# STK1100 Oblig 1

Daniel Heinesen, daniehei

27. mars 2017

## Oppgave 1)

### a)

$$X = \text{Antall komponenter som testes frem til og med første defekte}$$

Vi ser at de første $x$ vi tester, må de første $x-1$ ikke være defekte. Sjansen for at komponentene ikke er defekte er $(1-p)$. Så sannsynligheten for at dette inntreffer er $(1-p)^{x-1}$. Den siste vi tester må så være defekt, som har en sannsynelighet $p$. Dette gir oss en punktsannsynlighet

$$p_X(x;p) = p(1-p)^{x-1}$$

Dette er en negativ binomisk fordelig, hvor antall suksesser er 1. Vi kan derfor også finne punktsannsynligheten gjennom denne fordelingn.

### b)

Den momentgenererende funksjonen for denne fordeligen er gitt ved

$$M_X(t) = E[e^{tX}] = \sum_x e^{tx} p_X(x;p) = \sum_x e^{tx} p(1-p)^{x-1}$$

Vi bruker så formelen for en uendelig geometrisk serie og får

$$M_X(t) = pe^t \sum_x e^{t(x-1)}(1-p)^{x-1} = \frac{pe^t}{1-e^t(1-p)}$$

Vi kan nå finne $E[X]$ og $E[X]$ ved hjelp av den momentgenererende funksjonen. Først finner vi:

$$M'_X(t) = \frac{pe^t}{(1-(1-p)e^t)^2}$$

$$M''_X(t) = \frac{pe^t(1+(1-p)e^t)}{(1-(1-p)e^t)^3}$$

Vi finner så at

$$E[X] = M'_X(0) = \frac{p}{(1-(1-p))^2} = \frac{1}{p}$$

$$E[X^2] = M''_X(0) = \frac{p(1+(1-p))}{(1-(1-p))^3} = \frac{2-p}{p^2}$$

Vi kan så regne ut variasjonen:

$$V[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

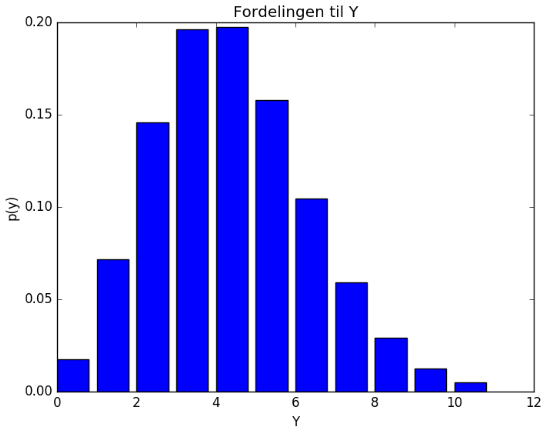
### c)

$$Y = \text{Antall defekte komponenter i et utvalg.}$$

Variablen over, $Y$, beskriver et eksperiment hvor vi har et antall forsøk hvor vi enten kan få suksess(defekt) eller ikke-suksess(ikke defekt).Hvorvidt komponentene er defekte eller ikke tar vi til å være uavhenging, og sannsynligheten for å finne en defekt komponent er $p$. Denne variablen forfyller alle kravene til en binomisk fordelig, og er derfor det. For denne oppgaven bruker vi at $p = 0.02$ og $n = 200$

### d)

Dette er programmert i python, og finnes i *opp1c.py*




Figur 1: Den binomiske fordelingen til Y

### e)

Fra side 134 i læreboken finner vi at

$$E[Y] = np = 200 \cdot 0.02 = \underline{\underline{4}}$$

og

$$V[Y] = np(1-p) = 4 \cdot (1 - 0.02) = \underline{\underline{3.92}}$$

### f)

En binomisk fordelig går mot en Poisson-fordeling når $n \to \infty$ og $p \to 0$. For vært eksperiment er vi lang unna, men som en tommelregel kan den binomiske fordelingen bli godt tilnærmet av Poisson-fordelingen når $n > 50$ og $np < 5$. Og siden vi har

$$n = 200 > 50$$

$$np = 4 < 5$$

Så er en Poisson-fordeling en god tilnærming til fordelingen til $Y$.

## g)

For en Poisson-fordeling

$$p(y; \lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$$

Hvor $\lambda$ er forventningsverdien til $Y$, som gjør at vi har

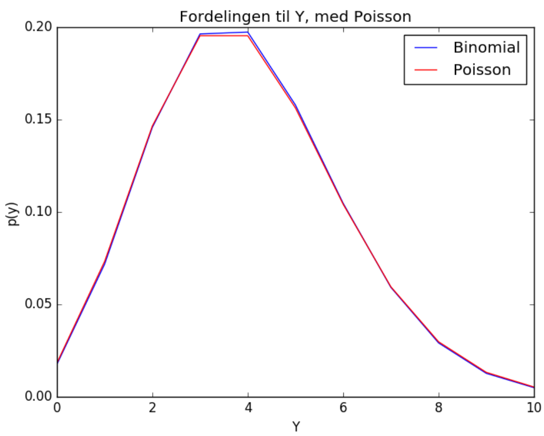
$$\lambda = np = 4$$

og

$$p(x) = \frac{e^{-4}4^y}{y!}$$

## h)

Dette er programmert i python, og finnes i *opp1c.py*




Figur 2: Jeg har nå valgt å plotte fordelingene som rene grafer, slik at det blir lettere å sammenlikne. Vi kan se at de ikke er helt identiske – som vi ikke forventer –, men tilnærmingen er veldig bra.

# Oppgave 2)

## a)

Vi har $X$ med sannsynlighetsfordelingen

$$f(x) = \begin{cases} \theta^{-1} x^{-\frac{\theta+1}{\theta}} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

Den kumulative fordelingen er gitt ved

$$F(x) = \int_0^x f(y) dy = \int_1^x \frac{y^{-\frac{\theta+1}{\theta}}}{\theta} dy$$

Dette reduseres til

$$F(x) = -y^{-1/\theta} \Big|_1^x = -x^{-1/\theta} - (-1) = 1 - x^{-1/\theta}$$

## b)

For $\theta = 0.45$ kan vi lett finne

$$P(2 \leq X \leq 5) = \int_2^5 f(x) dx = F(5) - F(2) = -5^{-1/0.45} + 2^{-1/0.45} = \underline{\underline{0.186}}$$

## c)

Medianen til en kontinuelig fordeling er det punktet $x$ hvor arealet til fordelingen er lik på hverside. Dette kan regnes ut på to måter

$$1) \qquad \int_1^x f(x) dx = \int_x^\infty f(x) dx$$

$$2) \qquad P(X \geq x) = P(X \leq x) = \frac{1}{2}$$

Det er lett å vise at disse ble det samme. Men den enkleste er 2). Fra den får vi

$$P(X \geq x) = F(x) = 1 - x^{-1/\theta} = \frac{1}{2}$$

$$\Rightarrow x = 2^\theta$$

Så $x = 2^\theta$ er medianen til fordelingen.

## d)

Vi ønsker så å finne forventningsverdien til fordelingen:

$$E[X] = \int_{-\infty}^\infty x f(x) dx = \int_1^\infty \frac{x^{-\frac{\theta+1}{\theta}+1}}{\theta} dx$$

$$= \frac{1}{-\frac{1}{\theta}+1} \frac{x^{-1/\theta+1}}{\theta} \Big|_1^\infty = \frac{1}{1-\theta}$$

Det er viktig å merke seg at siden $\theta < 1/2$ så er $\lim_{x\to\infty} x^{-1/\theta+1} = 0$, om ikke kunne integralet divergert.

### e)

For å finne variansen må vi først finne

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_1^{\infty} \frac{x^{-\frac{\theta+2}{\theta}+1}}{\theta} dx$$

$$= \int_1^{\infty} \frac{x^{\frac{\theta-1}{\theta}}}{\theta} dx = \frac{1}{2\theta-1} x^{\frac{2\theta-1}{\theta}} \Big|_1^{\infty} = \frac{1}{2\theta-1}$$

Akkurat som i deloppgaven over, vil faktumet at $\theta < 1/2$ forsikre oss om at integralet divergerer, siden $\frac{2\theta-1}{\theta} < 0$, som gjør at $\lim_{x\to\infty} x^{\frac{2\theta-1}{\theta}} = 0$.

Vi kan nå finne variansen

$$V[X] = E[X^2] - E[X]^2 = \frac{1}{2\theta-1} - \frac{1}{(1-\theta)^2}$$

## Oppgave 3)

### a)

For å finne den kumulative fordelingsfunksjonen til $X$ legger vi merke til at

$$F(x) = P(X \le x) = 1 - P(X > x) \tag{1}$$

$q_x$ er sannsynligheten for at en x år gammel person skal dø før personen fyller $x+1$ år. Det betyr at sjansen for at personen overlever dette året er

$$1 - q_x$$

$P(X > x)$ er sannsynligheten for at personen blir eldre enn $x$, dette betyr at personen på ha overlevd alle årene opp til og med år $x$. Dette er produktet av sannsynligheten for å overleve de enkelte årene. Vi er bare interessert i sjansen overlever når de er over 35 år, så:

$$P(X > x) = \prod_{y=0}^{x} (1 - q_{35+y})$$

Vi kan så sette dette inn i (1) og ser at:

$$F(x) = 1 - P(X > x) = 1 - \prod_{y=0}^{x} (1 - q_{35+y}) \tag{2}$$

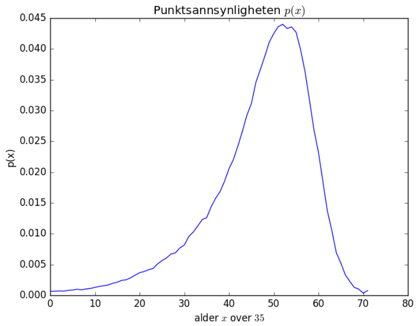
### b)

Punktsannsynligheten er sannsynligheten for at personen ble maks $x$ år gammel, men eldre enn $x-1$ år gammel, som gir oss:

$$p(x) = P(X = x) = P(x - 1 < X \le x) = F(x) - F(x-1)$$

Siden vi sier at ingen kan bli eldre enn 106 år gammel, så kan $x$ bare være definert i intervallet $0 \le x \le 71$.

## c)




Figur 3: Punktsannsynligheten $p(x)$ for å finne en person med alderen $x + 35$ år.

Vi kan se at denne punktsannsynligheten er tilnærmet en Poisson-fordeling.

## d)

$h(x)$ er nåverdien til den samlede pensjonsutbetalingen. Dør mannen før han fyller 35 år vil han ikke få noe utbetalt. Får han ikke noe utbetalt, vil nåverdien til dette beløpet også være 0. Det betyr at for $X \leq 34$ $h(X) = 0$.

Nåverdien er hvor mye penger som må settes inn i banken for at det etter $k$ år er blitt beløpet $K$. I vårt tilfelle mannen få $K = 100000$ i året. Nå verdien for disse pengene er gitt som

$$\frac{100000}{1.03^k}$$

med 3% rente. Han får dette beløpet verdt år etter han har fylt 70 år. For det første beløpet han får utbetalt har pengene stått 35 år i banken, og for det andre har de stått der 36 år. Den totale mengden penger han får utbetalt er $\sum_{x=35}^{x} 100000$ kr. (100000 for hvert år etter 70 år). Vi må finne nåverdien for hvert av disse årene. Dette betyr at vi får nåverdien:

$$h(X) = \sum_{x=35}^{x} \frac{100000}{1.03^k} \tag{3}$$

Vi kjenner igjen at dette er en geometisk rekke. Formelen for en slik rekke sier at

$$\sum_{a}^{b} r^k = \frac{r^a - r^{b+1}}{1 - r}$$

Bruker vi dette med (3) finner vi

$$h(X) = \sum_{x=35}^{x} \frac{10000}{1.03^k} = 100000 \frac{\frac{1}{1.03^{35}} - \frac{1}{1.03^{x+1}}}{1 - \frac{1}{1.03}}$$
$$= \frac{100000}{1.03^{35}} \frac{1 - (1/1.03)^{x+1-35}}{1 - (1/1.03)}$$

Dette gir oss endelig at nåverdien er

$$h(X) = \frac{100000}{1.03^{35}} \frac{1 - (1/1.03)^{x-34}}{1 - (1/1.03)} \tag{4}$$

For $X \geq 35$.

## e)

Forventingsverdien til nåverdien $h(X)$ er gitt ved

$$E[h(X)] = \sum_{x=0}^{71} h(X)p(x)$$

Men siden $h(X) = 0$ for $X < 35$ forenkles summen til

$$E[h(X)] = \sum_{x=35}^{71} h(x)p(x)$$

Vi kan så regne ut dette

$$\begin{aligned}\sum_{x=35}^{71} h(X)p(x) &= \sum_{x=35}^{71} \frac{100000}{1.03^{35}} \frac{1-(1/1.03)^{x-34}}{1-(1/1.03)} p(x) \\ &= \frac{100000}{1.03^{35}} \frac{\sum_{x=35}^{71}\left(1-(1/1.03)^{x-34}\right)}{1-(1/1.03)} p(x)\end{aligned}$$

Vi kan se at $\sum_{x=35}^{71} p(x) = P(X \geq 35)$, vi får derfor

$$E[h(X)] = \frac{100000}{1.03^{35}} \cdot \frac{P(X \geq 35) - \sum_{x=35}^{71}(1/1.03)^{x-34}p(x)}{1-(1/1.03)} \tag{5}$$

## f)

Løser jeg dette med koden du finner nedest i dokumentet, finner jeg at

$$E[h(X)] = \underline{\underline{387142}} \text{ kr}$$

## g)

Som i deloppgave *d)* er nåverdien gitt som

$$\frac{K}{1.03^k}$$

Der $K$ er premien, 1.03 er renten og $k$ er antall år.

Den samlede nåverdien er summen mannen betaler til han er 69 år eller dør. Så om $X < 35$, betaler han til han dør i en alder av $k = x$. Men om han blir 70 år eldre ($X \geq 35$) betaler han bare til han er 69 år – ved 70 år slutter han å betale, og får heller utbetalt fra forsikringen–. Dette betyr at han betaler til $k = \min(X, 34)$. Om nåverdien er gitt som $K \cdot g(X)$, så finner vi at

$$g(X) = \sum_{k=0}^{\min(X,34)} \frac{1}{1.03^k} \tag{6}$$

## h)

Vi gjenkjenner $g(X)$ som enda en geometrisk rekke, så vi skriver derfor den om som

$$g(X) = \frac{1-(1/1.03)^{\min(X,34)+1}}{1-(1/1.03)}$$

Vi kan så regne ut forventningsverdien til $g(X)$

$$E[g(X)] = \sum_{x=0}^{71} g(X)p(x) = \sum_{x=0}^{71} \frac{1-(1/1.03)^{\min(X,34)+1}}{1-(1/1.03)} p(x)$$

Vi kan splitte summen $\sum_{x=0}^{71}(1/1.03)^{\min(X,34)}$ inn i 2 deler: Den som gjelder for $X < 35$ og den delen hvor mannen har betalt i 34 år, og ikke lenger betaler premie($X \geq 35$):

$$E[g(X)] = \frac{1}{1-(1/1.03)} \left[ \underbrace{\sum_{x=0}^{71} p(x)}_{=1} - \left( \underbrace{\sum_{x=0}^{34} (1/1.03)^{x+1} p(x)}_{\text{For } X<35} + \underbrace{\sum_{x=35}^{71} (1/1.03)^{34+1} p(x)}_{\text{For } \geq 35} \right) \right]$$

Vi ser at det siste leddet blir:

$$\sum_{x=35}^{71} (1/1.03)^{34+1} p(x) = (1/1.03)^{35} \sum_{x=35}^{71} p(x) = (1/1.03)^{35} P(X \geq 35)$$

Dette gir oss

$$E[g(X)] = \frac{1-\sum_{x=0}^{34}(1/1.03)^{x+1}p(x) - (1/1.03)^{35}P(X \geq 35)}{1-(1/1.03)} \tag{7}$$

## i)

Bruker jeg koden gitt nederst i dokumentet, så finner jeg at

$$E[g(X)] = \underline{\underline{21.5649}}$$

## j)

Bruker vi at

$$K \cdot E[g(X)] = E[h(X)]$$

Så kan vi finne at premien mannen må betale i året er

$$K = \frac{E[h(X)]}{E[g(X)]} = \frac{387149}{21.5649} = \underline{\underline{17952.4}}$$

Det betyr at vi har funnet at mannen må betale en premie på $K = 17952.4$ kr per år for forsikringen.

## Appendix: Koder:

```
import numpy as np
import matplotlib.pyplot as plt

dodlighet = np.loadtxt("dodlighet.txt",skiprows = 1)
qk = dodlighet[:,1]/1000.


Fx=1-np.cumprod(1-qk[35:])

print len(Fx)



px=Fx-np.append(np.array([0]),Fx[:-1])
x = np.linspace(35,71,37)
print x




E = 100000./(1.03**35)*1.0/(1-1/1.03)*np.sum((1-(1.0/1.03)**(x-34))*px[35:])

x = np.linspace(0,71,72)
print x[35:]

Eg = 1.0/(1-1/1.03)*(1-np.sum((1/1.03)**(x[:35]+1)*px[:35]) - np.sum
    ((1/1.03)**35 * px[35:]))

print "Gjennomsnitt utbetalt: %g" %E
print "E[g(x)] = %g" %Eg

print "Premie som maa betales: %g" %(E/Eg)

plt.plot(px)
plt.title("Punktsannsynligheten $p(x)$")
plt.xlabel("alder $x$ over $35$")
plt.ylabel("p(x)")
plt.show()
```