



Norges teknisk-naturvitenskapelige universitet  
Institutt for matematiske fag

TMA4245 Statistikk

Vår 2015

Øving nummer 6, blokk I  
Løsningsskisse

### Oppgave 1

Vi antar  $X$  er normalfordelt,  $X \sim N(3315, 575^2)$ . Ved bruk av tabell A.3 finner vi

a) 1)

$$\begin{aligned}P(X > 3000) &= 1 - P(X < 3000) \\&= 1 - P\left(Z < \frac{3000 - 3315}{575}\right) \\&= 1 - P(Z < -0.55) \\&= 1 - 0.2912 \\&= 0.709.\end{aligned}$$

2)

$$\begin{aligned}P(3000 < X < 3500) &= P\left(\frac{3000 - 3315}{575} < Z < \frac{3500 - 3315}{575}\right) \\&= P(-0.55 < Z < 0.32) \\&= P(Z < 0.32) - P(Z < -0.55) \\&= 0.6255 - 0.2912 \\&= 0.335.\end{aligned}$$

3)

$$\begin{aligned}P(X > 3500 | X > 3000) &= \frac{P(X > 3500 \cap X > 3000)}{P(X > 3000)} \\&= \frac{P(X > 3500)}{P(X > 3000)} \\&= \frac{P(Z > 0.32)}{P(Z > -0.55)} \\&= \frac{1 - P(Z < 0.32)}{1 - P(Z < -0.55)} \\&= \frac{1 - 0.6255}{1 - 0.2912} \\&= 0.528.\end{aligned}$$

b)

$$P(Y > 0) = 1 - P(Y = 0) = 1 - 0.99^{100} = 0.6340$$

$$\begin{aligned} P(Y > 1 | Y > 0) &= \frac{P(Y > 1, Y > 0)}{P(Y > 0)} = \frac{1 - P(Y \leq 1)}{1 - P(Y = 0)} \\ &= \frac{1 - P(Y = 0) - P(Y = 1)}{1 - P(Y = 0)} = 0.4168 \end{aligned}$$

Det gamle løsningsforslaget (Overkill):

Dersom fødelvekten er mindre enn  $c$  gram,  $P(X < c) = 0.01$ , vil barnet bli klassifisert som undervektig. La  $n = 100$  være antall barn og  $Y$  antall av disse barna som var undervektige ved fødselen. Vi antar at fødelvekten for de ulike barna er uavhengige av hverandre. Fødelvekten til hvert barn kan da beskrives av en bernoulli variabel med suksesssannsynlighet (undervektig)  $p = P(X < c) = 0.01$ . Da vil  $Y$  være binomisk fordelt med parametrene  $n = 100, p = 0.01$ ,  $Y \sim b(y; 100, 0.01)$ .

Når  $Y$  er en binomisk fordelt variabel med fordeling  $b(y; n, p)$ ,  $n \rightarrow \infty$  og  $p \rightarrow 0$ ,  $np \rightarrow \mu$  vil vi kunne bruke Poisson-tilnærming til den binomiske fordelingen, dvs  $b(y; n, p) \rightarrow p(y; \mu)$ .

Vi bruker poisson-tilnærmingen  $Y \sim p(y; 1)$  og finner da ved bruk av tabell A.2

$$\begin{aligned} P(Y > 0) &= 1 - P(Y < 0) \\ &= 1 - 0.3679 \\ &= 0.63. \end{aligned}$$

$$\begin{aligned} P(Y > 1 | Y > 0) &= \frac{P(Y > 1 \cap Y > 0)}{P(Y > 0)} \\ &= \frac{P(Y > 1)}{P(Y > 0)} \\ &= \frac{1 - P(Y < 1)}{1 - P(Y < 0)} \\ &= \frac{1 - 0.7358}{1 - 0.3679} \\ &= 0.42. \end{aligned}$$

## Oppgave 2

$X$  og  $Y$  er uavhengige Poisson-fordelte stokastiske variable,  $X \sim p(x; 5)$  og  $Y \sim p(y; 10)$ . Fra tabell A.2 finner vi

$$P(X \leq 5) = \sum_{x=0}^5 p(x; 5) = 0.616$$

og

$$P(X \leq 3 | X \leq 5) = \frac{P(X \leq 3 \cap X \leq 5)}{P(X \leq 5)} = \frac{P(X \leq 3)}{P(X \leq 5)} = 0.430.$$

La  $Z = X + Y$ .  $X$  og  $Y$  er Poisson-fordelt og dermed vil variabelen  $Z$  være Poisson-fordelt med parameter  $\lambda = 5 + 10 = 15$ , dvs.  $Z \sim \text{Poisson}(15)$ . Ved bruk av tabell A.2 finner vi

$$P(X + Y > 10) = P(Z > 10) = 1 - P(Z \leq 10) = 1 - \sum_{x=0}^{10} = 1 - 0.1185 = 0.882.$$

### Oppgave 3

Definer

$X$  : antall tankskip som ankommer havnen i løpet av en dag

der

$$X \sim \text{poisson}(\lambda),$$

med

$$\lambda = E(X) = 2.$$

Havnen kan maksimalt betjene 3 tankskip per dag.

a) Da  $X$  er poissonfordelt har vi at

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} = \frac{2^x}{x!} e^{-2}, \quad x = 0, 1, 2, \dots$$

Med innsatte verdier for  $x$  har vi

$x$	0	1	2	3	4	5
$P(X = x)$	0.1353	0.2707	0.2707	0.1804	0.0902	0.0361

Vi ser dermed at det er størst sannsynlighet for at det ankommer ett eller to tankskip en bestemt dag. Tankskip må dirigeres til andre havner dersom det ankommer mer enn tre tankskip en dag. Vi har dermed

$$\begin{aligned} P(\text{Ett eller flere tankskip må omdirigeres}) &= P(X > 3) \\ &= 1 - P(X \leq 3) \stackrel{\text{tabell}}{=} 1 - 0.8571 = 0.1429 \end{aligned}$$

b) Definer

$Y$  : antall skip som betjenes en dag.

Da  $Y \leq 3$  har vi at

$$\begin{aligned} P(Y = y) &= P(X = x), \quad y = 0, 1, 2 \\ P(Y = 3) &= P(X \geq 3) = 1 - P(X \leq 2). \end{aligned}$$

Vi får dermed

$$\begin{aligned} E[Y] &= \sum_{y=0}^3 y P(Y = y) \\ &= 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot (1 - P(X \leq 2)) \\ &= 0.2707 + 2 \cdot 0.2707 + 3 \cdot (1 - 0.6767) = 1.782. \end{aligned}$$

c) Definer

$k$  : kapasitet.

Vi ønsker å finne  $k$  slik at  $P(X \leq k) \geq 0.90$ . Med innsatte verdier for  $k$  får vi

$k$	0	1	2	3	4	5
$P(X \leq k)$	0.1353	0.4060	0.6767	0.8571	0.9474	0.9834

Vi har dermed at

$$P(X \leq 4) = 0.9474 > 0.90,$$

og vi må dermed bygge ut havnen til kapasitet på fire skip for med minst 90% sannsynlighet å kunne betjene samtlige skip som ankommer en gitt dag.

#### Oppgave 4

En lottorekke kan oppfattes som et ikke-ordnet utvalg på 7 elementer blant tallene 1 til 34, der utvelgingen skjer uten tilbakelegging.

a) Det finnes

$$\binom{34}{7} = 5379616$$

forskjellige lotto-rekker.

Skal finne sannsynligheten for at lottorekka inneholder tallet 34. Denne kan beregnes som

$$P(\text{lottorekka inneholder tallet 34}) = 1 - P(\text{lottorekka inneholder IKKE tallet 34})$$

der antallet mulige lottorekker er  $\binom{34}{7}$  og antallet mulige lottorekker uten tallet 34 er  $\binom{33}{7}$  slik at

$$P(\text{lottorekka inneholder tallet 34}) = 1 - \frac{\binom{33}{7}}{\binom{34}{7}} = 1 - \frac{27}{34} = \frac{7}{34}.$$

Definer

$X$  : antall rette i lottorekken,

der  $X$  følger hypergeometrisk fordeling

$$h(x; N, n, k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$$

med  $N = 34$ ,  $n = 7$  og  $k = 7$ .

Sannsynligheten for at en lottorekke vil oppnå 7 rette er gitt ved

$$p = P(7 \text{ rette}) = h(7; 34, 7, 7) = \frac{\binom{7}{7} \binom{27}{0}}{\binom{34}{7}} = \frac{1}{\binom{34}{7}} = 1.859 \cdot 10^{-7}.$$

b) Det leveres hver uke inn  $n = 19\,000\,000$  rekker i lotto. Definer

$X$  : antall av de  $n$  innleverte rekkene som oppnår 7 rette i ukens trekning

$X$  er binomisk fordelt, men kan med god tilnærming regnes å være poissonfordelt da poissonfordelingen er grensefordelingen til binomisk fordeling når  $n \rightarrow \infty$  og  $p \rightarrow 0$ . Ifølge teorem 5.6 har vi at  $np \rightarrow \mu$  for  $n \rightarrow \infty$  slik at parameteren i poissonfordelingen er gitt ved

$$\mu = np = 19\,000\,000 \times 1.859 \cdot 10^{-7} = 3.53.$$

Sannsynlighetstettheten til  $X$  er nå gitt ved

$$f_X(x) = \frac{\mu^x}{x!} e^{-\mu}$$

og vi har

$$\begin{aligned} &P(\text{ingen av de innleverte rekkene har 7 rette}) \\ &= P(X = 0) = \frac{3.53^0}{0!} e^{-3.53} = e^{-3.53} = 0.029. \end{aligned}$$

Legg merke til at vi oppnår samme svar uten poissontilnærmingen, dvs dersom  $X$  er binomisk fordelt med parametre  $n = 19\,000\,000$  og  $p = 1.859 \cdot 10^{-7}$  gitt ved

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Vi har da

$$\begin{aligned} &P(\text{ingen av de innleverte rekkene har 7 rette}) \\ &= P(X = 0) = \binom{n}{0} p^0 (1-p)^{n-0} = (1-p)^n = (1 - 1.859 \cdot 10^{-7})^{19\,000\,000} = 0.029. \end{aligned}$$

Definer

$Y$  : antall Gull-lotto omganger pr. år

der vi antar at det er 52 uker pr. år.  $Y$  følger her en binomisk fordeling med  $n = 52$  og  $p = 0.029$ . Vi har dermed at

$$E(Y) = n p = 52 \times 0.029 = 1.5$$

og

$$P(Y = 0) = f_Y(0) = \binom{52}{0} p^0 (1-p)^{52-0} = (1 - 0.029)^{52} = 0.22.$$

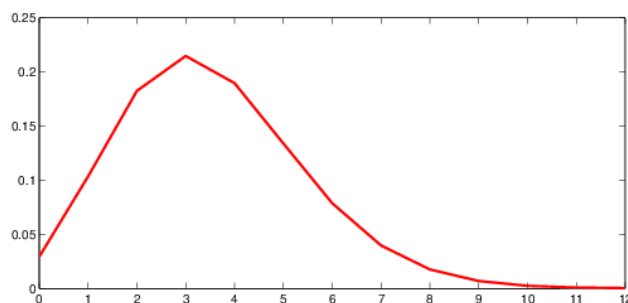
c) Eksempel på en mulig kode for Poisson sannsynlighetstetthetsfunksjon:

```
function funcval = poisson(x,mean)
funcval = ((mean^x)/(factorial(x)))*(exp(-mean));
```

Et plot av verdiene fra Poisson tetthetsfunksjonen for  $x$ -verdier fra 0 til 12 er vist i Fig 1. Her har vi brukt  $\mu = np = 3.53$  som vi fant i Oppgave 5b). Vi bruker heltall ettersom Poisson tetthetsfunksjonen er en diskret sannsynlighetsfordeling som bare er definert for heltallsverdier av  $x$ .

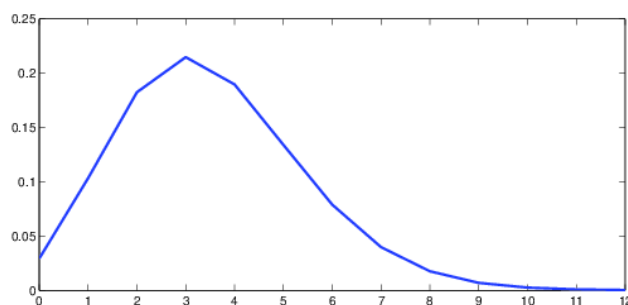
d) Eksempel på en mulig kode for den binomiske sannsynlighetstetthetsfunksjonen:

```
function funcval = binomial(n,x,p)
funcval = nchoosek(n,x) * (p^x) * ((1-p)^(n-x));
```



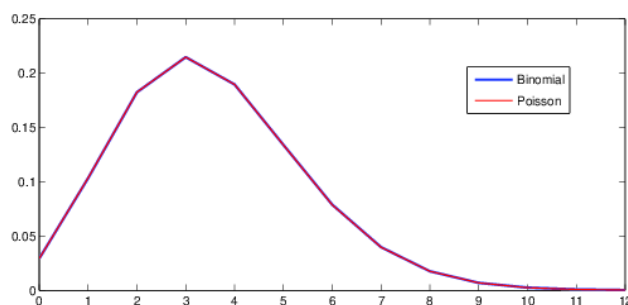
Figur 1: Poissontetthetsfunksjonen,  $f_X(x) = \frac{\mu^x}{x!} e^{-\mu}$  for heltallsverdier av  $x$  fra 0 til 12 med  $\mu = 3.53$ .

Et plot av verdiene fra den binomiske tetthetsfunksjonen for  $x$ -verdier fra 0 til 12 er vist i Fig 2. Her har vi brukt  $n = 19\,000\,000$  og  $p = 1.859 \cdot 10^{-7}$  som vi fant i Oppgave 5b).



Figur 2: Binomisk tetthetsfunksjon,  $f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$  for heltallsverdier av  $x$  fra 0 til 12 med  $n = 19\,000\,000$  og  $p = 1.859 \cdot 10^{-7}$ .

I Fig 3 har vi plottet de to tetthetsfunksjonene, Poisson og binomisk, mot hverandre. Vi ser at tetthetsfunksjonene er tilnærmet identiske ettersom de overlapper hverandre nesten perfekt.



Figur 3: Poissontetthetsfunksjonen av b) plottet mot den binomiske tetthetsfunksjonen.

- e) Vi skal nå øke antall valgbare tall fra 34 til  $m$ , der  $m > 34$  slik at det med sannsynlighet minst 0.1 ikke finnes rekker med 7 riktige i en uke der det innleveres  $n = 19\,000\,000$  rekker. Vi må i denne oppgaven bruke de tidligere definerte ligningene. Vi har at

$$\begin{aligned} P(\text{ingen av de innleverte rekkene har 7 rette}) \\ = P(X = 0) = (1 - p)^{19\,000\,000} \geq 0.1 \end{aligned}$$

Vi har videre at  $p$  er sannsynligheten for at en lottorekke vil oppnå 7 rette, og denne er gitt ved

$$p = P(7 \text{ rette}) = h(7; m, 7, 7) = \frac{\binom{7}{7} \binom{m-7}{0}}{\binom{m}{7}} = \frac{1}{\binom{m}{7}}$$

slik at

$$\left(1 - \frac{1}{\binom{m}{7}}\right)^{19\,000\,000} \geq 0.1$$

Prøver med innsatte verdier av  $m$ :

$m$	35	36
$\left(1 - \frac{1}{\binom{m}{7}}\right)^{19\,000\,000}$	0.059	0.103

Vi ser av tabellen at ulikheten er oppfylt for  $m = 36$ , og har dermed at  $m \geq 36$  dersom det med sannsynlighet minst 0.1 ikke skal finnes rekker med 7 riktige i en uke der det innleveres  $n = 19\,000\,000$  rekker.

## Oppgave 5

- a) Vi må i tillegg anta at hver gjennomlesning av teksten er uavhengig av hverandre.

Vi har  $\lambda = 2$  og  $s = 8$  og ønsker å finne sannsynligheten for at antall trykkfeil,  $N$ , er større enn 10. Vi har  $\mu = \lambda s = 2 \cdot 8 = 16$



$$\begin{aligned}P(N > 10) &= 1 - P(N \leq 10) \\&= 1 - \sum_{i=1}^{10} P(N = i) \\&= 1 - 0.0774 \\&= 0.923.\end{aligned}$$

Vi har nå gitt  $N = 12$  og  $p = 0.6$  og ønsker å finne sannsynligheten for at korrekturleseren oppdager alle trykkfeilene.

$$\begin{aligned}P(X = 12|N = 12) &= \binom{12}{12} \cdot 0.6^{12} \cdot 0.4^0 \\&= 0.6^{12} \\&= 0.0022.\end{aligned}$$

- b)  $Y_k$  = antall trykkfeil som gjenstår etter  $k$  uavhengige gjennomlesninger. Vi finner først simultanfordelingen til  $Y_1$  og  $N$ .

Vi har

$$P(X = x|N = n) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, \dots, n$$

Simultanfordelingen til  $Y_1$  og  $N$  er da gitt ved

$$\begin{aligned}P(Y_1 = u, N = n) &= P(Y_1 = u|N = n) \cdot P(N = n) \\&= P(N - X = u|N = n) \cdot P(N = n) \\&= P(X = n - u|N = n) \cdot P(N = n) \\&= \binom{n}{n-u} p^{n-u} (1-p)^u \cdot P(N = n)\end{aligned}$$

for  $u = 0, 1, \dots$  og  $n = u, u + 1, \dots$

Vi finner deretter marginalfordelingen til  $Y_1$ .

$$\begin{aligned}
 P(Y_1 = u) &= \sum_{n=u}^{\infty} P(Y_1 = u, N = n) \\
 &= \sum_{n=u}^{\infty} \binom{n}{n-u} p^{n-u} (1-p)^u e^{-\lambda s} \cdot \frac{(\lambda s)^n}{n!} \\
 &= \sum_{n=0}^{\infty} \binom{n+u}{n} p^n (1-p)^u e^{-\lambda s} \cdot \frac{(\lambda s)^{n+u}}{(n+u)!} \\
 &= \sum_{n=0}^{\infty} \frac{1}{n! u!} p^n (1-p)^u e^{-\lambda s} (\lambda s)^{n+u} \\
 &= \frac{(\lambda s)^u e^{-\lambda s}}{u!} (1-p)^u \sum_{n=0}^{\infty} \frac{(\lambda p s)^n}{n!} \\
 &= \frac{(\lambda s)^u e^{-\lambda s}}{u!} (1-p)^u e^{\lambda p s} \\
 &= \frac{(\lambda s(1-p))^u}{u!} e^{-\lambda s(1-p)}.
 \end{aligned}$$

Vi ser at marginalfordelingen til  $Y_1$  er  $Y_1 \sim \text{Poisson}(\lambda s(1-p))$ .

Siden fordelingen for  $Y_1$  er den samme som for  $N$  bortsett fra at parameteren er endret, vil mønsteret gjenta seg slik at også  $Y_2$  blir poissonfordelt, og parameteren i fordelinga til  $Y_2$  blir  $\lambda s(1-p)(1-p) = \lambda s(1-p)^2$ . Tilsvarende blir  $Y_3 \sim \text{Poisson}(\lambda s(1-p)^3)$  og generelt  $Y_k \sim \text{Poisson}(\lambda s(1-p)^k)$ .