

OPPG 1a

Bernoulli forsøksrekke

1) suksess (tauke-vask) / fiasko

2) uavhengige forsøk

3) sanns. $p=0.4$ for suksess
i alle forsøk

X : antall forsøk frem til og med første suksess

$X \sim \text{geom}(p=0.4)$

$$\text{i) } P(X=1) = 0.4 (1-0.4)^{1-1} = \underline{\underline{0.4}}$$

$$\text{ii) } P(X=3) = 0.4 (1-0.4)^{3-1} = 0.4 \cdot 0.6^2 = \underline{\underline{0.144}}$$

$$\begin{aligned} \text{iii) } P(X>3) &= 1 - P(X \leq 3) = 1 - (P(X=1) + P(X=2) + P(X=3)) \\ &= 1 - (0.4 + 0.24 + 0.144) = \underline{\underline{0.216}} \end{aligned}$$

Oppg 1b

Y : antall forsøk frem til og med tredje suksess

$$Y \sim \text{negbinom}(p=0.4, k=3)$$

$$i) E(Y) = \frac{3}{0.4} = 7.5$$

$$ii) Z \sim \text{negbinom}(p, k=3)$$

$$E(Z) = \frac{3}{p} \leq 4 \quad \text{hva blir } p?$$

$$p \geq 3/4 = 0.75 \quad \rightarrow \text{figur for } \boxed{p=0.75}$$

Poissonprosess (uformelt)

tidsintervaller
↓
områder

- * Uavhengige hendelser i ikke-ovelappende segmenter
- * konstant rate λ

eks : λ = forventet antall hendelser / time

- * to hendelser kan ikke
skje på nøyaktig samme
tid / sted

/ km

/ m²

/ sekund etc.

OPPG. 2

rate $\lambda = 5$ kunder / time

$\bar{X}$ = antall kunder i 1 time,

$$E(\bar{X}) = 5 \cdot 1 = 5$$

$$a) P(\bar{X} = 0) = \frac{5^0}{0!} e^{-5} = e^{-5} = 0,0067$$

b) $\bar{X}_3$ = antall kunder i 3 timer (0700-1000)

$$\bar{X}_3 \sim \text{Poisson}(\lambda 3)$$

$$E(\bar{X}_3) = \lambda 3 = 15$$

$$\begin{aligned} P(\bar{X}_3 \geq 20) &= 1 - P(\bar{X}_3 < 20) \\ &= 1 - P(\bar{X}_3 \leq 19) \\ &= 1 - 0.8752 \leftarrow \text{tabell} \\ &= \underline{\underline{0.1248}} \end{aligned}$$

Trillingfødsler - en Poissonprosess?

- Uavhengighet mellom ikke-overlappende tidsrom
- Samme rate så lenge populasjonen av mødre holder seg lik
- Data behrefter antagelse

$\bar{X}$ = antall trillingfødsler på ett år (1960-1985)

$\bar{X} \sim \text{Poisson}(6.3)$  snitt og varians i dataene er ca 6.3 (mer om estimering fra data i Blokk 2)

→ hvor mange trillingfødsler fra 1986-2018? 33 år

$\bar{X}_{33}$ = antall trillingfødsler på 33 år

$\bar{X}_{33} \sim \text{Poisson}(6.3 \cdot 33)$ $E(\bar{X}_{33}) = 207.9$

- nærmere 600 trillingfødsler
 - snitt 18 per år, varians 40 (likke lenger en Poissonprosess)
- * Hva har skjedd? *

Danmark 1960-1985, burde også være en Poissonprosess

$\lambda = 6.3$ i Norge $\leftarrow$ populasjon ca 3.8 mill

DK ca 4.9 mill (samme tid) $DK/Norge \approx 1.3$

$\lambda_{DK} \approx 6.3 \cdot 1.3 = 8.19$ (obs. god gjettning)

Sum av uavh.-Poissonfordelte variable

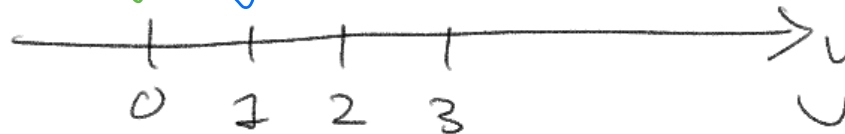
Stokastisk forsøk: observere to uavhengige Poisson-prosesser
i en tid t og telle hendelser

Prosess A har rate λ_A

Prosess B har rate λ_B

Utfallsrom: $\overset{A \downarrow}{\downarrow} \overset{B \downarrow}{\downarrow} \{ (0,0), (0,1), (1,0), (1,1), (2,0), (0,2), (2,1) \dots \}$

Y : en funksjon som tar inn
et enkeltutfall (f.eks $(0,1)$)
og returnerer totalt antall
hendelser (f.eks. $0+1=1$)



OSV.

$$P(Y=0) = P((0,0))$$

$$P(Y=1) = P((0,1) \cup (1,0)) \stackrel{\text{disjunkte}}{\downarrow} = P((0,1)) + P((1,0))$$

$$P(Y=2) = P((0,2) \cup (1,1) \cup (2,0)) \stackrel{\text{disjunkte}}{\uparrow} = \sum_{X_A=0}^2 P((X_A, 2-X_A))$$

forhåber litt

$$P(Y=y) = \sum_{X_A=0}^y P((X_A, y-X_A)) = \sum_{X_A=0}^y P(X_A=X_A, X_B=y-X_A)$$

men hva er dette?

$\uparrow$ uavh. $X_A=0$

X_A : antall hendelser i prosess A over tid t

$$X_A \sim \text{Poisson}(\lambda_A t)$$

X_B : antall hendelser i prosess B over tid t

$$X_B \sim \text{Poisson}(\lambda_B t)$$

X_A, X_B uavhengige

$$= \sum_{x_A=0}^y \frac{(\lambda_A t)^{x_A}}{x_A!} e^{-\lambda_A t} \frac{(\lambda_B t)^{y-x_A}}{(y-x_A)!} e^{-\lambda_B t} \cdot \frac{y!}{y!}$$

$$= \frac{e^{-(\lambda_A + \lambda_B)t}}{y!} \sum_{x_A=0}^y (\lambda_A t)^{x_A} (\lambda_B t)^{y-x_A} \frac{y!}{x_A! (y-x_A)!}$$

$$= \frac{e^{-(\lambda_A + \lambda_B)t}}{y!} \underbrace{\sum_{x_A=0}^y \binom{y}{x_A} (\lambda_A t)^{x_A} (\lambda_B t)^{y-x_A}}_{\substack{\text{Binom} \\ \text{formel}}} \quad \text{Binom} \\ \text{formel}$$

$$= (\lambda_A t + \lambda_B t)^y = ((\lambda_A + \lambda_B)t)^y$$

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

$$\Rightarrow Y \sim \text{Poisson}((\lambda_A + \lambda_B)t)$$

Celledeling

$n = 3 \times 10^9$ basepar langs genomet,
alle skal kopieres.

En feil ved kopiering kalles mutasjon



$$P(\text{mutasjon}) = 5 \cdot 10^{-10} \quad (\text{p\aa hver posisjon})$$

Vi kan anta uavhengighet mellom posisjoner

X = antall mutasjoner i en celledeling

$$X \sim \text{binom}(n = 3 \cdot 10^9, p = 5 \cdot 10^{-10})$$

$$E(X) = 1.5 \quad (\text{Plott } f(x))$$

$$Y \sim \text{Poisson}(1.5) \quad (\text{Plott } f(y)) \quad \text{obs. samme}$$

OPPG. 3

$$Z_1 = \begin{cases} 1 & \text{hvis } \bar{X}_1 = \bar{X}_3 = 0 \\ 0 & \text{ellers} \end{cases}$$

(uavhengig av $\bar{X}_2$)

$$P(Z_1 = 1) = P(\bar{X}_1 = 0, \bar{X}_3 = 0)$$

uavh. $\rightarrow$

$$\begin{aligned} &= P(\bar{X}_1 = 0) P(\bar{X}_3 = 0) \\ &= e^{-\mu_1} e^{-\mu_3} \\ &= e^{-(\mu_1 + \mu_3)} = p_1 \end{aligned}$$

$$P(X_1 = 0) = \frac{\mu_1^0}{0!} e^{-\mu_1}$$

$$P(Z_1 = 0) = 1 - P(\bar{X}_1 = 0, \bar{X}_3 = 0) = 1 - p_1$$

$\nwarrow$ Bernoulli /
 $\swarrow$ binomisk(1, p_1)

$$E(Z_1) = p_1 \quad \text{Var}(Z_1) = p_1(1 - p_1)$$

(np)

(Samme for Z_2 med μ_2)

$$\text{Cov}(Z_1, Z_2) = E(Z_1 Z_2) - E(Z_1) E(Z_2)$$

↑
ihke uavhengige,
begge inneholder
 X_3

↑
2

$$\begin{matrix} \uparrow & \uparrow \\ P_1 & P_2 \\ = e^{-(\mu_1 + \mu_3)} & = e^{-(\mu_2 + \mu_3)} \end{matrix}$$

$$E(Z_1 Z_2) = \sum_{z_1=0}^1 \sum_{z_2=0}^1 z_1 z_2 P(Z_1 = z_1, Z_2 = z_2)$$

$$\begin{aligned} &= 0 P(Z_1=0, Z_2=0) + 0 P(Z_1=0, Z_2=1) \\ &+ 0 P(Z_1=1, Z_2=0) + 1 P(Z_1=1, Z_2=1) \end{aligned}$$

$$\begin{aligned} P(Z_1=1, Z_2=1) &= P(X_1=0, X_2=0, X_3=0) \\ &= e^{-\mu_1} e^{-\mu_2} e^{-\mu_3} \\ &= e^{-(\mu_1 + \mu_2 + \mu_3)} \end{aligned}$$

$$\begin{aligned} \text{Cov}(Z_1, Z_2) &= e^{-(\mu_1 + \mu_2 + \mu_3)} - e^{-(\mu_1 + \mu_3 + \mu_2 + \mu_3)} = \\ &= e^{-(\mu_1 + \mu_2 + \mu_3)} (1 - e^{-\mu_3}) \end{aligned}$$