Løsningsforslag øving 3

3.1

Tettheter sier noe om hvor konsentrert en viss egenskap er i et område, dvs. areal, volum, ... Massetetthet sier da noe om hvor mye masse som er konsentrert i en volumenhet, ladningstetthet sier noe om hvor mye ladning som er konsentrert i en volumenhet, og befolkningstetthet sier noe om hvor mange folk som er konsentrert i en arealenhet. Sannsynlighetstetthet har ikke en like klar definisjon, selv om vi kan si at den beskriver hvor stor sannsynlighet som er konsentrert langs en linje eller i et areal eller volum osv. Sannsynligheten for en hendelse er for enkelte sannsynlighetsfunksjoner da lik integralet eller summen av sannsynlighetstettheten over hele hendelsen. Måleenheten til en sannsynlighetstetthet kommer an på hvilke utfallsrom den er definert over. Generelt vil dimensjonen for en sannsynlighetstetthet være 1/enhet for en kontinuerlig stokastisk variabel som har måleenhet 'enhet'. Dette er akkurat som for andre tettheter. Forskjellen er at telleren har dimensjon 1 i tilfellet med sannsynlighetstettheter fordi sannsynlighet har måleenhet 1 (versus masse som har enhet 1kg etc.). I f.eks et eksempel hvor vi kaster pil og X, Y er koordinater som måles med lengdeenhet meter, så vil sannsynlighetstettheten ha dimensjon 1/m^2.

3.5

Den empiriske fordelingen er ikke alltid uniform. Hvis vi f.eks antar at den som kaster på blinken i eksempel 3.2 er svært dårlig til å kaste, så har han eller hun kanskje bommet mange ganger på blinken. Da vil den empiriske fordelingen vise mye større sannsynlighet for å bomme enn for å treffe innenfor et spesifikt område på selve blinken. Den empiriske fordelingen vil alltid være diskret, siden vi har endelig mange data. Utfallsrommet til en diskret fordeling trenger ikke alltid å være diskret. Hvis vi f.eks definerer hendelsene $A_i = \{(x, y) : i - 1 \leq x^2 + y^2 < i\}$ fører dette til en diskret fordeling, selv om utfallsrommet er kontinuerlig. Det har ikke noe å si for utfallsrommet om den diskrete fordelingen er uniform.

3.6

Vi ser fra likning (3.12) at normalfordelingen avhenger av de to parametrene μ og σ. Den uniforme fordelingen avhenger av b og a, og eksponensialfordelingen avhenger av parameteren λ. En sannsynlighetstetthet kan ikke være negativ. Dette vil kunne føre til negative sannsynligheter, som ikke er lov. Derfor må parameteren σ i normalfordelingen være positiv. Sannsynlighetsfordelingen må også integreres til 1, som betyr at verken μ eller σ kan være lik uendelig. For den uniforme fordelingen må $b \geq a$ for at fordelingen skal være positiv. Heller ikke her kan parametrene ta verdien ∞. For eksponensialfordelingen må $\lambda > 0$, slik at fordelingen skal være positiv.

3.7

For å vise at dette er en tetthet må vi summere over alle verdier av n og vise at summen er lik 1. Vi kan se at f alltid er positiv når $\lambda \geq 0$. Vi regner ut

$$\sum_{n=0}^{\infty} f(n) = \sum_{n=0}^{\infty} \frac{1}{1+\lambda}(\frac{\lambda}{1+\lambda})^n.$$

Vi kjenner igjen dette som en geometrisk rekke. Siden $\lambda \geq 0$ har vi at $\lambda/(1+\lambda) < 1$, som betyr at

$$\sum_{n=0}^{\infty} \frac{1}{1+\lambda}(\frac{\lambda}{1+\lambda})^n = \frac{1}{1+\lambda}\frac{1}{1-\frac{\lambda}{1+\lambda}} = \frac{1}{1+\lambda}\frac{1+\lambda}{1} = 1.$$

Funksjonen f kan også bare ta tellbart mange forskjellige verdier. Derfor er f en diskret tetthet. Det er oppgitt at utfallsrommet er R. Dette er ikke et diskret utfallsrom. Det hadde derimot vært mulig å definere tettheten for et diskret utfallsrom, men det er ikke gjort i dette tilfellet. Hvis $\lambda = 0$ får vi $f(n) = 1$ hvis $n = 0$ og $f(n) = 0$ hvis $n > 0$. At $f(n) = 0$ for $n > 0$ er åpenbart. At $f(n) = 1$ hvis $n = 0$ er ikke like klart, men dette kan sees fra likningen

$$\lim_{\lambda \to 0} f(0) = \lim_{\lambda \to 0} \frac{1}{1+\lambda}\left(\frac{\lambda}{1+\lambda}\right)^0 = \lim_{\lambda \to 0} \frac{1}{1+\lambda} \cdot 1 = 1.$$

3.13

Vi kan bruke sannsynlighetstettheten for en eksponensialfordeling. Sannsynligheten for at det er mellom 50 og 100 dager mellom to gruveulykker blir da

$$P(50 \leq x \leq 100) = \int_{50}^{100} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx = -e^{-\frac{x}{\lambda}}\Big|_{50}^{100} = e^{-\frac{50}{\lambda}} - e^{-\frac{100}{\lambda}}.$$

Vi setter inn $\lambda = 241$ og får $P(50 \leq x \leq 100) = 0.15$.

4.1

Vi har $Q(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}$. Det er tre krav for å være en sannsynlighet. Vi sjekker alle tre.

1. Sannsynligheten til en tellbar disjunkt mengde av hendelser er lik summen av alle hendelsene. For disjunkte hendelser har vi at $P((A_1 \cup A_2) \cap B) = P((A_1 \cap B) \cup (A_2 \cap B)) = P(A_1 \cap B) + P(A_2 \cap B)$. Dette kan generaliseres til tellbart antall disjunkte hendelser, som vil si at $Q(A_1 \cup A_2 \cup \ldots) = Q(A_1) + Q(A_2) + \ldots$.

2. Vi har at $\Omega \cap B = B$ for alle mengder B. Da er $Q(\Omega) = P(B)P(B) = 1$.

3. $Q(A)$ må være positiv for alle A, siden den er definert som to positive sannsynligheter delt på hverandre.

Alle utfallsrom som inneholder settet B vil være naturlige for Q. Dette inkluderer det originale utfallsrommet Ω, og det minste mulige utfallsrommet B. Hvis vi bruker B som utfallsrom ser vi at Q kan ansees som en helt standard sannsynlighet.

4.2

For å finne sannsynligheten for sol bruker vi loven om total sannsynlighet (se f.eks eksempel 4.6). Vi får P(sol) = P(sol ∩ varslet sol) + P(sol ∩ varslet overskyet) + P(sol ∩ varslet regn) = 40%.

Sannsynligheten for regn dersom det er varslet sol kan skrives P(regn | varslet sol) = P(regn ∩ varslet sol) / P(varslet sol). Vi regner ut sannsynligheten for at det er varslet sol på samme måte som sannsynligheten for at det skal bli sol og får P(regn | varslet sol) = 10% / 44% = 23%.

Instituttet gjetter rett når det er samme vær som varslet. Dette skjer 30% + 20% + 20% = 70% av tiden. Da gjetter de galt 30% av tiden.

For å lage en slik tabell kunne man gått over n forskjellige værvarsler hvor man også kjenner til hvordan været ble. Så kan man rett og slett telle hvor ofte hver av de 9 hendelsene i tabellen inntreffer.

4.9

Sannsynligheten for å få dobbel 6'er i ett kast er lik 1/36, som betyr at sannsynligheten for å ikke få dobbel 6'er er lik 35/36. Hendelsen at vi får minst en dobbel 6'er på n kast er komplementet til at vi ikke får en eneste dobbel 6'er på n kast. Spørsmålet kan derfor endres til å spørre om hvor mange ganger man må kaste for at sannsynligheten for ingen doble 6'ere er mindre enn 1/2. Siden hvert kast er uavhengig er sannsynligheten for å ikke få noen doble 6'ere på n kast det samme som $(35/36)^n$. Vi får likningen

$$(35/36)^n < 1/2,$$

som gir

$$\begin{aligned} e^{n \log 35/36} &< 1/2 \\ \implies n \log 35/36 &< \log 1/2 \\ \implies n > (\log 1/2)/(\log 35/36) &\approx 24.6. \end{aligned}$$

Dette betyr at vi må gjennomføre 25 kast.

4.13

Ved å anta uavhengighet har vi at P(x_1 = mynt, x_2 = mynt, ..., x_9 = mynt, x_{10} = krone) = P(x_1 = mynt) · P(x_2 = mynt) · ... · P(x_9 = mynt) · P(x_{10} = krone). Vi antar at mynten er symmetrisk, som betyr at sannsynligheten for å få mynt er lik sannsynligheten for å få krone er lik 1/2. Da har vi at sannsynligheten for at du kaster 10 ganger før du får krone er $(1/2)^{10}$. Det er også mulig å tenke seg at absolutt alle utfall etter 10 kast er like sannsynlige. Det totale antallet kombinasjoner av myntutfall er lik 2^{10}. Det betyr at sannsynligheten for et hvilket som helt utfall er lik $(1/2)^{10}$.

4.15

Hvis A og B er uavhengige har vi P(A ∩ B) = P(A) P(B). Vi bruker loven om total sannsynlighet og får

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= P(A)P(B) + P(A|B^c)P(B^c) \\ &= P(A)P(B) + P(A|B^c)(1 - P(B)) \end{aligned}$$

$$\begin{aligned} &\implies P(A)(1 - P(B)) = P(A|P(B^c))(1 - P(B)) \\ &\implies P(A) = P(A|B^c). \end{aligned}$$

Dette betyr at A og B^c er uavhengige.