Oblig 1 - stk 1100

Kevin Alexander Aslesen

Oppgave 1

a) Vi har at hver person kan gå av på etasjene i mengden $E = \{2., 3., 4., \ldots, 12.\}$, totalt 11 forskjellige etasjer. Når de 6 personene skal velge etasje, så har får vi et ordnet utvalg med tilbakelegging. Dette betyr at totalt antall mulige utfall tilsvarer $tot = 11 \cdot 11 \cdot 11 \cdot 11 \cdot 11 \cdot 11 = 11^6$.

Spørsmålet er så, i hvor mange av disse utfallene går de 6 personene av i hver sin etasje? Den første personen står fritt til å velge fra alle 11 etasjer, mens den andre personen kan ikke velge etasjen den første gikk av på og står igjen med 10 etasjer. Slik fortsetter det til den 6. personen.

Da har vi at antall utfall hvor personene går av i hver sin etasje er lik $_{11}P_6 = 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6$

Da får vi sannsynligheten:

$$P(\text{hver sin etasje}) = \frac{{}^{11}P_6}{11^6} = \underline{\underline{0.188}}$$

Bra!

b) $P(\text{minst 2 går av i samme etasje})$

$$= 1 - P(\text{max 1 går av i samme etasje})$$

$$= 1 - P(\text{hver sin etasje})$$

$$= 1 - 0.188$$

$$= \underline{\underline{0.812}}$$

Dette er det samme som hvis X = "antall som går av i samme etasje" så har vi:

$$P(X \geq 2) = 1 - P(X < 2)$$

$$= 1 - P(X \leq 1)$$

c) Personene er $\{I, II, III, IV, V, VI\}$

Hvor mange ulike grupper på 3 kan vi få fra disse?

For det må være antall kombinasjoner på 3:

$$_6C_3 = \binom{6}{3} = \frac{6!}{(6-3)!\,3!} = \frac{6!}{3!\,3!} = \underline{\underline{20}}$$

Enig!

d) P(alarmen fungerer)

$$= P((1 \cup 2) \cup (3 \cap (4 \cup 5)))$$

$$= \underbrace{P(1 \cup 2)}_{\mathrm{I}} + \underbrace{P(3 \cap (4 \cup 5))}_{\mathrm{II}} - \underbrace{P((1 \cup 2) \cap (3 \cap (4 \cup 5)))}_{\mathrm{III}}$$

tar hver for seg:

uavheng.

$$\mathrm{I}: \; P(1 \cup 2) = P(1) + P(2) - P(1 \cap 2)$$

$$= P(1) + P(2) - P(1)P(2)$$

Bra at du nevner denne antakelsen her - viktig!

$$= 0.9 + 0.9 - 0.9 \cdot 0.9 = 0.99$$

uavheng.

$$\mathrm{II}: \; P(3 \cap (4 \cup 5)) = P(3) \cdot P(4 \cup 5)$$

$$= P(3)\left(P(4) + P(5) - P(4)P(5)\right)$$

$$= 0.9\,(0.9 + 0.9 - 0.9 \cdot 0.9)$$

$$= 0.891$$

III : $P((1 \cup 2) \cap (3 \cap (4 \cup 5)))$

$$= P(1 \cup 2) \cdot P(3 \cap (4 \cup 5)) \qquad \text{uavheng.}$$

$$= 0.99 \cdot 0.891$$

$$= 0.882$$

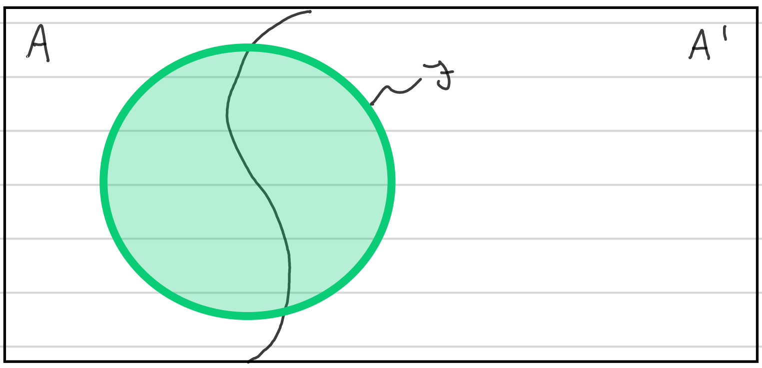
da får vi totalt :

$$P(\text{alarm fungerer}) = 0.99 + 0.891 - 0.882 = \underline{\underline{0.999}}$$

Bra!

Oppgave 2

Tenker at sammenhengene er noe lignende :



*not to scale

hvor A = "tekst er skrevet med AI"

og J = "programmet konkluder teksten som juks".

Da har vi følgende:

$$\cancel{P(J)} = 0.92, \quad P(A) = 0.05 \quad \text{og} \quad P(J \mid A') = 0.07$$

$P(J \mid A)$

a) Total sannsynlighet gir oss:

$$
\begin{aligned}
P(J) &= P(J \cap A) + P(J \cap A') \\
&= P(A)P(J \mid A) + P(A')P(J \mid A') \\
&= P(J)P(A \mid J) + (1 - P(A))P(J \mid A')
\end{aligned}
$$

løser for $P(A \mid J)$ og setter inn verdier:

$$
\begin{aligned}
P(J) - P(J)P(A \mid J) &= P(A')P(J \mid A') \\
1 - P(A \mid J) &= P(A')P(J \mid A') / P(J) \\
P(A \mid J) &= 1 - P(A')P(J \mid A') / P(J) \\
&= 1 - 0.95 \cdot 0.07 / 0.92 \\
&= \underline{\underline{0.928}}
\end{aligned}
$$

Her må du beregne $P(J)$ – ellers riktig!

Hvis programmet konkluderer at tekstet er juks, så er altså sannsynligheten for at det stemmer 92.8 %.

b) Setter sanns. for feilaktig fastslåelse til variabelen x, altså $x = P(J \mid A')$. Setter denne variabelen i uttrykket fra a) og løser uliketen for x:

$$P(A \mid J) > 0.90$$

$$1 - \frac{0.95 \cdot x}{0.92} > 0.90$$

$$\frac{0.95x}{0.92} < 0.10$$

$$x < \frac{0.92}{0.95} \cdot 0.10 = \underline{\underline{0.0968}}$$ 2

Sanns. for at programmet feilaktig fastslår at teksten er juks må være mindre enn 0.0968 for at sanns. fra a) skal være mer enn 90%.

Oppgave 3

a) $F_X(x) = P(X \leq x) = 1 - P(X > x)$

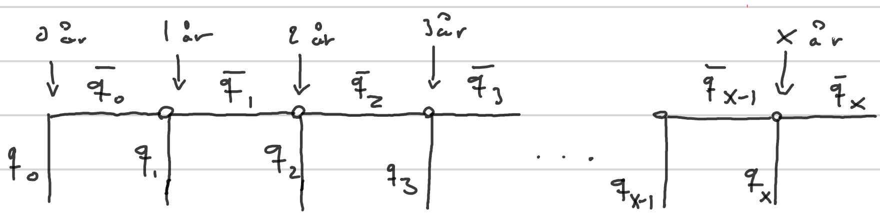
hvor

$P(X > x)$ = "sanns. for at en mann lever mer enn x år"

Denne sannsynligheten er produktet av sannsynlighetene for at mannen ikke dør de første x årene:

$$P(X>x) = (1-q_0)(1-q_1)\cdots(1-q_x)$$
$$= \prod_{z=0}^{x} (1-q_z)$$

Dette tilsvarer treet:



hvor $\bar q_x = 1 - q_x$ tilsvarer sanns. for at en x år gammel mann <u>ikke</u> dør ila. ett år.

fint!

b) Vi har at $p_X(x) = F_X(x) - F_X(x-1)$ ettersom $F_X(x)$ legger sammen sannsynlighetene frem til x mens $F_X(x-1)$ legger sammen alle før x. Vi får altså:

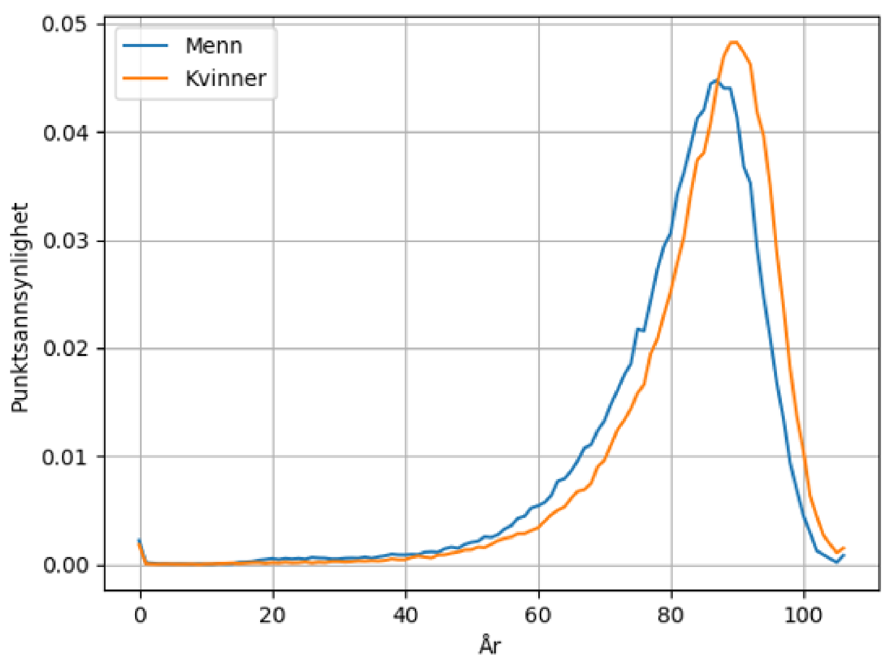
$$\underbrace{[\cancel{p_X(0)} + \ldots + \cancel{p(x-1)} + p_X(x)]}_{F_X(x)} - \underbrace{[\cancel{p_X(0)} + \ldots + p_X(\cancel{x-1})]}_{F_X(x-1)}$$

$$\Downarrow$$

$$p_X(x)$$

og tilsvarende er dermed $p_Y(y) = F_Y(y) - F_Y(y-1)$.

c) Plott av punktsannsynlighetene:



kode brukt :

```python
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

doed = pd.read_csv("https://www.uio.no/studier/emner/matnat/math/STK1100/data/doedelighet.txt",sep="\t")
alder= doed["alder"].values
menn = doed["menn"].values
kvinner = doed["kvinner"].values

qx_menn = menn/1000
qx_kvinner = kvinner/1000

Sx_menn = np.cumprod(1 - qx_menn)
Sx_kvinner = np.cumprod(1 - qx_kvinner)
Fx_menn = 1 - Sx_menn
Fx_kvinner = 1 - Sx_kvinner

px_menn = []
for i in range(len(Fx_menn)):
    if i > 0:
        px_menn.append(Fx_menn[i] - Fx_menn[i-1])
    else:
        px_menn.append(Fx_menn[i])

px_kvinner = []
for i in range(len(Fx_kvinner)):
    if i > 0:
        px_kvinner.append(Fx_kvinner[i] - Fx_kvinner[i-1])
    else:
        px_kvinner.append(Fx_kvinner[i])

plt.plot(px_menn, label="Menn")
plt.plot(px_kvinner, label="Kvinner")
plt.xlabel("År")
plt.ylabel("Punktsannsynlighet")
plt.grid()
plt.legend()
plt.show()
```

d) Hvis en mann er a år gammel, så har han maksimalt $106 - a$ år igjen å leve. Dermed vil $h(X) = X - a$ tilsvare gjenstående levealder for en tilfeldig valgt mann som er a år gammel. I tilfellet hvor $X < a$ så så må $h(X) = 0$ ettersom man ikke kan bli

50 år når man allerede er 55 år gammel.

Forventet gjenstående levealder er da $E(h(X))$

hvor $h(x) = \begin{cases} 0 & x < a \\ x-a & x \geq a \end{cases}$.

Bra!

e) Benytter følgende kode for beregningene:

```
def h(X, a):
    if (X < a):
        return 0
    else:
        return X - a

print("Forventet gjenstående levealder for menn:")
for a in [30, 50, 80]:
    E_hx = 0
    for x in range(0, 107):
        E_hx += h(x, a) * px_menn[x]
    print(f"a={a}: {E_hx:.2f} år")

Ex_menn = 0
for i in range(len(px_menn)):
    alder = i
    px = px_menn[i]
    Ex_menn += alder * px
print(f"Forventet levealder for menn: {Ex_menn:.2f} år")
```

og får følgende resultater:

```
Forventet gjenstående levealder for menn:
a=30: 50.88 år
a=50: 31.26 år
a=80: 5.24 år
Forventet levealder for menn: 80.73 år
```

Vi ser at antall år man er forventet å leve totalt

Denne oversikten var veldig fin!

$a=0$ $a=30$ $a=50$ $a=80$

øker med alderen: $80.73 < 80.88 < 81.26 < 85.24$.

Dette gir mening ettersom man ikke trenger å ta med seg usikkerheten å bli a år gammel når man er a år gammel. Usikkerheten med å dø ved yngere alder forsvinner.

Godt reflektert!

f) Benytter lik kode som i e) for beregningene:

```
print("Forventet gjenstående levealder for kvinner:")
for a in [30, 50, 80]:
    E_hx = 0
    for x in range(0, 107):
        E_hx += h(x, a) * px_kvinner[x]
    print(f"a={a}: {E_hx:.2f} år")

Ex_kvinner = 0
for i in range(len(px_kvinner)):
    alder = i
    px = px_kvinner[i]
    Ex_kvinner += alder * px
print(f"Forventet levealder for kvinner: {Ex_kvinner:.2f} år")
```

Resultatene:

```
Forventet gjenstående levealder for kvinner:
a=30: 54.21 år
a=50: 34.43 år
a=80: 7.14 år
Forventet levealder for kvinner: 84.11 år
```

Her øker også total forventet levealder når a øker som i e). I dette tilfellet har vi:

$a=0$ $a=30$ $a=50$ $a=80$

$84.11 < 84.21 < 84.43 < 87.14$.

Fra $a=0$ til $a=30$ er økningen på 0.15 hos menn og 0.10 hos kvinner.

Fra $a=30$ til $a=50$ er økningen på 0.38 hos menn og 0.22 hos kvinner.

Fra $a=50$ til $a=80$ er økningen på 3.98 hos menn og 2.71 hos kvinner.

Dette forteller oss at menn har større usikkerhet for å overleve ved yngre alder sammenlignet med kvinner.

↳ Dette klinger litt rart. Kunne du sagt at menn har større sannsynlighet for å tidlig, ref. plot over overlevelsessannsynligheter, slik at forskjellen i forventet gjenstående levealder blir stort menn!

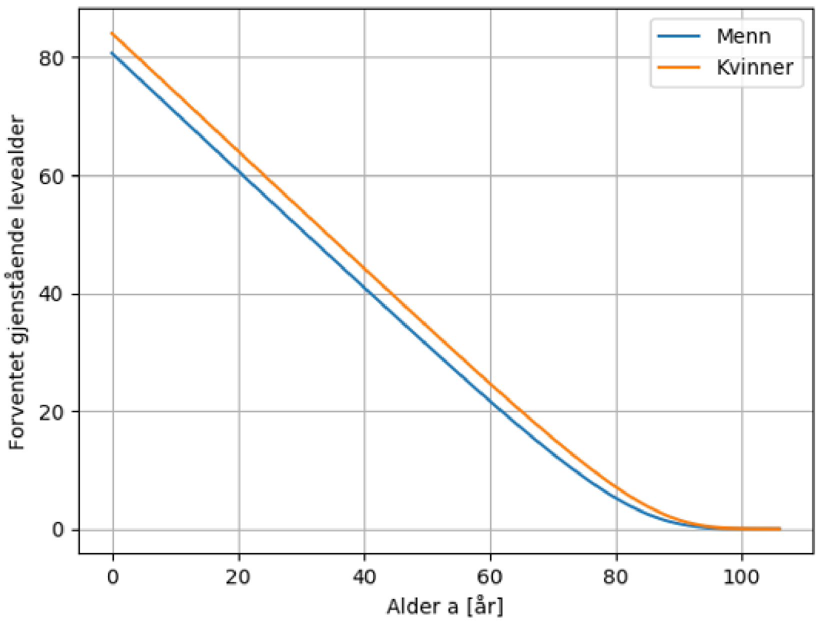
g) Kode brukt for å plotte forventet gjenstående levealder:

```
E_hx_menn_liste = []
E_hx_kvinner_liste = []

for a in range(0, 107):
    E_hx_menn = 0
    E_hx_kvinner = 0
    for x in range(0, 107):
        E_hx_menn += h(x, a) * px_menn[x]
        E_hx_kvinner += h(x, a) * px_kvinner[x]
    E_hx_menn_liste.append(E_hx_menn)
    E_hx_kvinner_liste.append(E_hx_kvinner)

a = np.arange(0, 107)
plt.plot(a, E_hx_menn_liste, label="Menn")
plt.plot(a, E_hx_kvinner_liste, label="Kvinner")
plt.xlabel("Alder a [år]")
plt.ylabel(f"Forventet gjenstående levealder")
plt.grid()
plt.legend()
plt.show()
```

Får følgende plott:



Forventningene blir mindre når a blir større og når $a \to 100$ så går $E(h(X)) \to 0$. Dette virker rimelig og stemmer med det vi så i e) og f).