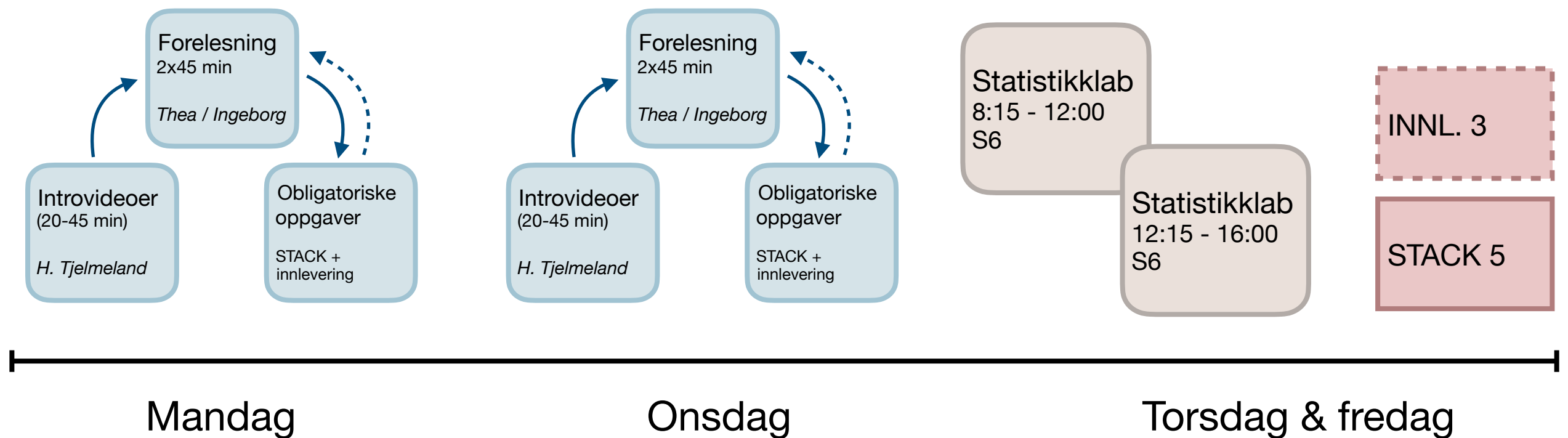


TMA4245 Statistikk

Uke 6 - mandag



Uke 6 (mandag 5. feb - fredag 9. feb)



Introvideoer

Geometrisk og negativ binomisk fordeling

TMA4240/TMA4245 Statistikk

Håkon Tjelmeland
Institutt for matematiske fag
Norges teknisk-naturvitenskapelige universitet

Poissonprosess og poissonfordeling

TMA4240/TMA4245 Statistikk

Håkon Tjelmeland
Institutt for matematiske fag
Norges teknisk-naturvitenskapelige universitet

Poissonfordeling som tilærming til binomisk fordeling

TMA4240/TMA4245 Statistikk

Håkon Tjelmeland
Institutt for matematiske fag
Norges teknisk-naturvitenskapelige universitet

Oppgaver

STACK 5: Oppg. 1-3
INNLL. 3: Oppg. 1(?)

Kjente / mye brukte sannsynlighetsfordelinger

Definisjon (Bernoulli forsøksrekke)

Anta at vi gjentar et stokastisk forsøk mange ganger. Vi sier da at vi har en Bernoulli forsøksrekke dersom

- ★ *forsøket har to mulige utfall, som vi kaller henholdsvis «suksess» og «fiasko»,*
- ★ *utfallet av forsøkene er uavhengige,*
- ★ *sannsynligheten for suksess er lik (p) i alle forsøkene.*

Binomisk(n, p) - antall suksesser i n forsøk

Geometrisk(p) - antall forsøk frem til og med første suksess

Negativ binomisk(k, p) - antall forsøk frem til og med suksess nr. k

Oppgave 1a

En gruppe ingeniørstudenter forsøker å lage en tavle-vask-robot.

Roboten er ikke ferdig, men de vil vise frem roboten til faglærer.

De vet fra testing at sannsynligheten for at roboten klarer tavle-vasken er ca 0.4.

- i) Hva er sannsynligheten for at roboten klarer tavle-vask på første forsøk?
- ii) Hva er sannsynligheten for at roboten ikke klarer tavle-vask før på tredje forsøk?
- iii) Hva er sannsynligheten for at roboten må prøve mer enn tre ganger før den klarer tavle-vasken?

Anta at forsøkene på tavle-vask er uavhengige.

Geometrisk fordeling

$$f(x) = g(x; p) = p(1 - p)^{x-1}, \quad x = 1, 2, \dots$$

$$E(X) = \frac{1}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2}, \quad M_X(t) = \frac{pe^t}{1 - (1-p)e^t} \text{ for } t < -\ln(1-p).$$

Oppgave 1b

En gruppe ingeniørstudenter forsøker å lage en tavle-vask-robot.

Studentene skal delta i en konkurranse der vinnerlaget er de som klarer 3 tavle-vask på færrest mulig forsøk.

- i) Hva er forventet antall forsøk roboten må gjøre for å klare 3 tavle-vask når suksess-sannsynligheten er 0.4?
- ii) Dersom studentene ønsker at forventet antall forsøk roboten må gjøre for å klare 3 tavle-vask skal være høyst 4, hvilken sannsynlighet for å klare tavle-vask bør de jobbe mot før konkurransen?

Negativ-binomisk fordeling

$$f(x) = b^*(x; k, p) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad x = k, k+1, \dots$$

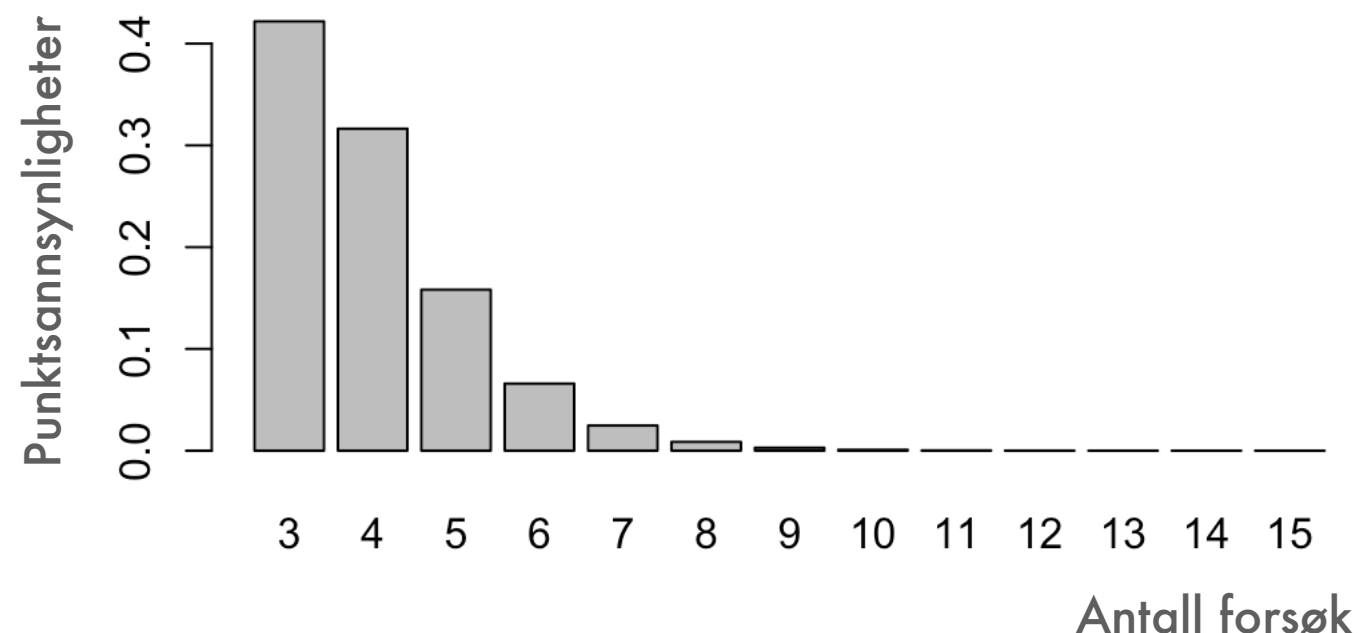
$$E(X) = \frac{k}{p}, \quad \text{Var}(X) = k \frac{1-p}{p^2}, \quad M_X(t) = \left(\frac{pe^t}{1-(1-p)e^t} \right)^k \text{ for } t < -\ln(1-p).$$

Oppgave 1b

En gruppe ingeniørstudenter forsøker å lage en tavle-vask-robot.

Studentene skal delta i en konkurranse der vinnerlaget er de som klarer 3 tavle-vask på færrest mulig forsøk.

- i) Hva er forventet antall forsøk roboten må gjøre for å klare 3 tavle-vask når suksess-sannsynligheten er 0.4?
- ii) Dersom studentene ønsker at forventet antall forsøk roboten må gjøre for å klare 3 tavle-vask skal være høyst 4, hvilken sannsynlighet for å klare tavle-vask bør de jobbe mot før konkurransen?



Kjente / mye brukte sannsynlighetsfordelinger

Definisjon (Bernoulli forsøksrekke)

Anta at vi gjentar et stokastisk forsøk mange ganger. Vi sier da at vi har en Bernoulli forsøksrekke dersom

- ★ forsøket har to mulige utfall, som vi kaller henholdsvis «suksess» og «fiasko»,
- ★ utfallet av forsøkene er uavhengige,
- ★ sannsynligheten for suksess er lik (p) i alle forsøkene.

Binomisk(n, p) - antall suksesser i n forsøk

Geometrisk(p) - antall forsøk frem til og med første suksess

Negativ binomisk(k, p) - antall forsøk frem til og med suksess nr. k

Definisjon (Poissonprosess)

En prosess $\{N(t); t \geq 0\}$ kalles in poissonprosess med intensitet λ hvis følgende krav er oppfylt

1. $N(0) = 0$.

2. For $0 \leq t_1 < t_2 \leq t_3 < t_4$ er $N(t_2) - N(t_1)$ og $N(t_4) - N(t_3)$ uavhengige.

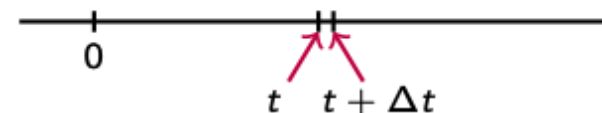
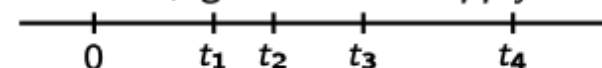
3. For $t \geq 0$ og $\Delta t > 0$ har vi at

$$P(N(t + \Delta t) - N(t) = 1) = \lambda \Delta t + o(\Delta t),$$

$$P(N(t + \Delta t) - N(t) \geq 2) = o(\Delta t),$$

der hver $o(\Delta t)$ angir en funksjon som oppfyller

$$\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0.$$



Er dette Poissonprosesser?

Eks 1 (tid):

Studenter som ankommer Rapido (kiosk Stripa) per minutt

Eks 2 (tid):

Hjerteinfarktpasienter per dag på St Olavs Hospital

Eks 3 (a) (rom):

Bilulykker per km² i Norge

Eks 3 (b) (rom):

Bilulykker per km av E6

Kjente / mye brukte sannsynlighetsfordelinger

Definisjon (Bernoulli forsøksrekke)

Anta at vi gjentar et stokastisk forsøk mange ganger. Vi sier da at vi har en Bernoulli forsøksrekke dersom

- ★ forsøket har to mulige utfall, som vi kaller henholdsvis «suksess» og «fiasko»,
- ★ utfallet av forsøkene er uavhengige,
- ★ sannsynligheten for suksess er lik (p) i alle forsøkene.

Binomisk(n, p) - antall suksesser i n forsøk

Geometrisk(p) - antall forsøk frem til og med første suksess

Negativ binomisk(k, p) - antall forsøk frem til og med suksess nr. k

Definisjon (Poissonprosess)

En prosess $\{N(t); t \geq 0\}$ kalles en poissonprosess med intensitet λ hvis følgende krav er oppfylt

1. $N(0) = 0$.

2. For $0 \leq t_1 < t_2 \leq t_3 < t_4$ er $N(t_2) - N(t_1)$ og $N(t_4) - N(t_3)$ uavhengige.

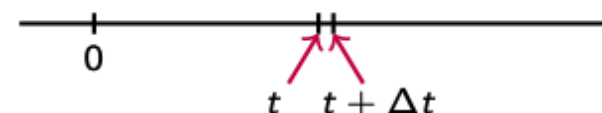
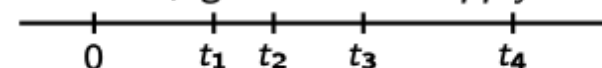
3. For $t \geq 0$ og $\Delta t > 0$ har vi at

$$P(N(t + \Delta t) - N(t) = 1) = \lambda \Delta t + o(\Delta t),$$

$$P(N(t + \Delta t) - N(t) \geq 2) = o(\Delta t),$$

der hver $o(\Delta t)$ angir en funksjon som oppfyller

$$\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0.$$



Poisson(λt) - antall hendelser i intervall/område av størrelse t

Oppgave 2

Busselskapet AtilÅ sitt kundesenter er åpent hver dag fra klokka 07:00 til 19:00. Kunder som vil ha hjelp med *billettkjøp* ankommer kundesenteret ifølge en Poissonprosess med rate 5 per time. La X være antallet slike kunder som ankommer i løpet av en time slik at

$$P(X = x) = \frac{5^x e^{-5}}{x!}, \quad x = 0, 1, 2, \dots$$

- a) Hva er sannsynligheten for at det ikke kommer en eneste kunde som trenger hjelp med billettkjøp mellom klokken 07:00 og 08:00?

Hva er sannsynligheten for at det kommer minst 20 kunder som trenger hjelp med billettkjøp før klokka 10:00?

Definisjon (Poissonfordeling)

La $N(t)$ være en poissonprosess med intensitet λ . Fikser en verdi $t > 0$ og la $X = N(t)$. Vi sier da at X er poissonfordelt med parameter λt .

Teorem (Punktsannsynlighet i poissonfordeling)

La X være poissonfordelt med parameter λt . Da er punktsannsynligheten til X gitt ved

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

Teorem (Forventningsverdi og varians i poissonfordeling)

La X være poissonfordelt med parameter λt . Da er

$$E[X] = \lambda t \quad \text{og} \quad \text{Var}[X] = \lambda t.$$

Poissonfordeling

$$P(X \leq x)$$

$x \backslash \mu$	10	11	12	13	14	15	16	17	18
0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0005	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
2	.0028	.0012	.0005	.0002	.0001	.0000	.0000	.0000	.0000
3	.0103	.0049	.0023	.0011	.0005	.0002	.0001	.0000	.0000
4	.0293	.0151	.0076	.0037	.0018	.0009	.0004	.0002	.0001
5	.0671	.0375	.0203	.0107	.0055	.0028	.0014	.0007	.0003
6	.1301	.0786	.0458	.0259	.0142	.0076	.0040	.0021	.0010
7	.2202	.1432	.0895	.0540	.0316	.0180	.0100	.0054	.0029
8	.3328	.2320	.1550	.0998	.0621	.0374	.0220	.0126	.0071
9	.4579	.3405	.2424	.1658	.1094	.0699	.0433	.0261	.0154
10	.5830	.4599	.3472	.2517	.1757	.1185	.0774	.0491	.0304
11	.6968	.5793	.4616	.3532	.2600	.1848	.1270	.0847	.0549
12	.7916	.6887	.5760	.4631	.3585	.2676	.1931	.1350	.0917
13	.8645	.7813	.6815	.5730	.4644	.3632	.2745	.2009	.1426
14	.9165	.8540	.7720	.6751	.5704	.4657	.3675	.2808	.2081
15	.9513	.9074	.8444	.7636	.6694	.5681	.4667	.3715	.2867
16	.9730	.9441	.8987	.8355	.7559	.6641	.5660	.4677	.3751
17	.9857	.9678	.9370	.8905	.8272	.7489	.6593	.5640	.4686
18	.9928	.9823	.9626	.9302	.8826	.8195	.7423	.6550	.5622
19	.9965	.9907	.9787	.9573	.9235	.8752	.8122	.7363	.6509
20	.9984	.9953	.9884	.9750	.9521	.9170	.8682	.8055	.7307
21	.9993	.9977	.9939	.9859	.9712	.9469	.9108	.8615	.7991
22	.9997	.9990	.9970	.9924	.9833	.9673	.9418	.9047	.8551
23	.9999	.9995	.9985	.9960	.9907	.9805	.9633	.9367	.8989
24	1.0000	.9998	.9993	.9980	.9950	.9888	.9777	.9594	.9317
25	1.0000	.9999	.9997	.9990	.9974	.9938	.9869	.9748	.9554
26	1.0000	1.0000	.9999	.9995	.9987	.9967	.9925	.9848	.9718
27	1.0000	1.0000	.9999	.9998	.9994	.9983	.9959	.9912	.9827
28	1.0000	1.0000	1.0000	.9999	.9997	.9991	.9978	.9950	.9897
29	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9989	.9973	.9941
30	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9986	.9967
31	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9993	.9982
32	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9990
33	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9995
34	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998
35	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
36	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
37	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Oppgave 2

Busselskapet AtilÅ sitt kundesenter er åpent hver dag fra klokka 07:00 til 19:00. Kunder som vil ha hjelp med *billettkjøp* ankommer kundesenteret ifølge en Poissonprosess med rate 5 per time. La X være antallet slike kunder som ankommer i løpet av en time slik at

$$P(X = x) = \frac{5^x e^{-5}}{x!}, \quad x = 0, 1, 2, \dots$$

- a) Hva er sannsynligheten for at det ikke kommer en eneste kunde som trenger hjelp med billettkjøp mellom klokken 07:00 og 08:00?

Hva er sannsynligheten for at det kommer minst 20 kunder som trenger hjelp med billettkjøp før klokka 10:00?

Definisjon (Poissonfordeling)

La $N(t)$ være en poissonprosess med intensitet λ . Fikser en verdi $t > 0$ og la $X = N(t)$. Vi sier da at X er poissonfordelt med parameter λt .

Teorem (Punktsannsynlighet i poissonfordeling)

La X være poissonfordelt med parameter λt . Da er punktsannsynligheten til X gitt ved

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

Teorem (Forventningsverdi og varians i poissonfordeling)

La X være poissonfordelt med parameter λt . Da er

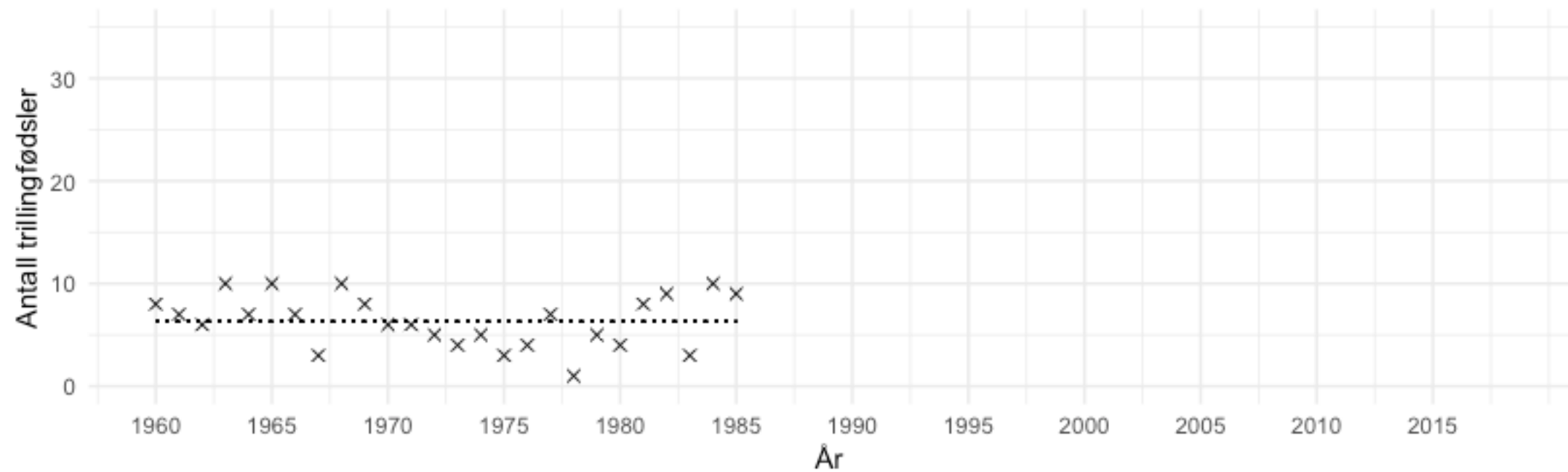
$$E[X] = \lambda t \quad \text{og} \quad \text{Var}[X] = \lambda t.$$

Poissonfordeling

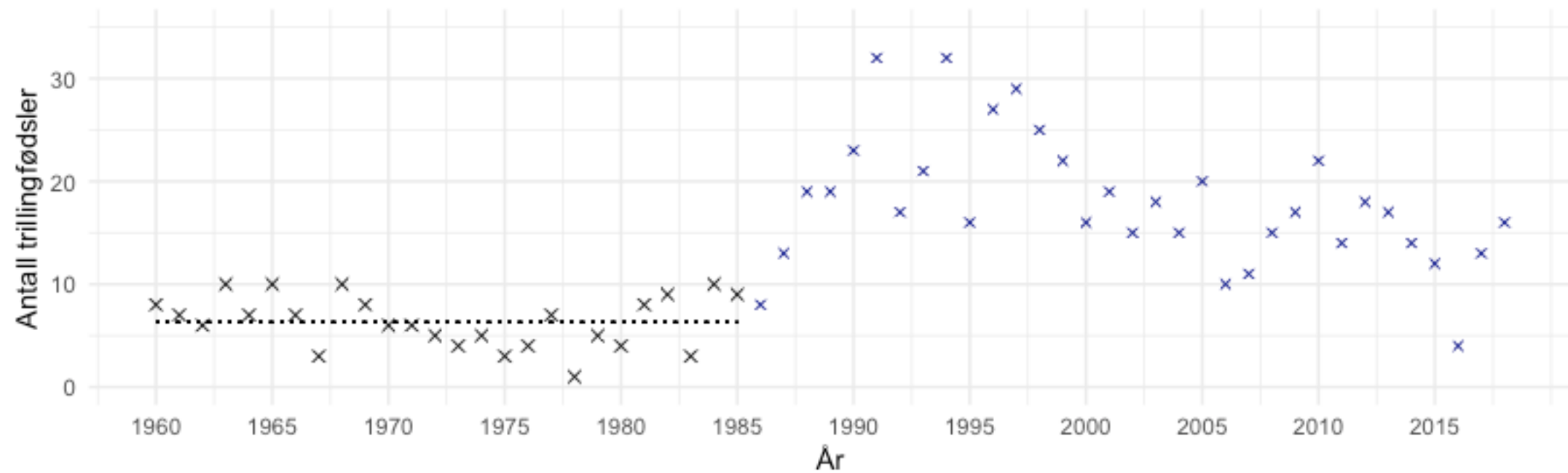
$$f(x) = p(x; \mu) = \frac{\mu^x}{x!} e^{-\mu}, \quad x = 0, 1, 2, \dots$$

$$E(X) = \mu, \quad \text{Var}(X) = \mu, \quad M_X(t) = e^{\mu(e^t - 1)}.$$

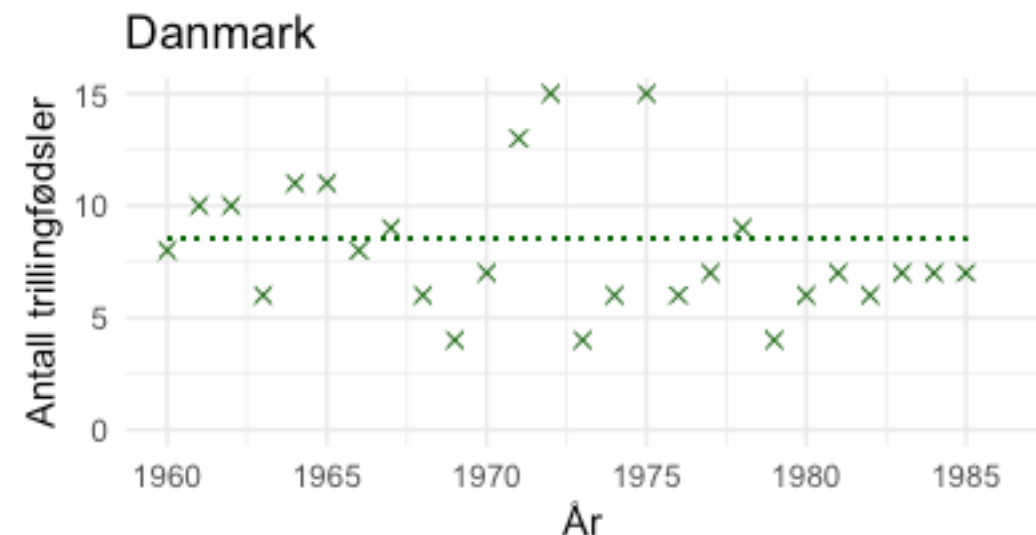
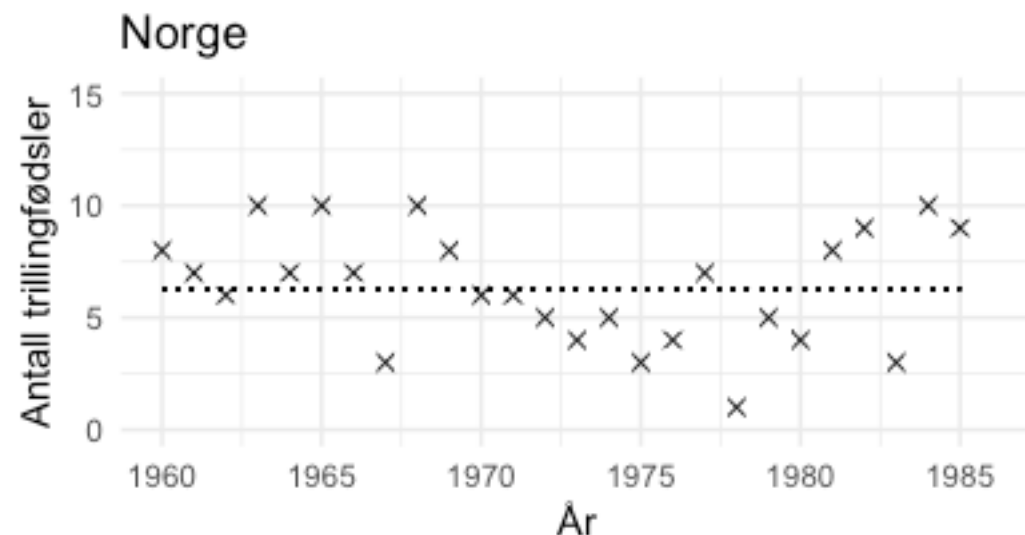
Trillingfødsler i Norge



Trillingfødsler i Norge



Trillingfødsler i Norge og Danmark



Hva om vi ser de to landene under ett?

X_A : trillingfødsler i løpet av t år i Norge, $X_A \sim \text{Poisson}(\lambda_A t)$

X_B : trillingfødsler i løpet av t år i Danmark, $X_B \sim \text{Poisson}(\lambda_B t)$

$$Y = X_A + X_B$$

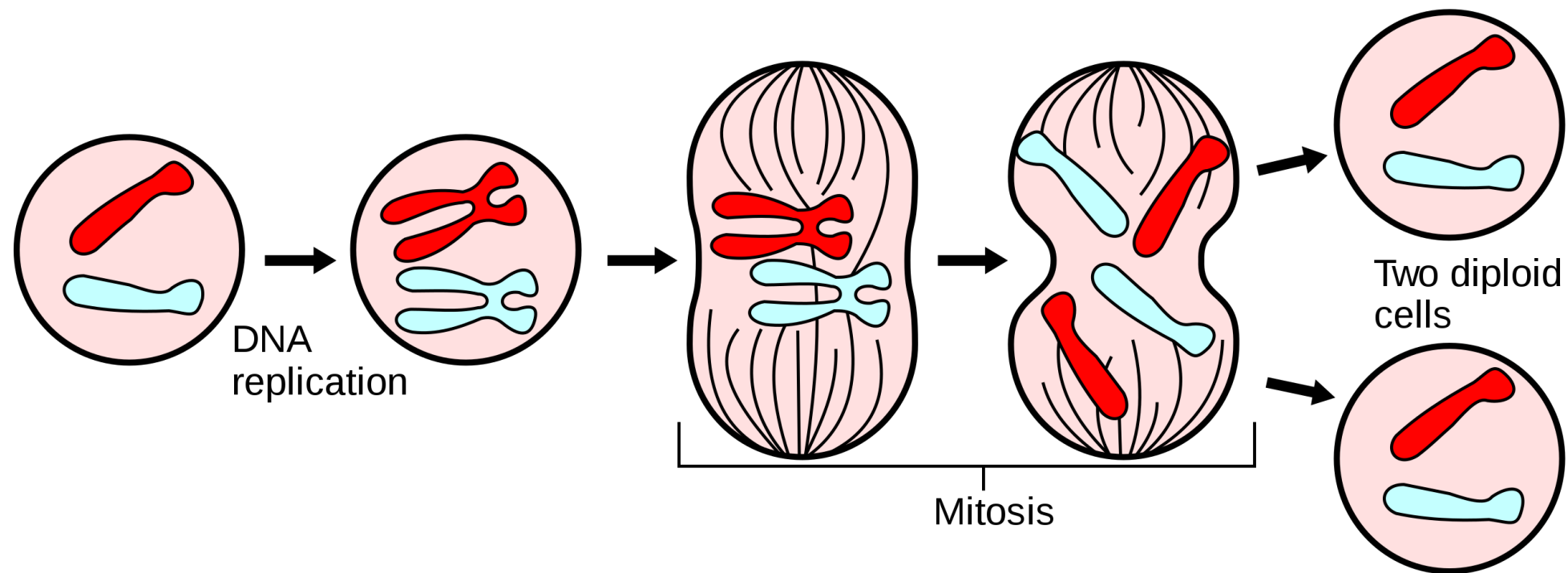
X_A og X_B uavhengige.

Hvilken fordeling har Y ?

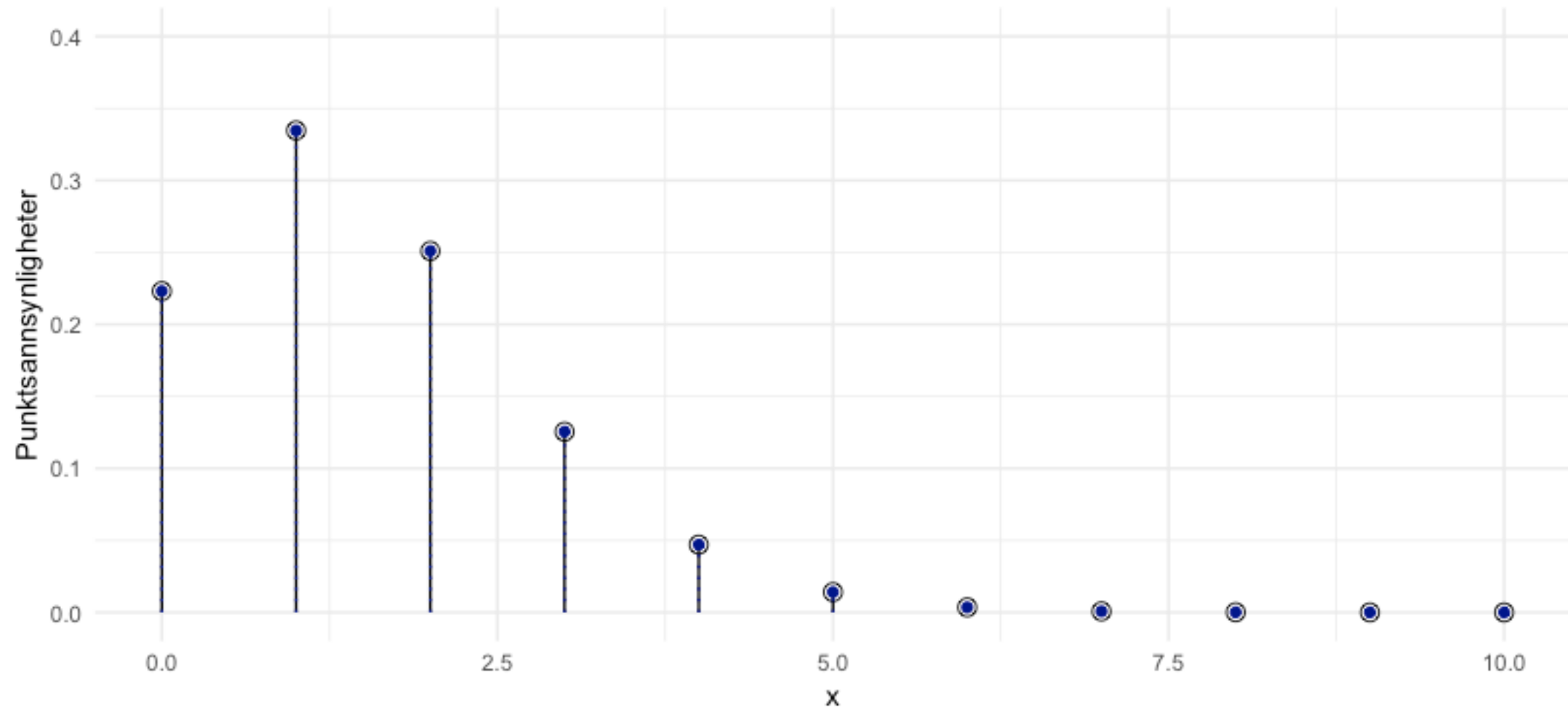
$$\sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = (a + b)^n$$

$$P(Y = y) = \frac{((\lambda_A + \lambda_B)t)^y}{y!} e^{-(\lambda_A + \lambda_B)t} \quad ?$$

Eks: Celledeling



Eks: Celledeling



Eks: Celledeling

Teorem (Poissonfordeling som tilnærming til binomisk fordeling)

Anta at X er binomisk fordelt med parametre n og p , slik at punktsannsynligheten til X er gitt ved

$$f(x; n, p) = \binom{n}{x} p^x (1 - p)^{n-x}, x = 0, 1, 2, \dots, n.$$

I grensen når $n \rightarrow \infty$ og $p \rightarrow 0$ på en slik måte at $\lambda = np$ holdes konstant har man da for hver $x = 0, 1, 2, \dots$ at

$$\lim_{\substack{n \rightarrow \infty, p \rightarrow 0 \\ np = \lambda}} f(x; n, p) = \frac{\lambda^x}{x!} e^{-\lambda}.$$

★ Bevis senere

— momentgenerende funksjoner

★ Dermed: Hvis n er stor og p liten, kan man med god approksimasjon regne på poissonfordeling i stedet for en binomisk fordeling

Oppgave 3

La X_1 , X_2 og X_3 være uavhengige poissonfordelte stokastiske variabler med forventningsverdi henholdsvis μ_1 , μ_2 og μ_3 .

La

$$Z_i = \begin{cases} 1 & \text{hvis } X_i = X_3 = 0 \\ 0 & \text{ellers.} \end{cases} \quad (5)$$

for $i = 1, 2$.

- a) Vis at Z_i er bernoullifordelt, det vil si binomisk fordelt med $n = 1$ delforsøk, med suksessannsynlighet $p_i = e^{-(\mu_i + \mu_3)}$, og bruk dette til å finne forventningsverdien og variansen til Z_i for $i = 1, 2$.

Bruk dette til å utlede et uttrykk for kovariansen mellom Z_1 og Z_2 , det vil si $\text{Cov}(Z_1, Z_2)$.

Binomisk fordeling

$$f(x) = b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n.$$

$$E(X) = np, \quad \text{Var}(X) = np(1-p), \quad M_X(t) = (pe^t + 1 - p)^n.$$