

Geometrisk og negativ binomisk fordeling

TMA4240/TMA4245 Statistikk

Håkon Tjelmeland

Institutt for matematiske fag

Norges teknisk-naturvitenskapelige universitet

Geometrisk og negativ binomisk fordeling

- ★ Husk: Definisjonen av en Bernoulli forsøksrekke

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.

Definisjon (Negativ binomisk fordeling)

Anta at vi har en Bernoulli forsøksrekke hvor vi gjentar forsøket uendelig mange ganger, og la X være antall forsøk som må til for å få k suksesser. Vi sier da at X har en negativ binomisk fordeling med parametre p og k

- ★ Realisasjon av en Bernoulli forsøksrekke med $p = 0.25$



- for $k = 2$ blir da $x = 6$

- ★ Hva blir $f(x) = P(X = x)$?

- regner først på tilfellet $k = 1$ (dette kalles en geometrisk fordeling med parameter p)

Utleddning av $f(x)$ for geometrisk fordeling

A og B uavhengige hendelser: $P(A \cap B) = P(A) \cdot P(B)$

★ Punktsannsynlighet:

$$\begin{aligned}f(x) &= P(X = x) \\&= P(FF \dots FS) \\&= P(F)P(F) \cdot \dots \cdot P(F)P(S) \\&= (1 - p)^{x-1}p \\&= p(1 - p)^{x-1} \text{ for } x = 1, 2, \dots\end{aligned}$$

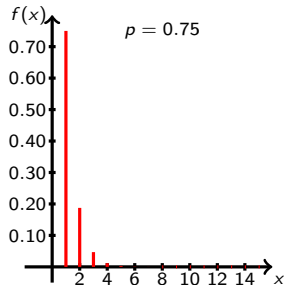
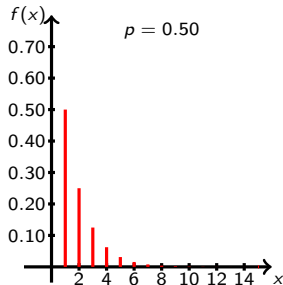
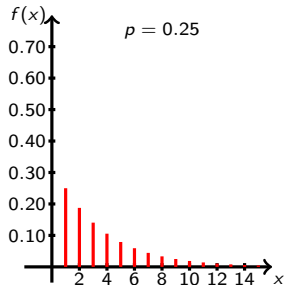
★ Forventningsverdi og varians: Man kan vise at

$$\begin{aligned}E[X] &= \frac{1}{p} \\ \text{Var}[X] &= \frac{1 - p}{p^2}\end{aligned}$$

Punktsannsynlighet i geometrisk fordeling

★ Punktsannsynlighet

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



Utleddning av $f(x)$ i negativ binomisk fordeling

A og B uavhengige hendelser: $P(A \cap B) = P(A) \cdot P(B)$

★ Punktsannsynlighet

Binomisk fordeling: $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$

$$\begin{aligned} f(x) &= P(X = x) \\ &= P(\dots S) \quad (\dots \text{ betyr } x-1 \text{ forsøk med nøyaktig } k-1 \text{ suksesser}) \\ &= P(\dots) \cdot P(S) \\ &= \binom{x-1}{k-1} p^{k-1} (1-p)^{(x-1)-(k-1)} \cdot p \\ &= \binom{x-1}{k-1} p^k (1-p)^{x-k} \text{ for } x = k, k+1, \dots \end{aligned}$$

★ La Z_1 : antall forsøk som må til for å få en suksess

Z_2 : antall ytterligere forsøk som må til for å få enda en suksess

$\vdots$

Z_k : antall forsøk etter suksess nummer $k-1$ som må til for å få enda en suksess



★ Da er $Z_1, Z_2, \dots, Z_k$ uavhengige og geometrisk fordelt med parameter p , og vi har at

$$X = \sum_{i=1}^k Z_i$$

★ Dermed

$$E[X] = E\left[\sum_{i=1}^k Z_i\right] = \sum_{i=1}^k E[Z_i] = \sum_{i=1}^k \frac{1}{p} = \frac{k}{p}, \quad \text{Var}[X] = \text{Var}\left[\sum_{i=1}^k Z_i\right] = \dots = k \cdot \frac{1-p}{p^2}$$

Kumulativ fordeling og kvantiler for negativ binomisk fordeling

- ★ Punktsannsynlighet

$$f_X(x) = P(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, x = k, k+1, \dots, n$$

- ★ Kumulativ fordelingsfunksjon for $x = 0, 1, 2, \dots, n$

$$F_X(x) = P(X \leq x) = \sum_{t=0}^x f_X(t) = \sum_{t=0}^x \binom{t-1}{k-1} p^k (1-p)^{t-k}$$

- for de fleste verdier av k og p har vi intet analytisk uttrykk for summen
- NB: for $k = 1$ (geometrisk fordeling) kan vi finne uttrykk for $F(x)$

- ★ For å finne en sannsynlighet $P(X \leq x)$ når $k > 1$ for gitte verdier av k , p og x kan man

- numerisk regne ut summen
- kalle en funksjon som gir verdien
- slå opp verdien i en tabell

- ★ For å finne en kvantil $x = \min\{x : F_X(x) \geq q\}$ når $k > 1$ for gitte verdier av n , p og q kan man

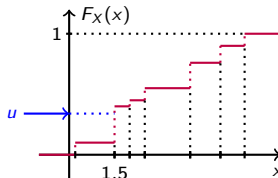
- regne den ut numerisk
- kalle en funksjon som gir verdien
- lete seg frem i en tabell over $P(X \leq x)$

Simulering fra negativ binomisk fordeling

- ★ Har tidligere diskutert generell algoritme for å simulere fra en diskret fordeling

$$U \sim \text{Unif}[0, 1]$$

$$X = \min\{x : F_X(x) \geq U\}$$



- ★ Alternativ: Simulere Bernoulli forsøksrekke og telle opp verdien til X

Initier $n\text{Sukcess} = 0$

Initier $X = 0$

While ($n\text{Sukcess} < k$)

 Simuler $U_i \sim \text{Unif}[0, 1]$

 Regn ut $Z_i = \mathbb{I}(U_i \leq p)$

 Regn ut $n\text{Sukcess} = n\text{Sukcess} + Z_i$

 Regn ut $X = X + 1$

End

Oppsummering

- ★ Har definert geometrisk og negativ binomisk fordeling
- ★ Har utledet
 - punktsannsynlighet, $f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$, $x = k, k+1, \dots$
 - forventningsverdi, $E[X] = \frac{k}{p}$
 - varians, $\text{Var}[X] = k \cdot \frac{1-p}{p^2}$
- ★ Har diskutert hvordan man kan
 - finne sannsynligheter $P(X \leq x)$
 - NB: For geometrisk fordeling kan man finne analytisk uttrykk for $F(x)$
 - finne kvantiler, $x = \min\{F(x) \geq q\}$
 - simulere en realisasjon ved å simulere en Bernoulli forsøksrekke