betingede sannsynlighetene er definert.

(d) La A og B være hendelser, med $\Pr(B) > 0$. Utled Bayes' formel Bayes' formel

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B)}.$$

Det er vanlig å kalle $\Pr(A)$ *a priori* sannsynligheten for A, og $\Pr(A \mid B)$ *a posteriori* sannsynligheten for A, eller kort bare prior og posterior. prior og posterior

Oppgave 3.2 La $A_1, \ldots, A_n$ være en partisjon av utfallsrommet, det vil si at disse er disjunkte hendelser hvis union utgjør hele utfallsrommet.

(a) La B være en hendelse. Vis *Loven om total sannsynlighet* Loven om total sannsynlighet

$$\Pr(B) = \Pr(B \mid A_1)\Pr(A_1) + \cdots + \Pr(B \mid A_n)\Pr(A_n).$$

Forklar hvorfor vi kan bruke denne formelen uten å først sjekke at $\Pr(A_j)$ er positiv for hver bidige j.

(b) Anta at $\Pr(B) > 0$. Utled fra det over at for enhver hendelse A_j tilhørende partisjonen $\{A_1, \ldots, A_n\}$ kan Bayes' formel skrives

$$\Pr(A_j \mid B) = \frac{\Pr(B \mid A_j)\Pr(A_j)}{\Pr(B \mid A_1)\Pr(A_1) + \cdots + \Pr(B \mid A_n)\Pr(A_n)}.$$

Utled også formelen for spesialtilfellet der partisjonen av utfallsrommet som vi er interessert i simpelthen er $\{A, A^c\}$.

Oppgave 3.3 En sykdom rammer 1 prosent av befolkningen. Det finnes en test for denne sykdommen som har en *sensitivitet* $= \Pr(\text{test sier syk} \mid \text{faktisk syk})$ på 95 prosent, og en *spesifisitet* $= \Pr(\text{test sier frisk} \mid \text{faktisk frisk})$ på 90 prosent. Et tilfeldig individ tar denne testen, og testen sier 'syk' (som noe kontraintuitivt i sensitivitet og spesifisitet

denne settingen kalles et 'positivt' testresultat). Hva er sannsynligheten for at denne personen faktisk er syk?

– Bruken av 'tilfeldig' i denne oppgaven er litt slepphendt, men vanlig. Det vi mener her, og som ofte menes med 'tilfeldig', er at alle individider/enheter i en populasjon har samme sannsynlighet for å bli trukket. Som regel vil det være tydelig fra konteksten om 'tilfeldig' brukes på denne måten.

Oppgave 3.4 Tenk på alle familier i verden med to og bare to barn.

(a) Fra denne populasjonen trekker du tilfeldig en familie, men alt du får vite om denne familien er at minst ett av barna er en jente. Hva er sannsynligheten for at familien du trakk har to jentebarn?

(b) Igjen, fra populasjonen beskrevet over trekker du tilfeldig en familie, og denne gangen er alt du får vite at minst ett av barna er en jente født på en tirsdag. Hva er sannsynligheten for at familien du trakk har to jentebarn?

(c) Skriv Python-kode der du ved hjelp av simuleringer 'sjekker' det du kom frem til i **(b)**. I Python, for å tilfeldig trekke ti tifeldige tall fra $\{1, \ldots, 7\}$ med tilbakelegg kan du skrive

```
import numpy as np
pr = np.zeros(7) + 1/7.
np.random.choice([1,2,3,4,5,6,7],size=10,p = pr,replace=True)
```

4. Uavhengige hendelser

Dette er notater knytte til Kap. 2.5 i Devore et al. (2021). Noen ganger vil det være slik at informasjon om én hendelse, ikke påvirker sannsynligheten vi gir en annen.

4.1. **Definisjon av uavhengighet.** Hendelsene A og B er uavhengige dersom

$$\Pr(A \cap B) = \Pr(A)\Pr(B). \tag{2}$$

To hendelser som *ikke* er uavhengighe, er avhengige. For å slippe å skrive at A og B er uavhengige, skriver vi $A \perp\!\!\!\perp B$. Så $A \perp\!\!\!\perp B$ betyr at A og B tilfredsstiller (2). Merk at $A \perp\!\!\!\perp B$ betyr det samme som $B \perp\!\!\!\perp A$.

Def. uavhengige hendelser

Nå definerer vi uavhengighet for flere enn to hendelser. Hendelsene $A_1, \ldots, A_m$ er uavhengige dersom

$$\Pr(\cap_{j\in J} A_j) = \prod_{j\in J} \Pr(A_j), \quad \text{for alle } J \subset \{1, \ldots, m\}. \tag{3}$$

Dette beskriver 2^m likheter, men det er ikke alle disse vi trenger å sjekke (som en liten telleoppgave: hvor mange likheter må vi sjekke?). Et eksempel: hendelsene A_1, A_2, A_3 er uavhengige dersom $\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1)\Pr(A_2)\Pr(A_3)$, og $\Pr(A_1 \cap A_2) = \Pr(A_1)\Pr(A_2)$, og $\Pr(A_1 \cap A_3) = \Pr(A_1)\Pr(A_3)$ og $\Pr(A_2 \cap A_3) = \Pr(A_2)\Pr(A_3)$. Dette er 4 likheter, ikke $2^3 = 8$.

Enda mer generelt: Hendelsene $A_1, A_2, \ldots$ er uavhengige dersom $A_{n_1}, \ldots A_{n_m}$ er uavhengige for enhver endelig delmengde $\{n_1, \ldots, n_m\} \subset \{1, 2, \ldots\}$.

En ting definisjonene i (2) og (3) skjuler litt, er at uavhengighet mellom to eller flere hendelser ikke kun er en egenskap ved hendelsene, men en egenskap ved kombinasjonen av hendelser *og* sannsynlighetsmål (se for eksempel Oppgave 4.1**(g)**). Dette innebærer blant annet at vi *ikke* kan 'tegne uavhengigehet' i et Venn-diagram.

Dersom $A_1, \ldots, A_m$ er uavhengige, ser vi fra definisjonen at de er *parvis uavhengige*, det vil si at $A_i \perp\!\!\!\perp A_j$ for alle $i \neq j$, som vi definerte det over. Den motsatte implikasjonen er *ikke* sann: Parvis uavhengige hendelser er ikke nødvendigvis uavhengige, som vi skal se i Oppgave 4.2.

parvis uavhengige

4.2. **Betinget uavhengighet.** Fra Oppgave 3.1 vet vi at $\Pr(\cdot \mid B)$ er et sannsynlighetsmål. Definisjonen av betinget uavhengighet er derfor strengt tatt overflødig, men vi tar den med. La B være en hendelse med $\Pr(B) > 0$. Hendelsene $A_1, \ldots, A_m$ er betinget uavhengige gitt B dersom (3) holder for $\Pr(\cdot \mid B)$, det vil si at $\Pr(\cap_{j\in J} A_j \mid B) = \prod_{j\in J} \Pr(A_j \mid B)$ for alle $J \subset \{1, \ldots, m\}$. For hendelsene A_1, A_2, og B med $\Pr(B) > 0$, betyr dette at A_1 er uavhengige av A_2 gitt B, dersom

$$\Pr(A_1 \cap A_2 \mid B) = \Pr(A_1 \mid B) \Pr(A_2 \mid B),$$

som jo er den samme definisjonen som i (2), bare under et annet sannsynlighetsmål.

4.3. **Oppgaver.**

Oppgave 4.1 La A og B være to hendelser.

(a) Vis at $A \perp\!\!\!\perp B$, $A \perp\!\!\!\perp B^c$, $A^c \perp\!\!\!\perp B$, og $A^c \perp\!\!\!\perp B^c$ sier det samme (er ekvivalente).

(b) La B være en hendelse med $\Pr(B) = 0$. Vis at $A \perp\!\!\!\perp B$ for alle hendelser A.

(c) La B være en hendelse med $\Pr(B) = 1$. Vis at $A \perp\!\!\!\perp B$ for alle hendelser A.

(d) Dersom $A \perp\!\!\!\perp A$, hva kan vi si om hendelsen A?

(e) Vis at dersom $\Pr(A) > 0$ og $\Pr(B) > 0$ og $A \cap B = \emptyset$, så er A og B avhengige.

(f) Vis at er det følgende ekvivalent

(i) $\Pr(A \cap B) = \Pr(A) \Pr(B)$;
(ii) $\Pr(A \mid B) = \Pr(A)$;
(iii) $\Pr(B \mid A) = \Pr(B)$,

såfremt de betingede sannsynlighetene er definert.

(g) To hendelser som er uavhengige under et sannsynlighetsmål, kan være avhengige under et annet. Se på delmengdene $A = \{1, 3, 5\}$ og $B = \{5, 6\}$ av utfallsrommet $\Omega = \{1, 2, 3, 4, 5, 6\}$. La Q_1 og Q_2 være sannsynlighetsmålene definert i Oppgave 2.3**(a)**&**(b)**. Vis at $A \perp\!\!\!\perp B$ under Q_1, men at A og B er avhengige under Q_2.

(h) Kan du konstruere et tredje sannsynlighetsmål som er slik at A og B er uavhengige hendelser?

Oppgave 4.2 Vi kaster to kronestykker, et blått og et rødt, som begge er slik at sannsynligheten for kron er $1/2$.

(a) La $B = \{\text{blå kron}\}$, $R = \{\text{rød kron}\}$ og $L = \{\text{to like utfall}\}$. Det er naturlig å anta at B og R er uavhengige, og vi antar det. Vis at hendelsene B, R, L er parvis uavhengige, men ikke uavhengige.

(b) Du har to rettferdige terninger, og kaster begge. Lag et eksempel på tre hendelser som er parvis uavhengige, men ikke uavhengige.

(c) I **(a)** og **(b)** brydde vi oss ikke med å presist beskrive sannsynlighetsrommene som vi implisitt jobber med. Forsøk å gi en presis beskrivelse av dem.

Oppgave 4.3 La A_1, A_2, og B være tre hendelser, med $\Pr(B) > 0$.

(a) Vis at multiplikasjonsregelen for betinget sannsynlighet tar følgende form

$$\Pr(A_1 \cap A_2 \mid B) = \Pr(A_1 \mid A_2 \cap B) \Pr(A_2 \mid B) = \Pr(A_2 \mid A_1 \cap B) \Pr(A_1 \mid B).$$

(b) Vis at dersom A_1 og A_2 er uavhengige gitt B, så er

$$\Pr(A_2 \mid A_1 \cap B) = \Pr(A_2 \mid B),$$

såfremt begge sider er definert.

(c) La A_1, A_2 og B være hendelser med sannsynligheter som verken er lik 0 eller 1. Vi tenker oss at A_1 skjer først, deretter A_2, og at vi oppdaterer sannsynligheten til B ettersom hendelsene inntreffer. Definer $Q(\cdot) = \Pr(\cdot \mid A_1)$. Dersom A_1 inntreffer er $Q(B) = \Pr(B \mid A_1)$ vår oppdaterte sannsynlighet for B. Hvis så A_2 intreffer, oppdaterer vi på nytt, og vår oppdaterte sannsynlighet for B er (vis dette)

$$Q(B \mid A_2) = \Pr(B \mid A_1 \cap A_2).$$

Anta nå at A_1 og A_2 er uavhengige gitt B, og gitt B^c. Vis at

$$Q(B \mid A_2) = \frac{\Pr(A_2 \mid B)Q(B)}{\Pr(A_2 \mid B)Q(B) + \Pr(A_2 \mid B^c)Q(B^c)}.$$

Her ser vi at vår posterior etter A_1 er blitt vår prior når A_2 skjer.

(d) Se tilbake på Oppgave 3.3. Personen som fikk positivt testresultat bestemmer seg for å ta testen en gang til. Nok en gang gir testen positivt resultat. Hva er nå sannsynligheten for at personen faktisk er syk? Hvilken svært rimelige antagelse gjør du her?

(e) Konstruer et enkelt eksempel der to hendelser A og B er uavhengige gitt C, men ikke uavhengige (altså avhengige) gitt C^c.

Oppgave 4.4 Vi har en hatt med to kronestykker som ser identiske ut, men ikke er det. Det ene kronestykket er rettferdig: sannynligheten for kron er den samme som sannynligheten for mynt, altså 1/2; det andre kronetsykket er ikke rettferdig, det har sannsynlighet 3/4 for kron.

(a) Du trekker et tilfeldig kronestykke fra hatten (som igjen betyr at begge kronestykkene har samme sannsynlighet for å bli trukket), og slår kron-eller-mynt. Hva er sannsynlighetene for kron?

(b) Det ble kron. Hva er sannsynligheten for at du trakk det rettferdige kronestykket?

(c) Du bestemmer deg for å slå kron-eller-mynt med det samme kronestykket en gang til. Hva er nå sannsynligheten for kron? Husk at du fikk kron på første kast. Hvilken (helt rimelige!) antagelse gjør du her?

(d) Det ble kron igjen. Hva er nå sannsynligheten for at du trakk det rettferdige kronestykket?

Oppgave 4.5 La A_1, A_2, A_3 være uavhengige hendelser.

(a) Vis at A_1, A_2, A_3^c er uavhengige; og at A_1, A_2^c, A_3^c er uavhengige; og så videre.

(b) Vis at $A_1 \cap A_2$ og A_3 er uavhengige.

(c) Vis at $A_1 \cup A_2$ og A_3 er uavhengige.

(d) Generaliser.

Oppgave 4.6 Med samme sannsynlighet trekker jeg tallet 4 eller 5, og kaster mitt rettferdige kronestykke så mange ganger som tallet jeg trakk. Du får vite at jeg fikk to kron, ikke flere. Hvor mange ganger tror du jeg kastet kronestykket mitt?

5. Tilfeldige variable og fordelinger

La $(\Omega, \mathcal{A}, \Pr)$ være et sannsynlighetsrom. En *tilfeldig variabel* er en funksjon $X(\omega)$ med utfallsrommet som domene og som antar reelle verdier, def. tilfeldig variabel

$$X\colon \Omega \to \mathbb{R},$$

slik at mengden $\{\omega \in \Omega\colon X(\omega) \leq x\}$ en hendelse for enhver $x \in \mathbb{R}$. Dersom $B \subset \mathbb{R}$ vil vi skrive $\{X \in B\}$ for $\{\omega \in \Omega\colon X(\omega) \in B\}$, siden det er vanlig og tar mindre plass. Spesielt er $\{X \leq x\} = \{\omega \in \Omega\colon X(\omega) \leq x\}$ med denne notasjonen.

Vi kommer til å møte *indikatorfunksjonen* flere ganger. Indikatorfunksjonen til delmengden A av Ω er indikatorfunksjonen

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \in A^c. \end{cases}$$

Dersom A er en hendelse, er I_A en tilfeldig variabel.

Eksempel. Vi triller en rettferdig terning, så utfallsrommet er $\Omega = \{1,2,3,4,5,6\}$ og sannsynlighetsmålet er $Q(A) = |A|/6$ for $A \in 2^\Omega$ (det er dette vi mener med at terningen er rettferdig, se Oppgave 2.3). Da har vi sannsynlighetsrommet $(\Omega, 2^\Omega, Q)$. La $A = \{1,3,5\}$ og definer funksjonen $X = I_A$, altså er X definert som indikatorfunksjonen til mengden A. Merk at

$$\{X \leq x\} = \begin{cases} \emptyset, & x < 0, \\ A^c, & 0 \leq x < 1, \\ \Omega, & 1 \leq x. \end{cases}$$

Siden $\emptyset$, A^c, og Ω er inneholdt in potensmengden 2^Ω viser dette at $\{X \leq x\}$ er en hendelse for alle $x \in \mathbb{R}$, som betyr at X er en tilfeldig variabel.

Å kreve at $\{X \leq x\}$ er en hendelse for alle $x \in \mathbb{R}$, er det samme som å kreve at $\{X \in B\}$ er en hendelse for alle $B \in \mathcal{B}$, der $\mathcal{B}$ er den minste samlingen hendelser på $\mathbb{R}$ som inneholder alle de åpne mengdene i $\mathbb{R}$.[1] Med at $\mathcal{B}$ er en samling hendelser mener vi her at $\mathcal{B}$ er en σ-algebra, slik det er definert på side 3.

Hvis vi har en tilfeldig variabel $X\colon \Omega \to \mathbb{R}$, og en funksjon $g\colon \mathbb{R} \to \mathbb{R}$, virker det rimelig at $g(X)$ også er en tilfeldig variabel: Siden X er tilfeldig, og $g(X)$ er en funksjon av noe tilfeldig, så er $g(X)$ selv tilfeldig. Ta for eksempel den tilfeldige variabelene $X = I_A$ definert over, og se på funksjonene $g(x) = 2x - 1$. Setter vi $Y = g(X)$, så er Y en funksjon definert på Ω med verdimengde $\{-1, 1\}$. Og på samme måte som med X kan vi sjekke at $\{Y \leq y\}$ er en hendelse for alle $y \in \mathbb{R}$. Dette leder til et utsagn som i dette emnet alltid vil være sant (men det er altså ikke helt sant): *enhver reell funksjon av tilfeldige variable er en tilfeldig variabel.* funksjoner av tilfeldige variable er tilfeldige variable

5.1. Fordelinger.

Nå har vi altså et sannsynlighetsrom $(\Omega, \mathcal{A}, \Pr)$ en funksjon $X\colon \Omega \to \mathbb{R}$, der $\mathbb{R}$ er utstyrt samlingen hendelser $\mathcal{B}$ beskrevet over, og X er slik at $\{X \in B\} \in \mathcal{A}$ for alle $B \in \mathcal{B}$. Definer

$$P_X(B) = \Pr(X \in B), \quad B \in \mathcal{B}, \tag{4}$$

der $\Pr(X \in B)$ er en rask måte å skrive $\Pr(\{X \leq x\}) = \Pr(\{\omega \in \Omega\colon X(\omega) \in B\})$ på. Vi kunne også skrevet $P_X(B) = \Pr(X^{-1}(B))$, ved hjelp av det inverse bildet $X^{-1}(B) = \{\omega \in \Omega\colon X(\omega) \in B\}$. Uansett, P_X som gitt i (4) definerer en funksjon med domene $\mathcal{B}$ som antar relle verdier (fordi Pr gjør det). Vi vil vise at P_X er

[1] Denne samlingen $\mathcal{B}$ er kjent som *Borel σ-algebraen* på $\mathbb{R}$, og du kan lære mer om den i senere matematikk- og sannsynlighetsemner på Blindern.

et sannsynlighetsmål. Det betyr at vi må vise at P_X tilfredsstiller Kolmogorovs aksiomer (i), (ii), og (iii), gitt på side 6. Når vi viser dette bruker vi aktivt at Pr *er* et sannsynlighetsmål. At $P_X(B) \geq 0$ for alle $B \in \mathcal{B}$ følger av at $\Pr(A) \geq 0$ for alle $A \in \mathcal{A}$. Vi ser også at

$$P_X(\mathbb{R}) = \Pr(X \in \mathbb{R}) = \Pr(\Omega) = 1.$$

Da gjenstår det å vise egenskap (iii): Altså at $P_X(\cup_{n\geq 1} B_N) = \sum_{n\geq 1} P_X(B_n)$ for enhver følge $B_1, B_2, \ldots$ av disjunkte mengder inneholdt i $\mathcal{B}$. La $B_1, B_2, \ldots$ være en slik følge. Siden $B_i \cap B_j = \emptyset$ for $i \neq j$, har vi at

$$\{X \in B_i\} \cap \{X \in B_j\} = \{X \in B_i \cap B_j\} = \{X \in \emptyset\} = \emptyset,$$

for alle $i \neq j$. Altså er $\{X \in B_1\}, \{X \in B_2\}, \ldots$ en disjunkt følge av mengder i $\mathcal{A}$. Men siden

$$\{X \in \cup_{n\geq 1} B_n\} = \cup_{n\geq 1} \{X \in B_n\},$$

ser vi at

$$\begin{aligned} P_X(\cup_{n\geq 1} B_n) &= \Pr(X \in \cup_{n\geq 1} B_n) \\ &= \Pr(\cup_{n\geq 1}\{X \in B_n\}) = \sum_{n\geq 1} \Pr(X \in B_n) = \sum_{n\geq 1} P_X(B_n), \end{aligned}$$

der den tredje likheten bruker at Pr er et sannsynlighetsmål, og derfor tilfredstiller egenskap (iii). Ved hjelp av den tilfeldige variabelen X har vi nå et nytt sannsynlighetsrom $(\mathbb{R}, \mathcal{B}, P_X)$, der sannsynlighetsmålet P_X kalles *fordelingen* til X (*distribution*, på engelsk). Fordelingen til en tilfeldig variabel forteller oss alt det fra et sannsynlighetsperspektiv er verdt å vite om en tilfeldig variabel. For å si at X er en tilfeldig variabel med fordeling P_X skriver vi

fordeling

$$X \sim P_X.$$

Dersom $X \sim P_X$ og $Y \sim P_Y$ og $P_X(B) = P_Y(B)$ for alle $B \in \mathcal{B}$, sier vi at X og Y har samme fordeling, eller er *identisk fordelte*.

identisk fordelte tilfeldige variable

Eksempel. Terningkast igjen: $\Omega = \{1, 2, 3, 4, 5, 6\}$, sannsynlighetsmål $Q(A) = |A|/6$ for alle A i potensmengden 2^Ω. La $A = \{1, 3, 5\}$ og definer $X(\omega) = I_A(\omega)$ og $Y(\omega) = I_{A^c}(\omega)$ for $\omega \in \Omega$. Da er X og Y tilfeldige variable, og siden $X = 1$ når $Y = 0$ og motsatt er $X(\omega) \neq Y(\omega)$ for alle $\omega \in \Omega$. De har imidlertid samme fordeling. For å se dette holder det å merke seg at

$$\Pr(X = 1) = \Pr(A) = \tfrac{1}{2} = \Pr(A^c) = \Pr(Y = 1).$$

For at to (eller flere) tilfeldige variable er identisk fordelte trenger de ikke være definert på det samme sannsynlighetsrommet.

5.2. **Den kumulative fordelingsfunksjonen.** La X være en tilfeldig variabel med fordeling P. Den kumulative fordelingsfunksjonen til X er definert ved

$$F(x) = P((-\infty, x]), \quad x \in \mathbb{R}.$$

Som vi ser av denne definisjonene spiller ikke X egentlig noen rolle her, kun fordelingen P. Men siden P (i tilfellene vi ser på) alltid er fordelingen til en tilfeldig variabel, sier vi at F er den kumulative fordelingsfunksjonen til denne tilfeldige variabelen. Merk også at her har vi ikke spesifisert noe bakenforliggende sannsynlighetsrom. Et bakenforliggende sannsynlighetsrom er der, men vi trenger ikke bry oss om det,

med unntak av at vi skriver Pr for sannsynlighetsmålet på dette bakenforliggende rommet, slik at $F(x) = P((-\infty, x]) = \Pr(X \leq x)$.

Lemma 1. *Den kumulative fordelingsfunksjonen F er ikke-avtagende, høyrekontinuerlig, har grenser fra venstre, og $\lim_{x\to-\infty} F(x) = 0$ og $\lim_{x\to\infty} F(x) = 1$.*

Bevis. La $x \leq y$, da er $(-\infty, x] \subset (-\infty, y]$, så det følger av Oppgave 2.1**(g)** at $F(x) = P(-\infty, x] \leq P(-\infty, y] = F(y)$, altså er F ikke-avtagende. For å vise at F er høyrekontinuerlig bruker vi kontinuitet av mål (se side 6): La $\varepsilon_1 > \varepsilon_2 > \dots$ være en følge positive tall som går mot null når n vokser. Da er $(-\infty, x + \varepsilon_n] \supset (-\infty, x + \varepsilon_{n+1}]$ for alle n, og $\cap_{n\geq 1}(-\infty, x + \varepsilon_n] = (-\infty, x]$. For enhver x er da

$$\lim_n F(x + \varepsilon_n) = \lim_n P(-\infty, x + \varepsilon_n] = P(\cap_{n\geq 1}(-\infty, x + \varepsilon_n]) = P(-\infty, x] = F(x),$$

der den andre likheten bruker kontinuitet av mål. De tre andre egenskapene bevises på tilsvarende måter, se Oppgave 5.1. □

Enhver funksjon som tilfredsstiller egenskapenen i Lemma 1 er fordelingsfunksjonene til en tilfeldig variabel. Det vil si, gitt en fordelingsfunksjon kan man konstruere en fordelingen med denne fordelingsfunksjonen, og med en fordeling kan man lage seg en tilfeldig variabel med denne fordelingen.

5.3. **Diskret og kontinuerlige tilfeldige variable.** La X være en tilfeldig variabel som er slik at $\Pr(X = x) > 0$ for $x \in S = \{a_1, a_2, \dots\}$, og $P(X = x) = 0$ for $x \notin S$. Da er X en *diskret tilfeldig variabel*. Vi definerer sannsynlighetsmassefunksjonen (*probability mass function*, pmf) f til den diskret tilfeldige variabelen X ved pmf

$$f(x) = \Pr(X = x), \quad x \in \mathbb{R}.$$

Da kan vi utrykke den kumulative fordelingsfunksjonen F til X som

$$F(x) = \sum_{y\in S, y\leq x} f(y).$$

Lag en tegning av en slik funskjon. Det blir en trappefunskjon.

La Y være en tilfeldig variabel med kumulativ fordelingsfunksjon F. Dersom det finnes en ikke-negativ funksjon $f(y)$ slik at

$$F(x) = \int_{-\infty}^{x} f(y), \mathrm{d}y,$$

kaller vi Y en *kontinuerlig tilfeldig variabel*, og sier at f er tettheten til Y (eller sannsynlighetstettheten, *probability density function*, forkortet pdf, på engelsk). pdf

Vi ser fra dette at en pmf/pdf f er slik at $f(x) \geq 0$ for all $x \in \mathbb{R}$, og at f enten summerer eller integrerer til 1. Enhver funksjon med disse egenskapene er en pmf/pdf. Og blir vi gitt en pmf eller pdf kan vi lage oss en fordelingsfunksjon, og følgelig en tilfeldig variabel som har denne fordelingsfunksjonen.

5.4. **Uavhengige tilfeldige variable.** To tilfeldige variable X og Y er uavhengige dersom hendelsene $\{X \leq x\}$ og $\{Y \leq y\}$ er uavhengige for alle $x, y \in \mathbb{R}$. De tilfeldig variablene $X_1, X_2, \dots$ er uavhengige dersom hendelsene

$$\{X_1 \leq x_1\}, \dots, \{X_n \leq x_n\},$$

er uavhengige for alle $x_1, \ldots, x_n \in \mathbb{R}$. Dette betyr det samme som tidligere (se Seksjon 4.1), nemlig at for alle $x_1, \ldots, x_n \in \mathbb{R}$ er

$$\Pr\left(\cap_{j\in J}\{X_j \le x_j\}\right) = \prod_{j\in J} \Pr(X_j \le x_j), \quad \text{for alle } J \subset \{1,\ldots,n\}.$$

Vi ser på et eksempel.

Eksempel. La utfallsrommet være $\Omega = \{1,2,3,4,5,6\}$, utstyrt med potensmengden 2^Ω og sannsynlighetsmålet $Q(A) = |A|/6$ for $A \in 2^\Omega$. Vi har sett at $A = \{1,3,5\}$ og $B = \{5,6\}$ er uavhengige hendelser. Definer de to tilfeldige variablene $X = aI_A$ og $Y = bI_B$ for a og b som begge positive (for enkelhets skyld). Da er

$$\{X \le x\} = \begin{cases} \emptyset, & x < 0, \\ A^c, & 0 \le x < a \\ \Omega, & a \le x, \end{cases} \quad \text{og} \quad \{Y \le y\} = \begin{cases} \emptyset, & y < 0, \\ B^c, & 0 \le y < b \\ \Omega, & b \le y, \end{cases}$$

Siden A^c og B^c er uavhengige, og alle hendelser er uavhengige av $\emptyset$ og av Ω (se Oppgave 4.1), så ser vi at X og Y er uavhengige.

Tar vi dette eksemplet videre skjønner vi at to diskrete tilfeldige variable X og Y er uavhengige dersom hendelsene $\{X = x\}$ og $\{Y = y\}$ er uavhengige for alle $x, y \in \mathbb{R}$. For å sjekke uavhengighet av to diskrete tilfeldige variable er det nok å se på de x, y som er slik at $\Pr(X = x) > 0$ og $\Pr(Y = y) > 0$ (Se Oppgave 4.1 igjen).

uif,iid

En forkortelse som dukker opp overalt i sannsynlighet og statstikk er *uif*, eller *iid* på engelsk. Vi sier at de tilfeldige variablene $X_1, \ldots, X_n$ er *uavhengige og identisk fordelte* (uif) dersom de er uavhengige og $\Pr(X_i \le x) = \Pr(X_j \le x)$ for alle $x \in \mathbb{R}$ og all $i, j = 1, \ldots, n$, som betyr at de har samme fordeling.

5.5. **Oppgaver.**

Oppgave 5.1 Bevis resten av Lemma 1 om egenskapene til kumulative fordelingsfunksjoner.

Oppgave 5.2 La F være en fordelingfunskjon til en tilfeldig variabel X.

(a) Anta at F er kontinuerlig i punktet x. Siden F er er kontinuerlig fra høyre i alle punkter, betyr dette at i x er F også kontinuerlig fra venstre. Vis at $\Pr(X = x) = 0$.

(b) Vis at $\Pr\{X \in (a,b]\} = F(b) - F(a)$ for alle $a < b$.

(c) Anta at i punktet x er $\Pr(X = x) > 0$. Vis at $\Pr(X = x) = F(x) - \lim_{y\uparrow x} F(y)$.

(d) La $a < b$ og anta at F er kontinuerlig i a og i b. Vis at da er $\Pr\{X \in (a,b]\} = \Pr\{X \in [a,b)\} = \Pr\{X \in [a,b]\} = \Pr\{X \in (a,b)\}$.

(e) Ikke pensum (men fint å vite): La $J = \{x \in \mathbb{R} \colon \Pr(X = x) > 0\}$, det vil si at J er alle punktene der F gjør et hopp. Vis at J maksimalt inneholder tellbart mange punkter.

Oppgave 5.3 For et tall $p \in (0,1)$, se på funksjonen

$$f(x) = (1-p)^{x-1}p, \quad \text{for } x = 1, 2, \ldots,$$

og $f(x) = 0$ ellers.

(a) Vis at f er en sannsynlighetsmassefunksjon. Dette er pmfen til en geometrisk fordelt tilfeldig variabel.

(b) Finn den kumulative fordelingsfunksjonen F med pmf f.

Oppgave 5.4 For tall $a, b > 0$, se på funksjonen

$$f(x) = \frac{a}{b}\big(\frac{x}{b}\big)^{a-1}, \quad x \in [0, b],$$

og $f(x) = 0$ for $x \notin [0, b]$.

(a) Forklar at f er en pdf.

(b) Finn den kumulative fordelingsfunksjonen F med pdf f.

Oppgave 5.5 La $X_1, \ldots, X_n$ være uavhengige og identisk fordelte tilfeldige variable, med fordelingsfunksjon F, dvs. $F(x) = \Pr(X_i \le x)$ for all i.

(a) Definer $M_n = \max_{i \le n} X_i$. Forklar at

$$\Pr(M_n \le z) = F(z)^n.$$

(b) Anta at $F(x) = 0$ for $x < 0$, $F(x) = x$ for $x \in [0, 1)$ og $F(x) = 1$ for $x \ge 1$ (dette betyr at hver X_i har en uniform fordeling på enhetsintervallet, mer om det senere). Finn et utrykk for fordelingsfunskjonen til den tilfeldige variabelen $Y_n = n(1 - M_n)$.

6. Forventing og varians

Vi skal definere forventningen til en tilfeldig variabel i stadig mer generelle tilfeller (men helt generelt greier vi ikke i dette emnet). En viktig ting å ha i bakhodet er at egenskapene vi etablerer for forventningen av de enkleste tilfeldige variabler, også gjelder for mer avanserte tilfeldige variable. Dog er det slik at løftet fra de enkle til de mer avanserte krever matematiske verktøy vi ikke har tid til å studere i dette emnet. Fortvil ikke, du kommer til å lære masse mer om dette i senere matematikk- og sannsynlighetsemner.

6.1. Forventning på endelige utfallsrom.

Det gir innsikt å starte med tilfeldige variable på endelige utfallsrom. La $\Omega = \{\omega_1, \ldots, \omega_m\}$ være et utfallsrom utstyrt med potensmengden 2^Ω av hendelser, og et sannsynlighetsmål Pr. La X være en tilfeldig variabel på dette utfallsrommet. La $\{x_1, \ldots, x_k\}$ være verdimengden til X (så $k \le m$), og Uttrykk X som en *enkel tilfeldig variabel*

enkel tilfeldig variabel

$$X(\omega) = x_1 I_{A_1}(\omega) + \cdots + x_k I_{A_k}(\omega), \quad \omega \in \Omega, \tag{5}$$

dere I_A er indikatorfunksjonen til mengden A, og $A_j = \{\omega \in \Omega \colon X(\omega) = x_j\}$ for $j = 1, \ldots, k$. Forventningen til X skrives $\mathrm{E}\, X$, eller $\mathrm{E}\,(X)$, eller $\mathrm{E}\,[X]$, eller $\mathrm{E}\,\{X\}$ eller lignende, og er definert ved

forventning

$$\mathrm{E}\, X = \sum_{\omega \in \Omega} X(\omega) \Pr(\{\omega\}). \tag{6}$$

Legg merke til at dersom $X = I_A$ for en hendelse A, så følger det av additivteten til sannsynlighetsmål at $\mathrm{E}\, I_A = \Pr(A)$. Siden denne forventningen er en rekke (en sum), gjelder alt du kan om rekker også for forventningen. For eksempel er forventningen til en konstnat lik konstanten, $\mathrm{E}\,(a) = \sum_{\omega \in \Omega} a \Pr(\{\omega\}) = a$. Videre ser vi at

$$\mathrm{E}\, X \ge 0, \quad \text{dersom } X \ge 0,$$

fordi hvert ledd $X(\omega_j) \Pr(\{\omega_j\}) \ge 0$; og at dersom X og Y er to tilfeldige variable, og a og b er to reelle tall, så er

$$\mathrm{E}\,(aX + bY) = a\,\mathrm{E}\, X + b\,\mathrm{E}\, Y.$$

Kombinerer vi disse to egenskapene ser vi at hvis $X(\omega) \le Y(\omega)$ for alle ω, så er $\mathrm{E}\, X \le \mathrm{E}\, Y$.

6.2. **Forventning av enkle funksjoner.** Setter vi høyresiden av (5) inn i (6), og bruker disse egenskapene, ser vi at $\mathrm{E}\, X = x_1 \Pr(A_1) + \cdots + x_k \Pr(A_k)$. Tar vi dette som definisjonen av forventningen av enkle tilfeldige variable, har vi en definisjon der vi ikke bruker at utfallsrommet Ω er endelig (som i (6)). Mer presist, la $(\Omega, \mathcal{A}, \Pr)$ være et sannsynlighetsrom og sett $X = \sum_{i=1}^{k} x_i I_{A_i}$ for en endelig partisjon $\{A_1, \ldots, A_k\}$ av Ω, der $A_i \in \mathcal{A}$ for alle i, så X er en enkel tilfeldig variabel. Da definerer vi forventningen til X ved

$$\mathrm{E}\, X = x_1 \Pr(A_1) + \cdots + x_k \Pr(A_k). \tag{7}$$

Fra denne definisjonen er det umiddelbart at $\mathrm{E}\, I_A = \Pr(A)$ for en hendelse A; og at $\mathrm{E}\,(a) = a$ for en konstant a. I Oppgave 6.2 utleder vi flere egenskaper.

6.3. **Forventning av diskrete og kontinuerlige tilfeldig variable.** Forventningen til tilfeldige variable med tellbart uendelige verdimengder (f.eks. $\{1, 2, \ldots\}$) eller ikke-tellbart uendelige verdimengder (f.eks. $[0, 1]$, $[0, \infty)$, $\mathbb{R}$) er definert som en grense av forventningen til enkle tilfeldige variable som tilnærmer den tilfeldige variabelen vi vil regne forventningen til. Den presise definisjonen vil du møte i senere kurs. Poenget her er at denne definisjonen gjør at egenskaper ved forventningen til enkle tilfeldige variable overføres til forventningen av mer generelle tilfeldige variable. Så vi behøver egentlig ikke bevise noe på nytt for de definisjonene vi nå skal gi, selv om vi kan gjøre det.

Her er definisjonene (som er de gitt i boka (Devore et al., 2021, s. 127 og s. 203)): Hvis X er en tilfeldig variabel med verdimengde $\{x_1, x_2, \ldots\}$ og pmf $f(x) = \Pr(X = x)$ definerer vi forventninge til X ved

$$\mathrm{E}\, X = \sum_{j=1}^{\infty} x_j f(x_j),$$

såfremt $\sum_{j=1}^{\infty} |x_j| f(x_j)$ er konvergent. Dersom X er en kontinuerlig tilfeldig variabel med pdf $f(x)$ definerer vi forventningen til X ved

$$\mathrm{E}\, X = \int_{-\infty}^{\infty} x f(x)\, \mathrm{d}x,$$

såfremt $\int_{-\infty}^{\infty} |x| f(x)\, \mathrm{d}x$ er endelig.

6.4. **Varians og kovarians.** Variansen til en tilfeldig variabel er definert så lenge $\mathrm{E}\, X^2$ er endelig. Vi skriver $\mathrm{Var}(X)$ for variansen til X, og den er definert ved

varians

$$\mathrm{Var}(X) = \mathrm{E}\,\{(X - \mathrm{E}\,[X])^2\}.$$

kovarians

Kovariansen til X og Y er definert så lenge $\mathrm{E}\, X^2$ og $\mathrm{E}\, Y^2$ er endelige. Den er gitt ved

$$\mathrm{Cov}(X, Y) = \mathrm{E}\,\{(X - \mathrm{E}\,[X])(Y - \mathrm{E}\,[Y])\}.$$

Begge disse definisjonen gjelder uavhengig av om en eller begge av X og Y er endelige, diskrete, eller kontinuerlige.

6.5. **Oppgaver.**

Oppgave 6.1 Vis at definisjonene av $\mathrm{E}\, X = \mathrm{E} \sum_{j=1}^{k} x_j A_j$ gitt i (7) ikke avhenger av representasjonen vi gir X. Det vil si, vis at dersom $X = \sum_{j=1}^{k} x_j I_{A_j} = \sum_{i=1}^{m} y_i I_{B_i}$ for to forskjellige partisjoner $\{A_1, \ldots, A_k\}$ og $\{B_1, \ldots, B_m\}$ av Ω, så er $\sum_{j=1}^{k} x_j \Pr(A_j) = \sum_{i=1}^{m} y_i \Pr(B_i)$

Oppgave 6.2 La $X = \sum_{j=1}^{k} x_j I_{A_j}$ og $Y = \sum_{i=1}^{m} y_j I_{B_j}$ være to enkle tilfeldige variable. Dette betyr at $\{A_1, \ldots, A_k\}$ og $\{B_1, \ldots, B_m\}$ er partisjoner av Ω, og at A_j og B_i er hendelser.

(a) La a, b være konstanter. Vis at forventningen er lineær

$$\mathrm{E}\,(aX + bY) = a\,\mathrm{E}\,X + b\,\mathrm{E}\,Y.$$

(b) Vis at dersom $X(\omega) \leq Y(\omega)$ for all ω, så er $\mathrm{E}\,X \leq \mathrm{E}\,Y$.

(c) La g være en funksjon. Vis at $\mathrm{E}\,g(X) = \sum_{j=1}^{k} g(x_j)\,\Pr(A_j)$.

(d) Vis at $|\mathrm{E}\,X| \leq \mathrm{E}\,|X|$.

(e) Anta at X og Y er uavhengige. Vis at

$$\mathrm{E}\,(XY) = \mathrm{E}\,(X)\mathrm{E}\,(Y).$$

(f) Forsøk å konstruer et eksempel som viser at implikasjonen i **(e)** ikke går det andre veien; det vil si, $\mathrm{E}\,(XY) = \mathrm{E}\,(X)\mathrm{E}\,(Y)$ betyr *ikke* at X og Y er uavhengige.

Oppgave 6.3 La X og Y være tilfeldige variable (med endelige andremomenter), og a og b konstanter.

(a) Vis at $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$, og at $\mathrm{Var}(X + b) = \mathrm{Var}(X)$.

(b) Vis at $\mathrm{Var}(X) \geq 0$, med likhet kun i tilfeller der X er konstant.

(c) Vis at

$$\mathrm{Var}(X) = \mathrm{E}\,X^2 - (\mathrm{E}\,X)^2,$$

og utled at $\mathrm{E}\,X^2 \geq (\mathrm{E}\,X)^2$.

(d) På tilsvarende vis, vis at

$$\mathrm{Cov}(X, Y) = \mathrm{E}\,(XY) - \mathrm{E}\,(X)\mathrm{E}\,(Y),$$

og utled fra dette at dersom X og Y er uavhengige, så er $\mathrm{Cov}(X, Y) = 0$.

(e) Vis at

$$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y),$$

så spesielt er $\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$ dersom X og Y er uavhengige.

Oppgave 6.4 La $X_1, \ldots, X_n$ være tilfeldige variable (med endelige andremomenter når nødvendig) og $a_1, \ldots, a_n$ være konstanter.

(a) Utled fra det over at

$$\mathrm{E} \sum_{i=1}^{n} a_i X_i = \sum_{i=1}^{n} a_i \mathrm{E}\,X_i.$$

For det empiriske gjennomsnittet av n tilfeldige variable skriver vi gjerne $\bar{X}_n$, så $\bar{X}_n = (1/n)\sum_{i=1}^{n} X_i$. Spesielt, vis dersom $X_1, \ldots, X_n$ er identisk fordelte med forventning $\mathrm{E}\,X_1 = \mu$, så er

$$\mathrm{E}\,\bar{X}_n = \mu,$$

(b) Vis at

$$\mathrm{Var}(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_i a_j \mathrm{Cov}(X_i, X_j).$$

Spesielt, dersom $X_1, \ldots, X_n$ er uavhengige så er $\mathrm{Var}(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i)$.

(c) Vis at dersom $X_1, \ldots, X_n$ er uavhengige og identisk fordelte med varians $\mathrm{Var}(X_1) = \sigma^2$, så er

$$\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n},$$

der $\bar{X}_n$ er det empiriske gjennomsnittet.

Bernoulli-fordelingen

Oppgave 6.5 La X være en tilfeldig variabel med sannsynlighetsmassefunksjon $f(x) = p^x(1-p)^{1-x}$ for $x = 0, 1$, og $f(x) = 0$ eller, der $0 \leq p \leq 1$. Vi sier at X har en Bernoulli fordeling med suksessannsynlighet p.

(a) Sjekk først at $\Pr(X = 1) = p$, som forklarer at p kalles suksessannsynligheten.

(b) Vis at forventningen og variansen til X er

$$\mathrm{E}\, X = p, \quad \text{og} \quad \mathrm{Var}(X) = p(1-p).$$

(c) La X_1, X_2 være uavhengige og identisk fordelte Bernoullivariable. Vis at

$$\Pr(X_1 = 1, X_2 = 0) = p(1-p),$$

men at

$$\Pr(X_1 + X_2 = 1) = 2p^2(1-p) = \binom{2}{1} p^2(1-p).$$

(d) Tilsvarende, forklar at dersom $X_1, \ldots, X_n$ er uavhengige og identisk fordelte Bernoullivariable med suksesssannynlighet p, så for $k \in \{0, \ldots, n\}$ er

$$\Pr(X_1 = \cdots = X_k = 1, X_{k+1} = \cdots = X_n = 0) = p^k(1-p)^{n-k},$$

binomisk fordeling

mens

$$\Pr(X_1 + \cdots + X_n = k) = \binom{n}{k} p^k(1-p)^{n-k}.$$

Dette er pmfen til en binomisk fordeling med parametere (n, p).

(e) La X være binomisk fordelt med parametere (n, p). Vis at

$$\mathrm{E}\, X = np, \quad \text{og} \quad \mathrm{Var}(X) = np(1-p).$$

Oppgave 6.6 La X være en geometrisk fordelt tilfeldig variabel, det vil si at den har pmf på formen $f(x) = (1-p)^{x-1}p$ for $x = 1, 2, \ldots$, der $p \in (0, 1)$.

(a) Finn forventingen til X.

(b) Finn variansen til X.

(c) La $X_1, X_2, \ldots$ være uavhengige og identisk foredelte Bernoulli variable med suksessannsynlighet p. La $Y = \min\{k \geq 1 \colon \sum_{i=1}^{k} X_i = 1\}$. Hva er sannsynligheten for at $Y = 1$; for at $Y = 2$; for at $Y = 3$? Forklar at Y har en geometrisk fordeling med pmf $f(x)$ som gitt over.

Poissonfordeling

Oppgave 6.7 En tilfeldig variabel X har en Poissonfordeling med parameter $\lambda > 0$ dersom den har pmf

$$f(x) = \frac{1}{x!}\lambda^x \exp(-\lambda), \quad x = 0, 1, 2, \ldots,$$

og $f(x) = 0$ ellers.

(a) Sjekk at $f(x)$ er en pmf.

(b) Vis at både forventningen og variansen til X er λ.

Oppgave 6.8 Se på sannsynlighetstettheten gitt i Oppgave 5.4, nemlig $f(x) = (a/b)(x/b)^{a-1}$ for $x \in [0, b]$ og $f(x) = 0$ ellers, der $a, b > 0$. La X være en tilfeldig variabel med fordeling gitt av f.

(a) Finn forventningen til X.

(b) Finn variansen til X.

(c) La $X_1, \ldots, X_n$ være uif med samme fordeling som X. Definer $Y = \max_{i \le n} X_i$. Finn fordelingsfunksjonen til Y, og pdfen til Y. Hva er forventningen og variansen til Y?

Oppgave 6.9 En tilfeldig variabel X er uniformt fordelt på intervallet $[a, b]$ for $a < b$, dersom dens sannsynlighetstetthet (pdf) er

uniform fordeling

$$f(x) = \frac{1}{b-a}, \quad x \in [a, b],$$

og $f(x) = 0$ utenfor dette intervallet.

(a) Sjekk at $f(x)$ er en pdf.

(b) Sjekk også at forventningen og variansen til

$$\mathrm{E}\, X = \frac{a+b}{2}, \quad \text{og} \quad \mathrm{Var}\, X = \frac{(b-a)^2}{12}$$

(c) Finn den kumulativ fordelingsfunksjonen F til X. Spesielt, sjekk at hvis $a = 0$ og $b = 1$, så er $F(x) = x$ for $x \in [0, 1)$.

(d) La Y være en tilfeldig variabel med fordelingen vi studerte i Oppgave 6.8, den med tetthet $f_Y(y) = (a/b)(x/b)^{a-1}$ for $y \in [0, b]$. Skriv F_Y for dens kumulative fordelingsfunksjon. La U være en tilfeldig variabel uniformt fordelt på $[0, 1]$. Vis at

$$F_Y^{-1}(U) \sim F_Y,$$

som betyr at $F_Y^{-1}(U)$ har samme fordeling som Y. Forsøk å generaliser dette eksemplet.

Oppgave 6.10 La $f(x) = 2x$ for $x \in [0, 1]$.

(a) Skjekk at dette virkelig er en sannsynlighetstetthet, for en stokastisk variabel X, og finn dens forventning og varians.

(b) Finn formel for den kumulative $F(x)$, og bruk dette til å finne medianen.

(c) Finn et 95 prosent-intervall $[a, b]$, slik at sannsynligheten for å havne der er 0.95.

(d) Se på U_1, U_2, uavhengige fra den vanlige uniforme fordelingen på $[0, 1]$, og på den som er størst, $X = \max(U_1, U_2)$. Vis at X har tettheten over. Dette er altså et eksempel på et 'men hvor komme denne tetteheten fra' spørsmål.

Oppgave 6.11 Noen oppgaver om normalfordelingen.

(a) For X normalfordelt med forventning μ og varians σ^2, vis at dens mgf er $M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$.

(b) For X og Y uavhengige normalfordelte, med parameterpar (μ_1, σ_1^2) og (μ_2, σ_2^2), vis at variabelen $Z = X + Y$ også er normal, med parameterpar $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

(c) Vi besøker Mars, der høyden M for en mann er $\mathrm{N}(1.80, 0.20^2)$ og K for en kvinne er $\mathrm{N}(1.70, 0.15^2)$. Vi besøker nå et tilfeldig par, blant de millioner av par der. Hva er sjansen for at mannen er høyere enn kvinnen?

7. Konvergens i sannsynlighet

I denne seksjonen gjentar vi noen ting dere har vært gjennnom i obligen, og går litt videre.

7.1. **Markov og Tsjebysjov.** La X være en ikke-negativ tilfeldig variabel med endelige forventning. For enhver $\varepsilon > 0$ vil det holde at

$$\mathrm{E}\, X \geq \mathrm{E}\, X I_{X \geq \varepsilon} \geq \varepsilon \Pr(X \geq \varepsilon),$$

Markovs ulikhet

som gir Markovs ulikhet,

$$\Pr(X \geq \varepsilon) \leq \frac{\mathrm{E}\, X}{\varepsilon},$$

som altså holder for enhver positiv ε. Siden X er ikke-negativ er $\{X \geq \varepsilon\} = \{X^2 \geq \varepsilon^2\}$, som betyr at $\Pr(X \geq \varepsilon) \leq \mathrm{E}\, X^2/\varepsilon^2$, såfremt X^2 har endelig forventning. Fra dette utleder vi Tsjebysjovs ulikhet: La Y være en tilfeldig variabel endelig varians (så ikke nødvendigvis ikke-negativ), da holder det at

Tsjebysjovs ulikhet

$$\Pr(|Y - \mathrm{E}\, Y| \geq \varepsilon) \leq \frac{\mathrm{Var}\, Y}{\varepsilon^2},$$

for enhver $\varepsilon > 0$. Dersom $Y_1, \ldots, Y_n$ er uavhengige og identisk fordelte tilfeldig variable med forventing μ og varians σ^2, og $\bar{Y}_n = (1/n)\sum_{i=1}^n Y_i$ er gjennomsnittet, så vil (vis det)

$$\Pr(|\bar{Y}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2},$$

for enhver $\varepsilon > 0$. Denne ulikhetene forklarer langt på vei hvorfor større tilfeldige utvalg er mer til å stole på enn mindre tilfeldige utvalg.

7.2. **Definisjonen.** At en følge $x_1, x_2, \ldots$ av reelle tall konvergerer til x, betyr at vi for enhver $\varepsilon > 0$ kan finne en n_0 slik at

$$|x_n - x| < \varepsilon, \quad \text{for alle } n \geq n_0.$$

Vi skriver gjerne $x_n \to x$ eller $\lim_{n\to\infty} x_n = x$.

La $X_1, X_2, \ldots$ være en følge av tilfeldige variable. Vi sier at X_n konvergerer i sannsynlighet til X dersom vi uansett hvilken $\varepsilon > 0$ vi blir utfordret med, vil

def. konv. i sannsynlighet

$$\Pr(|X_n - X| < \varepsilon) \to 1, \quad \text{når } n \text{ går mot uendelig.}$$

Dette er det samme som å si at for enhver $\varepsilon > 0$ vil

$$\Pr(|X_n - X| \geq \varepsilon) \to 0, \quad \text{når } n \text{ går mot uendelig}$$

Vi skriver gjerne $X_n \to_p X$ når dette er tilfellet, og svært ofte vil X være en konstant, altså ikke en tilfeldig variabel.

7.3. **Kontinuitetsteoremet.** Her skal vi se at konvergens i sannsynlighet oppfører seg 'som forventet'. La $X_1, X_2, \ldots$ være en følge av tilfeldige variable slik at

$$X_n \to_p a,$$

for en konstant a, og la $g(x)$ være en funksjon som er kontinuerlig i a. At $g(x)$ er kontinuerlig i a betyr at for enhver $\varepsilon > 0$ kan vi finne $\delta > 0$ slik at

$$\text{hvis } |x - a| < \delta \text{ så er } |g(x) - g(a)| < \varepsilon.$$

Men dette medfører at for enhver n har vi $\{|X_n - a| < \delta\} \subset \{|g(X_n) - g(a)| < \varepsilon\}$, som gir ulikheten

$$\Pr(|g(X_n) - g(a)| \geq \varepsilon) \leq \Pr(|X_n - a| \geq \delta),$$

og siden $\Pr(|X_n - a| \geq \delta) \to 0$ ser vi at også $\Pr(|g(X_n) - g(a)| \geq \varepsilon) \to 0$, som er det vi mener når vi skriver

$$g(X_n) \to_p g(a).$$

Dette er et eksempel på det som kalles kontinuitetsteoremet (continuous mapping theorem på engelsk). Et meget nyttig teorem.

Kontinuitetsteoremet forteller oss også at dersom $X_n \to_p a$ og $Y_n \to_p b$, og $g(x, y)$ er en reell funksjon som er kontinuerlig i (a, b), så vil

$$g(X_n, Y_n) \to_p g(a, b),$$

når n går mot uendelig. Det er ikke pensum å kunne vise dette, men det er uansett gøy å prøve, se Oppgave 7.5. En konsekvens av kontinuitetsteoremet som *er* pensum at dersom $X_n \to_p a$ og $Y_n \to_p b$ så vil

$$X_n + Y_n \to_p a + b, \quad X_n Y_n \to_p ab, \quad \text{og} \quad X_n / Y_n \to_p a/b, \tag{8}$$

den siste såfremt $b \neq 0$.

Oppgave 7.1 La $Y_1, \ldots, Y_n$ være uif tilfeldige variable med endelig varians (da har de også endelig forventning). Vis de store talls lov, nemlig at $\bar{Y}_n \to_p \mathrm{E}\, Y_1$ når n går mot uendelig. store talls lov

Oppgave 7.2 Her ser vi på noen anvendelser av de store talls lov.

(a) La $Y_1, \ldots, Y_n$ være uif Bernoulliforsøk med suksessannsynlighet θ. Finn sannsynlighetsgrensen til $(1/n)\sum_{i=1}^n Y_i$, det vil si det denne konvergerer mot i sannsynlighet.

(b) La $Z_1, Z_2, \ldots$ være uif standard normalfordelte tilfeldige variable, det vil si at Z_1 har pdf $\phi(z) = (2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)$ og cdf $\Phi(z) = \int_{-\infty}^z \phi(u)\,\mathrm{d}u$, for $z \in \mathbb{R}$. Finn sannsynlighetsgrensen til $(1/n)\sum_{i=1}^n I\{Z_i \geq 0\}$.

(c) La $X_1, \ldots, X_n$ være uif tilfeldige variable med tetthet $f(x; \mu) = \phi(x - \mu)$. La $u(x; \mu) = \partial \log f(x; \mu)/\partial\mu$. Finn sannsynlighetsgrensen til $(1/n)\sum_{i=1}^n u(X_i; \mu)$.

(d) La $U_1, \ldots, U_n$ være uif uniformt fordelte tilfeldig variable på $[0, 1]$. Finn sannsynlighetsgrensen til $-(1/n)\sum_{i=1}^n \log(1 - U_i)/\theta$.

(e) I Python (og i R) er det funksjoner som lar deg trekke fra den uniform fordelingen, fra normalfordelingen, og mange andre. Simuler litt og sjekk at det du har kommet frem til i oppgavene stemmer.

Oppgave 7.3 Noen flere konvergens i sannsynlighetsoppgaver.

(a) La $Y_1, \ldots, Y_n$ være uif Bernoulliforsøk med suksessannsynlighet θ. Vis at $\bar{Y}_n^2 \to_p \theta^2$, og bruk dette til å vise at $(1/n)\sum_{i=1}^n (Y_i - \bar{Y}_n)^2 \to_p \theta(1 - \theta) = \mathrm{Var}(Y_1)$.

(b) Mer generelt, la $Y_1, \ldots, Y_n$ være uif tilfeldige variable med varians σ^2 og $\mathrm{E}\, Y_1^4$ endelig. Vis at

$$\frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2 \to_p \sigma^2.$$

(c) La $X_1, \ldots, X_n$ være uif normalfordelte tilfeldige variable med forventning μ og varians σ^2, vis at $(1/n)\sum_{i=1}^n \exp(tX_i) \to_p \exp(t\mu + \frac{1}{2}t^2\sigma^2)$, og at $\exp(t\bar{X}_n) \to_p \exp(t\mu)$.

Oppgave 7.4 En god måte å 'jukse' på når vi skal regne ut en vanskelig sannsynlighet er å bruke simuleringer. Ved de store talls lov vil simuleringene, dersom de er gjort riktig, gi oss en indikasjon på om vi har regnet riktig. Poenget er å oversette sannsynlighetsspørsmålet vi ønsker å besvare til et spørsmål om sannsynligheten for at en tilfeldig variable X lander i A, altså $\Pr(X \in A)$. Deretter går vi til datamaskinen vår og simulerer mange uif tilfeldige variable $X_1, \ldots, X_n$, der

hver av disse har samme fordeling som X. Så regner vi ut $(1/n)\sum_{i=1}^n I\{X_i \in A\}$. Siden denne konvergerer i sannsynlighet til $\Pr(X \in A)$, vil den være et godt anslag på sannsynligheten vi ønsker å regne ut. Some et eksempel, ta følgende spørsmål fra prøvemidtveisen: *En familie har to (og bare to) barn. Du blir fortalt at det ene barnet er en jente født på en tirsdag før lunsj. Hva er sannsynligheten for at denne familien har to jenter?* Vi vet at svaret er 27/55, men lat som du ikke vet det, og gå til Python eller R og forsøk å simuler deg frem til dette svaret.

Oppgave 7.5 Denne oppgaven er ikke pensum. Funksjonen $g\colon \mathbb{R}^2 \to \mathbb{R}$ er kontinuerlig i (a,b) dersom vi for enhver $\varepsilon > 0$ kan finne $\delta > 0$ slik at hvis $\|(x,y)-(a,b)\| = \sqrt{(x-a)^2+(y-b)^2} < \delta$ så er $|g(x,y)-g(a,b)| < \varepsilon$.

(a) Anta at $X_n \to_p a$ og $Y_n \to_p b$. Vis at for $\varepsilon > 0$ har vi inklusjonen

$$\{\|(X_n,Y_n)-(a,b)\| \geq \varepsilon\} \subset \{|X_n-a| \geq \varepsilon/2\} \cup \{|Y_n-b| \geq \varepsilon/2\},$$

og utled fra dette at $\|(X_n,Y_n)-(a,b)\| \to_p 0$.

(b) Anta at den reelle funksjonen $g(x,y)$ er kontinuerlig i (a,b). Vis at $g(X_n,Y_n)$ konvergerer i sannsynlighet til $g(a,b)$.

(c) Forklar hvorfor (8) følger fra det du nå har vist.

(d) Generaliser til flere dimensjoner.

8. Den bivariate normalfordeligen (og tilfeldige vektorer)

En tilfeldig vektor X er på er en vektor $X = (X_1,\ldots,X_p)^{\mathrm{t}}$ der elementene X_i er tilfeldig variable. Forventningen til en tilfeldig vektor $X = (X_1,\ldots,X_p)^{\mathrm{t}}$ er

$$\mathrm{E}\, X = (\mathrm{E}\, X_1,\ldots,\mathrm{E}\, X_p)^{\mathrm{t}}$$

der vi antar at $\mathrm{E}\, X_i$ er endelig for alle i. På tilsvarende vis definerer vi forventningen til en tilfeldig matrise (en matrise der elementene er tilfeldige variable), ved at forventningen tas elementvis. Variansen til en tilfeldig vektor er definert ved

$$\mathrm{Var}(X) = \mathrm{E}\,(X-\mathrm{E}\, X)(X-\mathrm{E}\, X)^{\mathrm{t}},$$

som når X er en p-dimensjonal vektor, er en $p \times p$ matrise. For eksempel, med $p=2$, og $\mu = (\mu_1,\mu_2)^{\mathrm{t}} = (\mathrm{E}\, X_1, \mathrm{E}\, X_2)^{\mathrm{t}}$, så er

$$\begin{aligned}\mathrm{E}\,(X-\mu)(X-\mu)^{\mathrm{t}} &= \begin{pmatrix} \mathrm{E}\,(X_1-\mu_1)^2 & \mathrm{E}\,(X_1-\mu_1)(X_2-\mu_2) \\ \mathrm{E}\,(X_2-\mu_2)(X_1-\mu_1) & \mathrm{E}\,(X_2-\mu_2)^2 \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1,X_2) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Var}(X_2) \end{pmatrix}\end{aligned}$$

I forelesning var det spørsmål om vi kan vite at denne matrisen er invertibel. Fra lineær algebra husker vi at en symmetrisk matrise A er positiv definitt dersom $w^{\mathrm{t}}Aw > 0$ for alle vektorer w der minst ett element er ulikt null; og at en positiv definitt matrise er invertibel.

I to dimensjoner har vi at (sjekk det)

$$(w_1,w_2)\begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1,X_2) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Var}(X_2) \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \mathrm{Var}(w_1X_1+w_2X_2),$$

og vi vet at $\mathrm{Var}(w_1X_1+w_2X_2) > 0$ så lenge $w_1X_1+w_2X_2$ ikke er konstant. Dette kan du generalisere til flere dimensjoner.

Momentgenererende funksjoner er nyttige også i flere dimensjoner. Som i én dimensjon, er det slik at en mgf karakteriserer en fordeling (når mgfen finnes). Dersom $Y = (Y_1, \ldots, Y_p)^{\mathrm{t}}$ er en tilfeldig vektor, er dens mgf

$$M_Y(t_1, \ldots, t_p) = \mathrm{E}\, \exp(t^{\mathrm{t}} Y) = \mathrm{E}\, \exp\big(\sum_{j=1}^{p} t_j Y_j\big).$$

At mgfen karakteriserer en fordeling betyr at dersom de tilfeldige vektorene X og Y har samme mgf, så har de samme fordeling.

8.1. **Normalfordelte vektorer.** Den tilfeldige vektoren $X = (X_1, \ldots, X_n)$ er multivariat normalfordelt med forventning $\mu = (\mu_1, \ldots, \mu_p)^{\mathrm{t}}$ og varians Σ dersom dens fordeling har simultantetthet

$$f(x_1, \ldots, x_p) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\{-\tfrac{1}{2}(x-\mu)^{\mathrm{t}}\Sigma^{-1}(x-\mu)\}, \quad x \in \mathbb{R}^p. \tag{9}$$

Vi skriver $X \sim \mathrm{N}_p(\mu, \Sigma)$ for å si at X er multivariat normalfordelt med forventning μ og varians Σ. Når $X \sim \mathrm{N}_p(\mu, \Sigma)$, så er dens mgf

$$M_X(t) = \exp(t^{\mathrm{t}}\mu + \tfrac{1}{2}t^{\mathrm{t}}\Sigma t), \quad t \in \mathbb{R}^p.$$

En bivariate normalfordeling er tilfellet $X \sim \mathrm{N}_2(\mu, \Sigma)$, som vi kan skrive

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathrm{N}_2\Big(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\Big). \tag{10}$$

Her er $\sigma_1^2 = \mathrm{Var}(X_1)$ og $\sigma_2^2 = \mathrm{Var}(X_2)$, mens ρ er korrelasjonen mellom X_1 og X_2, det vil si $\rho = \mathrm{Cov}(X_1, X_2)/(\sigma_1\sigma_2)$, og $|\rho| \leq 1$. Som en liten oppgave i matriseregning, kan du bruke den generelle formen for tettheten i (9) for å finne tettheten til en bivariat normal.

9. Oppgaver

Oppgave 9.1 La $X = (X_1, \ldots, X_p)^{\mathrm{t}}$ være en tilfeldig vektor. La A være en $q \times p$ matrise av konstanter (en vanlig, ikke-tilfeldig matrise). Vis at

$$\mathrm{E}\,(AX) = A\mathrm{E}\,(X), \quad \text{og at} \quad \mathrm{Var}(AX) = A\mathrm{Var}(X)A^{\mathrm{t}}.$$

Vis også at $\mathrm{Var}(\sum_{j=1}^{p} a_j X_j) = a^{\mathrm{t}}\mathrm{Var}(X)a$.

Oppgave 9.2 Vi vet at uavhengige variable har kovarians null, men at det motsatte ikke nødvendigvis er sant. For multivariat normalfordelte variable er det sant. La $f(x_1, \ldots, x_p)$ være tettheten i (9), og la $f_j(x)$ være tettheten til en $\mathrm{N}(\mu_j, \sigma_j^2)$ tilfeldig variabel. Sjekk at når Σ er en diagonal matrise med $\sigma_1^2, \ldots, \sigma_p^2$ på diagonalen, så er

$$f(x_1, \ldots, x_p) = f_1(x_1) \cdots f_p(x_p).$$

Oppgave 9.3 Anta at $X \sim \mathrm{N}_p(\mu, \Sigma)$, og la A være en konstant $q \times p$ matrise slik at $A\Sigma A^{\mathrm{t}}$ er positiv definitt. Vis at $Y = AX$ er normalfordelt.

Oppgave 9.4 Anta at $X \sim \mathrm{N}_p(\mu, \Sigma)$. La $a_1, \ldots, a_p$ være konstanter (minst en $a_j \neq 0$). Vis at $\sum_{j=1}^{p} a_j X_j$ er normalfordelt. Andre veien. Vis at dersom $\sum_{j=1}^{p} a_j X_j$ er normalfordelt for enhver $(a_1, \ldots, a_p)$ med minst en $a_j \neq 0$, så er $(X_1, \ldots, X_p)^{\mathrm{t}}$ multivariate normalfordelt.

Oppgave 9.5 La Z_1, Z_2 være uif standarnormalfordelte tilfeldige variable.

(a) Vis at $(Z_1, Z_2)^{\mathrm{t}} \sim N_2(0, I_2)$, der I_2 er 2×2 identitetsmatrisen.

(b) La $\rho \in (-1, 1)$. Vis at

$$\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathrm{N}_2\Big(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \Big),$$

og finn en transformasjon av $(W_1, W_2)^{\mathrm{t}}$ som gir en tilfeldig vektor $(X_1, X_2)^{\mathrm{t}}$ med den bivariate normalfordelingen i (10).

Oppgave 9.6 La $(W_1, W_2)^{\mathrm{t}}$ være bivariat normalfordelt med forventning $(0, 0)$, $\mathrm{Var}(W_j) = 1$ for $j = 1, 2$, og $\mathrm{Cov}(W_1, W_2) = \rho$.

(a) Vis at marginalt er $W_1 \sim \mathrm{N}(0, 1)$, og $W_2 \sim \mathrm{N}(0, 1)$.

(b) Vis at $W_1 \,|\, W_2 = w_2 \sim \mathrm{N}(\rho w_2, 1 - \rho^2)$.

(c) La $(X_1, X_2)^{\mathrm{t}}$ være bivariate normalfordelt som i (10). Vis at

$$X_1 \,|\, X_2 = x_2 \sim \mathrm{N}\big(\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1 - \rho^2)\big).$$

Referanser

Devore, J. L., Berk, K. N., and Carlton, M. A. (2021). *Modern Mathematical Statistics with Applications*. Springer Nature, Cham, Switzerland.

Matematisk institutt, Universitetet i Oslo
Email address: emilas@math.uio.no