Repetisjon av negativ binomisk fordeling, geometrisk fordeling og poissonfordeling

- Hvis X er antall forsøk til suksess k oppnås i en Bernoulliprosess med suksessannsynlighet p, har X en **negativ binomisk fordeling**. Vi skriver $X \sim \text{NBinomial}(k, p)$ og punktsannsynligheten er gitt ved

$$b^*(x; k, p) = \begin{cases} \binom{x-1}{k-1} p^k (1-p)^{x-k}, & x = k, k+1, \dots, \\ 0, & \text{ellers.} \end{cases}$$

- Hvis X er antall forsøk til første suksess oppnås i en Bernoulliprosess med suksessannsynlighet p, har X en **geometrisk fordeling**. Vi skriver $X \sim \text{Geo}(p)$ og punktsannsynligheten er gitt ved

$$g(x; p) = \begin{cases} p^k (1-p)^{x-1}, & x = 1, 2, \dots, \\ 0, & \text{ellers.} \end{cases}$$

- **Viktig**:
  - Her betegner X «antall forsøk» slik at $x = k, k+1, \dots$, men **vær obs** på at det også er vanlig å la X betegne «antall fiasko» til suksess k er oppnådd. Da er $x = 0, 1, \dots$ uavhengig av verdien av k.
  - Den geometriske fordelingen er et spesialtilfelle av den negativ binomiske fordelingen, $\text{Geo}(p) \sim \text{NBinomial}(1, p)$.

- Hvis $X \sim \text{Geo}(p)$ så er
  1. $\text{E}[X] = 1/p$
  2. $\text{Var}[X] = (1-p)/p^2$

- En **poissonprosess** i tid med rate $\lambda > 0$ kan defineres (vi gjør det uformelt i dette emnet) utfra de tre betingelsene:
  - Antall hendelser i to disjunkte tidintervaller er uavhengige stokastiske variabler.
  - Sannsynligheten for at 1 hendelse inntreffer i et lite interval $(t, t + \Delta t]$ er omtrent $\lambda \Delta t$.
  - Sannsynligheten for at mer enn 1 hendelse inntreffer på et lite tidsintervall $(t, t + \Delta t]$ er neglisjerbar.

- Antall hendelser i tidsintervallet $(0, t]$ for en poissonprosess med rate $\lambda > 0$ har en **poissonfordeling**. Vi skriver $X \sim \text{Poisson}(\lambda t)$ og punktsannsynligheten er gitt ved

$$p(x; \lambda t) = \begin{cases} \frac{(\lambda t)^x}{x!} \mathrm{e}^{-\lambda t}, & x = 0, 1, \dots, \\ 0, & \text{ellers.} \end{cases}$$

  Raten λ betegner forventet (gjennomsnittlig) antall hendelser per tidsenhet.

- For $X \sim \text{Poisson}(\lambda t)$ har vi at

  1. $\mu = \mathrm{E}[X] = \lambda t$
  2. $\mathrm{Var}[X] = \lambda t = \mu$

  Dette betyr at poissonfordelingen er parametrisert gjennom sin forventningsverdi μ.

- For n stor og p liten kan $\text{Binomial}(n, p)$ tilnærmes ved fordelingen $\text{Poisson}(np)$. Dvs, at for n stor og p liten er

$$b(x; n, p) \approx p(x; np), \quad x = 0, 1, \dots, n.$$

- Man kan konstruere en **poissonprosess** på tidsintervallet $(0, t]$ med rate $\lambda > 0$ som en grense av en bernoulliprosess ved følgende steg:

  1. **Regulær oppdeling**:
     * $t_i = t \cdot i/n$ for $i = 0, 1, \dots, n$.
     * $I_i = \begin{cases} 1, & \text{hendelse på } (t_{i-1}, t_i], \\ 0, & \text{ingen hendelse på } (t_{i-1}, t_i] \end{cases}, \quad i = 1, 2, \dots, n.$
  2. **Suksessannsynlighet proposjonal med intervalbredde**:
     * $p_n = \lambda \cdot t/n$
  3. **Uavhengige tidsintervaller**:
     * $I_1, I_2, \dots, I_n \sim \text{Bernoulli}(p_n)$ og uavhengige
  4. **Ta grensen $n \to \infty$ (kontinuerlig tid)**:
     * La $X(s) =$ «Antall hendelser på $(0, s]$» $= \sum_{i: t_i \leq s} I_i$
     * Under grensen $n \to \infty$ er $X(s) \sim \text{Poisson}(\lambda s)$ for $0 \leq s \leq t$

  Formelt er en poissonprosess en **«stokastisk prosess»** og en realisering av en poissonprosess er en kurve $s \mapsto x(s)$ i tid; se for eksempel «TMA4265» for detaljer.