LØSNINGER UKE 6, STK1100

SAMMENDRAG. Løsningsforslag for noen av de oppgavene jeg ikke rakk igjennom på plenumen. Originalt skrevet av Ingunn Fride Tvete. Send mail til `steffeng@math.uio.no` hvis du ser noe feil, så blir det rettet opp.

EKSTRAOPPGAVE 5

a) Sannsynligheten for at en 75 år gammel kvinne skal bli minst 80 år

$= P(\text{bli 76 år})P(\text{bli 77 år}|\text{blitt 76})P(\text{bli 78 år}|\text{blitt 77 år})$
$P(\text{bli 79 år}|\text{blitt 78 år})P(b\text{li 80 år}|\text{blitt 79 år})$
$= (1-26.17/1000)(1-28.45/1000)(1-30.12/1000)(1-37.74/1000)(1-41.57/1000)$
$= 0.8462899 \approx 0.846$

b) Sannsynligheten for at en 75 år gammel kvinne skal bli nøyaktig 80 år

$= P(\text{bli 80 år})P(\text{død før 81 år}|\text{blitt 80 år})$
$= 0.8462899 * 0.0441 = 0.03732138 \approx 0.037$

EKSTRAOPPGAVE 6

Vi definerer noen begivenheter:

- S HIV positiv
- F HIV negativ
- $+$ person testet positiv
- $-$ person testen negativ

Vi vet at $P(+|S) = 0.98$ og $P(+|F) = 0.002$.
Vi har en høyrisikogruppe der 10 % er HIV positive. Et tilfeldig valgt individ blir testet, og vi har at $P(S) = 0.10$ og $P(F) = 0.90$.

a) Sannsynligheten for at testen vil vise positiv:
$P(+) = P(+ \cap S) + P(+ \cap F) = P(+|S)P(S) + P(+|F)P(F)$
$= 0.98 * 0.10 + 0.002 * 0.90 = 0.0998.$

b) Vi antar at testen viser positiv. Den betingede sannsynligheten for at personen virkelig er smittet er gitt ved:

$$P(S|+) = \frac{P(+\cap S)}{P(+)} = \frac{P(+|S)P(S)}{P(+)} = \frac{0.98*0.10}{0.0998} = 0.9820$$

c) Vi har en lavrisikogruppe der 1 av 10000 er HIV positive, slik at $P(S) = 0.0001$. Et tilfeldig valgt individ blir testet og tester positivt. Den betingede sannsynligheten for at personen virkelig er smittet er gitt ved:

$$P(S|+) = \frac{P(+|S)P(S)}{P(+)} = \frac{0.98*0.0001}{0.98*0.0001+0.002*0.9999} = 0.0468$$

Ekstraoppgave 7

Lar X=antall gutter i en tolvbarnsfamilie.

a) Vi antar at X er binomisk fordelt med $n = 12$ og $p = 0.5$. Vi antar at vi har 12 uavhengige forsøk, hver med *suksess sannsynlighet* 0.5 .

Vi skriver i Matlab:

```
x=0:12;
bingutter1=binopdf(x,12,0.5);
```

og får punktsannsynlighetene:

0.0002 0.0029 0.0161 0.0537 0.1208 0.1934
0.2256 0.1934 0.1208 0.0537 0.0161 0.0029 0.0002

Vi kan sjekke hva den relative frekvensen av guttefødsler i de 6115 familiene var:

I Matlab:

```
totbarn= 6115*12;
totgutter=sum(x.*familier);
andelgutter=totgutter/totbarn
```

Som gir svaret andelgutter= 0.5192 $\approx$ 0.52.

b) Vi antar at X er binomisk fordelt med $n = 12$ og $p = 0.52$. Vi antar at vi har 12 uavhengige forsøk, hver med *suksess sannsynlighet* 0.52 .

Vi skriver i Matlab:

```
x=0:12;
bingutter2=binopdf(x,12,0.52);
```

og får punktsannsynlighetene:

0.0001 0.0019 0.0116 0.0418 0.1020 0.1768
0.2234 0.2075 0.1405 0.0676 0.0220 0.0043 0.0004

c) Vi beregner de relative frekvensene av familier med k gutter for $k = 0, 1, 2, ..., 12$

familier=[3 24 104 286 670 1033 1343 1112 829 478 181 45 7];

totantfam=sum(familier);

relfrek=familier/totantfam;

og får de relative frekvensene:

0.0005 0.0039 0.0170 0.0468 0.1096 0.1689
0.2196 0.1818 0.1356 0.0782 0.0296 0.0074 0.0011

c) og d): Vi kan plotte resultatene fra a)-c), og en mulig måte å gjøre dette på i Matlab er:

```
plot(relfrek)
hold on
plot(bingutter1,'–')
plot(bingutter2,':')
hold off
legend('relativ frekvens','binopdf(X,12,0.5)','binopdf(X,12,0.52)')
title('Gutter i tolvbarsnsfamilier.')
xlabel('Antall gutter')
ylabel('Punktsannsynlighet')
```

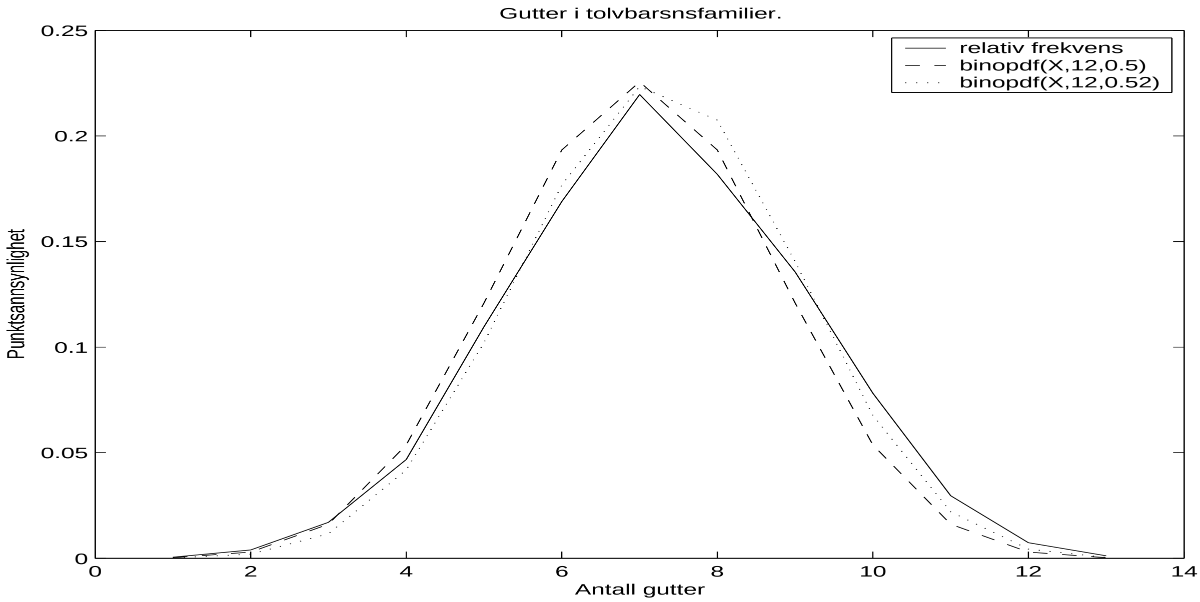
Vi får da følgende figur:

Vi ser av figuren at både modellen i a) og b) ikke er så aller verst i forhold til de relative frekvensene (heltrukket linje). Bortsett fra x-verdier mellom 7 og 9 ser modellen i b) ut til å være litt bedre enn modellen i a) for å beskrive vårt datamateriale.

Vi har i a) og b) anntatt at guttenefødslene i en familie er uavhengige begivenheter. Man kan kanskje spørre seg om dette er en rimelig antagelse. Det er vel ikke urimelig å tro at noen av disse guttene er tvillinger. Ja, kanskje til om med eneggede tvillinger. Da vet man jo at man ikke har uavhengige begivenheter. I tillegg er det mulig at sannsynligheten for guttefødsel kan variere noe fra familie til familie. Vi kan se denne (svake) avhengigheten i figuren, da vi i vårt datamateriale har litt flere familier med mange barn av samme kjønn enn det modellene våre skulle tilsi.

Oppgave 1.8.50

Vi kaster to terninger, og summen skal være 6. Hva er sannsynligheten for at minst en av terningere var 3?



Figur 1. Venn diagram

Antall mulige måter å få sum = 6 på: $(5,1),(4,2),(3,3),(2,4),(1,5)$. Det er en mulig måte å få sum = 6 på når minst en av terningere er 3, nemlig $(3,3)$. Dette gir den betingede sannsynligheten $1/5 = 0.2$.

Oppgave 1.8.65

Vi vet at A og B er uavhengige. Da er $P(A \cap B) = P(A)P(B)$, jmf. definisjonen på side 22 i læreboka. Hva med A og B^c?

$$P(A) = P(A \cap B) + P(A \cap B^c) = P(A)P(B) + P(A \cap B^c)$$

$$P(A \cap B^c) = P(A)(1 - P(B)) = P(A)P(B^c)$$

Så A og B^c er uavhengige.

Vi vet nå at A og B er uavhengige, og at A og B^c er uavhengige. Hva med A^c og B^c?

På lignende måte som ovenfor har vi at:

$$P(B^c) = P(B^c \cap A) + P(B^c \cap A^c) = P(B^c)P(A) + P(B^c \cap A^c)$$

Dette gir at
$P(B^c \cap A^c) = P(B^c)(1 - P(A)) = P(B^c)P(A^c)$

Så B^c og A^c er uavhengige.

Oppgave 1.8.77

Sannsynligheten for å treffe *Bulls Eye* $= P(BE) = 0.05$,
som gir at sannsynligheten for ikke å treffe *Bulls Eye* $= P(BE^c) = 0.95$.
Hvor mange ganger må det kastes for at sannsynligheten for å treffe minst en gang er 0.5?

Vi har at:
P(ikke treffe *Bulls Eye* minst en gang i n forsøk) $= 0.95^n$, og videre
P(treffe minst en gang i n forsøk) $= 1 - 0.95^n$. Vi vil at denne sannsynligheten skal være 0.5, og løser vi ut med hensyn på n får vi $n = \frac{log(1-0.5)}{log(0.95)} = 13.51$, slik at det må kastes 14 ganger.

Dette kan også løses slik:
P(treffe minst 1 gang)
$= P$(treffe 1. gang eller treffe 2. gang .... eller treffe n.te gang)
$= 0.05 + 0.95 * 0.05 + 0.95 * 0.95 * 0.05 + ... + 0.95 * ... * 0.95 * 0.05$
$= \sum_{i=1}^{n} 0.05 * 0.95^{i-1} = 0.5$, og dermed at $\sum_{i=1}^{n} 0.95^{i-1} = 10$.

Fra formelsamling vet vi at $\sum_{k=1}^{n} q^{k-1} = \frac{1-q^n}{1-q}$ som gir i vårt tilfelle at $\frac{1-0.95^n}{1-0.95} = 10$ og løser vi ut med hensyn på n får vi, som ovenfor, at $n = 13.51$.

Oppgave 2.5.6

Vi kaster en rettferdig mynt 4 ganger, og vi skal finne frekvensfunksjonen (eller punktsannsynligheten) og kummulativ fordelingsfunksjon (cdf) for noen stokastiske variable. La oss først liste opp alle de mulige utfallene vi kan få:

$$\left\{ \begin{array}{llll} (kkkk) & (kkkm) & (kkmk) & (kmkk) \\ (mkkk) & (kkmm) & (kmmk) & (mmkk) \\ (kmkm) & (mkmk) & (mkkm) & (mmmk) \\ (mmkm) & (mkmm) & (kmmm) & (mmmm) \end{array} \right\}$$

a) X= antall kron før første mynt. Mulige verdier X kan ta er $0, 1, 2, 3$ og 4.

<u>Frekvensfunksjon</u>

$P(X = 0) = \frac{8}{16} = \frac{1}{2}$
(Vi har her mulighetene $(mkkk), (mmkk), (mkmk), (mkkm), (mmmk), (mmkm), (mkmm)$ og $(mmmm)$.)

$P(X = 1) = \frac{4}{16} = \frac{1}{4}$
(Vi har her mulighetene $(kmmm), (kmkm), (kmmk)$ og $(kmkk)$.)

$P(X = 2) = \frac{2}{16} = \frac{1}{8}$
(Vi har her mulighetene $(kkmm)$ og $(kkmk)$.)

$P(X = 3) = \frac{1}{16}$
(Vi har her muligheten $(kkkm)$.)

$P(X = 4) = \frac{1}{16}$
(Vi har her muligheten $(kkkk)$.)

<u>cdf</u>

$F(x) = P(X \leq x)$, og vi får
$P(X \leq 0) = P(X = 0) = \frac{1}{2}$
$P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$
$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$
$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}$
$P(X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$
$= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} = \frac{16}{16} = 1$

b) X= Antall kron etter første mynt. Mulige verdier X kan ta er $0, 1, 2$ og 3.

Vi tolker dette som antall kron som tilsammen kommer etter første mynt.

<u>Frekvensfunksjon</u>

$P(X = 0) = \frac{5}{16}$
(Vi har her mulighetene $(kkkk), (kkkm), (kkmm), (kmmm)$ og $(mmmm)$.)

$P(X = 1) = \frac{6}{16}$
(Vi har her mulighetene $(kkmk), (kmmk), (kmkm), (mmmk), (mmkm)$ og $(mkmm)$.)

$P(X = 2) = \frac{4}{16}$
(Vi har her mulighetene $(kmkk), (mmkk), (mkkm)$ og $(mkmk)$.)

$P(X = 3) = \frac{1}{16}$
(Vi har her muligheten $(mkkk)$.)

<u>cdf</u>

$P(X \leq 0) = \frac{5}{16}$
$P(X \leq 1) = \frac{5}{16} + \frac{6}{16} = \frac{11}{16}$
$P(X \leq 2) = \frac{5}{16} + \frac{6}{16} + \frac{4}{16} = \frac{15}{16}$
$P(X \leq 3) = \frac{5}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16} = \frac{16}{16} = 1$

c) X= Antall mynt - antall kron. Mulige verdier X kan ta er $-4, -2, 0, 2$ og 4.

<u>Frekvensfunksjon</u>

$P(X=-4)=\frac{1}{16}$
(Vi har her muligheten $(kkkk)$.)

$P(X=-2)=\frac{4}{16}$
(Vi har her mulighetene $(kkkm), (kkmk), (kmkk)$ og $(mkkk)$.)

$P(X=0)=\frac{6}{16}$
(Vi har her mulighetene $(kkmm), (kmmk), (mmkk), (kmkm), (mkmk)$ og $(mkkm)$.)

$P(X=2)=\frac{4}{16}$
(Vi har her mulighetene $(mmmk), (mmkm), (mkmm)$ og $(kmmm)$.)

$P(X=4)=\frac{1}{16}$
(Vi har her muligheten $(mmmm)$.)

<u>cdf</u>

$P(X\leq -4)=\frac{1}{16}$
$P(X\leq -2)=\frac{1}{16}+\frac{4}{16}=\frac{5}{16}$
$P(X\leq 0)=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}=\frac{11}{16}$
$P(X\leq 2)=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}=\frac{15}{16}$
$P(X\leq 4)=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}+\frac{1}{16}=\frac{16}{16}=1$

d) X= Antall kron*antall mynt. Mulige verdier X kan ta er $0, 3$ og 4.

<u>Frekvensfunksjon</u>

$P(X=0)=\frac{2}{16}$
(Vi har her mulighetene $(kkkk)$ og $(mmmm)$.)

$P(X=3)=\frac{8}{16}$
] (Vi har her mulighetene $(kkkm), (kkmk), (kmkk), (mkkk)$,
$(mmmk), (mmkm), (mkmm)$ og $(kmmm)$.)

$P(X=4)=\frac{6}{16}$
(Vi har her mulighetene $(kkmm), (kmmk), (mmkk), (kmkm), (mkmk)$ og $(mkkm)$.)

<u>cdf</u>
$P(X\leq 0)=\frac{2}{16}$
$P(X\leq 3)=\frac{2}{16}+\frac{8}{16}=\frac{10}{16}$
$P(X\leq 4)=\frac{2}{16}+\frac{8}{16}+\frac{6}{16}=\frac{16}{16}=1$