

Obligatorisk innlevering 1 i STAT110/622

Basert på kapittel 1-3.3 i boken. Les informasjon “Obligatoriske innleveringer” på MittUiB->Hjem. For noen oppgaver er det oppgitt at de er basert på en tidligere eksamensoppgave. Hvis du trenger hjelp kan du kikke på disse, men legg merke til at oppgavene **er endret**. Du får ingen poeng hvis bare skriver rett av løsningen av eksamensoppgaven uten å vise at du skjønner hvordan du må endre oppgaven.

1. Inntektsfordeling i befolkningen (Modifisert eksamen Juni 2021, oppg. 2a)

Vi ser på inntekt x (målt i millioner) i en befolkning. Tabellen under viser inntektsfordelingen for et utvalg på $n = 8000$ personer.

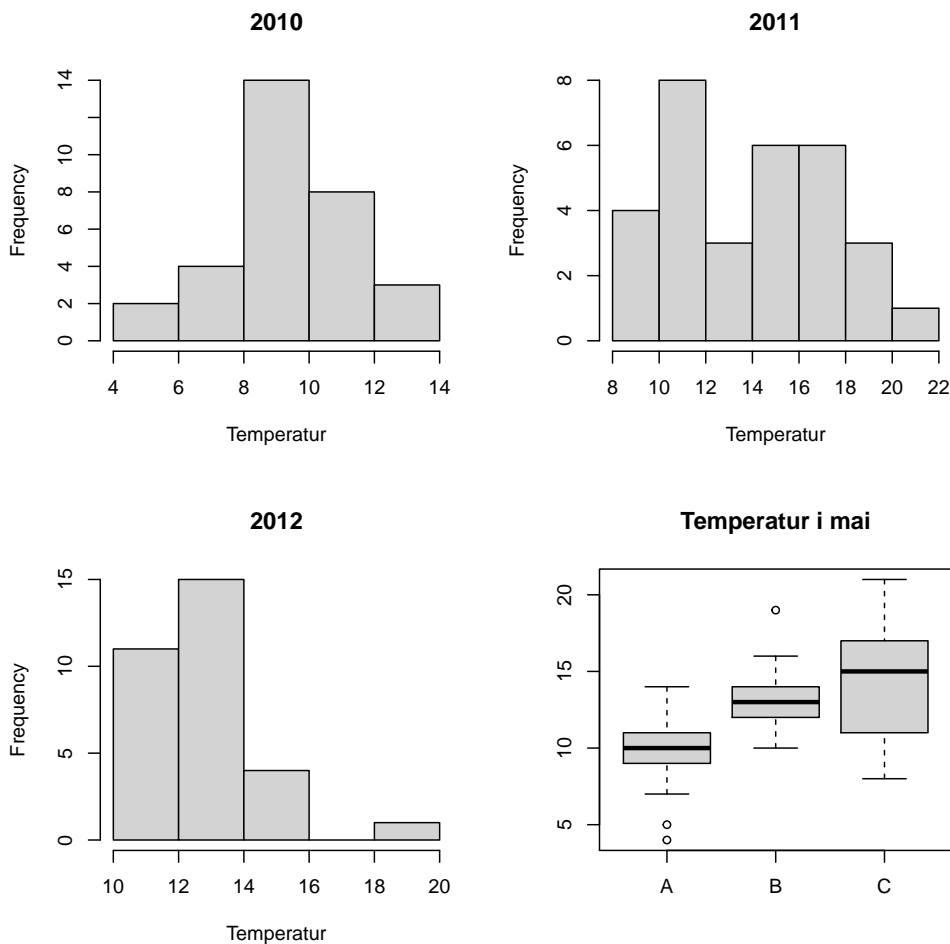
	Midtp.	Frek	Kumulativ
(0.0,0.1]	0.05	1	1
(0.1,0.2]	0.15	35	36
(0.2,0.3]	0.25	230	266
(0.3,0.4]	0.35	609	875
(0.4,0.5]	0.45	1029	1904
(0.5,0.6]	0.55	1226	3130
(0.6,0.7]	0.65	1247	4377
(0.7,0.8]	0.75	1090	5467
(0.8,0.9]	0.85	858	6325
(0.9,1.0]	0.95	596	6921
(1.0,1.1]	1.05	426	7347
(1.1,1.2]	1.15	300	7647
(1.2,1.3]	1.25	167	7814
(1.3,1.4]	1.35	96	7910
(1.4,1.5]	1.45	37	7947
(1.5,1.6]	1.55	26	7973
(1.6,1.7]	1.65	13	7986
(1.7,1.8]	1.75	8	7994
(1.8,1.9]	1.85	4	7998
(1.9,2.0]	1.95	2	8000

Tabellforklaring:

- Kolonnen til venstre (uten navn) viser inntekt (i millioner) som et intervall (fra,til].
- Midtp.: Midtpunktet i inntektsintervallet. Eksempel: intervallet $(0.0, 0.1]$ har midtpunkt 0.05.
- Frek: absolutt frekvens. Kolonnesum = 8000.
- Kumulativ viser kumulativ frekvens dvs. antall observasjoner som er mindre enn eller lik den aktuelle x -verdien (f.eks. $266 = 1 + 35 + 230$).

- (a) Hvor stor andel av befolkningen tjener mindre eller lik 1 million?
- (b) Hvis vi lager et histogram med tetthetsskala på y -aksen, hva er verdien (høyden) av histogrammet i intervallet $(1.3, 1.5]$?
- (c) Hva er gjennomsnittsinntekt blant de som tjener mer enn 1.5 millioner (1.5 ikke inkludert)? (Her kan du bruke midtpunktet i intervallene i utregningen)

2. **Fordelinger** Figuren under viser histogrammer over dagstemperaturer i mai for årene 2010-2012 (ikke virkelige data). Du ser også et boksplott som sammenlikner de tre årene, men rekkefølgen er stokket om tilfeldig.



- (a) Hva er gjennomsnittstemperatur i 2010? (Bruk midtpunkt i søylene.) Hva er medianen i 2011?
 (b) Hvilke boksplott (A, B og C) svarer til hvilke år med data? Begrunn svaret.

3. **Oljeleting** (Modifisert eksamen H2016, oppg. 2)

Først en liten oppvarming. La D og C være to hendelser i et utfallsrom. Det oppgis at $P(D) = 0.4$, $P(C) = 0.5$ og $P(D \cup C) = 0.8$.

- (a) Er hendelsene D og C disjunkte? Er hendelsene D og C uavhengige?

Så til oljeleting. Vi tenker oss to oljefelt. Hendelsen B = “oljefunn på felt 1”, mens den komplementære hendelsen B' = “ingen olje på felt 1”. Tilsvarende er hendelsen A = “oljefunn på felt 2”, mens A' = “ingen olje på felt 2”. Vi får oppgitt at $P(B \cap A) = 0.07$, $P(B' \cap A) = 0.05$, $P(B \cap A') = 0.17$ og $P(B' \cap A') = 0.7$.

- (b) Finn sannsynligheten for olje på felt 1. Anta at man har funnet olje på felt 2. Hva er nå sannsynligheten for olje på felt 1? Anta at man har påvist at felt 2 ikke inneholder olje. Hva er nå sannsynligheten for olje på felt 1? Er hendelsene B og A uavhengige?

4. **Elektrisk krets**

En elektrisk kobling består av 6 komponenter (se figur 1), som virker uavhengig av hverandre. La A_i være begivenheten at komponent nr. i virker ($i = 1, \dots, 6$), og anta at disse har sannsynligheter $P(A_1) = 0.7$, $P(A_2) = 0.4$, $P(A_3) = 0.5$, $P(A_4) = 0.9$, $P(A_5) = 0.3$ og $P(A_6) = 0.6$.

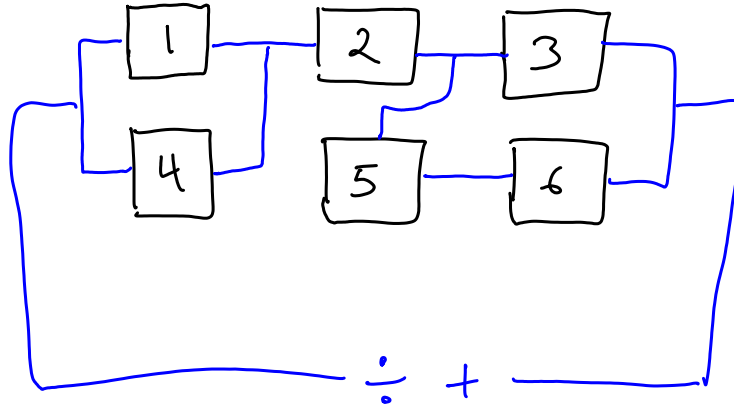


Figure 1: Kobling av elektriske komponenter (blå strek betyr ledning)

- (a) Vi definerer $B = A_1 \cup A_4$. Regn ut sannsynligheten for begivenheten B , dvs. $P(B)$.
 - (b) Definer $C = (A_5 \cap A_6) \cup A_3$. Regn ut $P(C)$.
 - (c) Finn sannsynligheten for at det går strøm i kretsen.
5. **Kuler i urne** En urne har 10 røde, 11 hvite og 13 gule baller. Trekk en ball tilfeldig fra urnen. La $X = -2$ hvis ballen er rød, la $X = 0$ hvis den er hvit og la $X = 2$ hvis den er gul. Finn følgende:
- (a) Punktsannsynlighetsfunksjonen til X .
 - (b) Forventningen til X .
 - (c) Variansen til X .

6. Studenter som spiller kort

Fire studenter sitter rundt et kvadratisk bord.

- (a) På hvor mange måter kan de sette seg? Hva er sannsynligheten for at de to eldste sitter rett overfor hverandre? Hva er sannsynligheten for at de sitter ved siden av hverandre?

Hver student har fått utdelt 5 kort av en kortstokk som består av 52 kort.

- (b) Hva er sannsynligheten for at det blant de fem først utdelte kortene befinner seg akkurat to ess? Skriv opp en formel for antall mulige kortfordelinger når det deles ut fem kort til hver student. Det kreves ikke et numerisk svar.

7. Forventning Følgende tabell viser sannsynlighetsfordelingen til den tilfeldige variabelen X :

x	1	2	3	4
$p(x)$	0.4	0.1	0.3	0.2

der $p(x) = P(X = x)$.

- (a) Finn forventning $E(X)$ og varians $V(X)$.
- (b) Finn $E[X/(1+X)]$, dvs. $E[h(X)]$ der $h(x) = x/(1+x)$. (Hint: se side 115 i læreboken.)
8. **Skjelden sykdom** En skjedent forekommende genvariant a disponerer for en sykdom. Den relative frekvensen av individer som er bærere av to kopier av genvarianten (har genotype aa) er 0.0001, mens frekvensen av bærere av én kopi (genotype Aa) er 0.0198 og frekvensen av de som ikke har a i det hele tatt (dvs har genotypen AA) er 0.9801. Selv om du er bærer av a er det ikke sikkert du får sykdommen, og du kan faktisk få sykdommen selv om du ikke har a . Sannsynligheten for å ha sykdommen blant personer med genotype aa , aA og AA er henholdsvis 0.6, 0.02 og 0.01.
- (a) Hva er sannsynligheten for at et tilfeldig valgt individ har sykdommen?
- (b) Gitt at du har sykdommen, hva er sannsynlighetene for du har genotype aa ? Finn de tilsvarende sannsynlighetene for genotypene Aa og AA .