Random walks and branching processes

Bo Friis Nielsen

1 DTU Informatics

02407 Stochastic Processes 2, September 8 2020

Simple random walk with two reflecting barriers 0 and \( N \)

\[
P = \begin{pmatrix}
1 & 0 & 0 & \ldots & 0 & 0 & 0 \\
q & 0 & p & \ldots & 0 & 0 & 0 \\
0 & q & 0 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & q & 0 & p \\
0 & 0 & 0 & \ldots & 0 & 0 & 1 \\
\end{pmatrix}
\]

\( T = \min\{ n \geq 0; X_n \in \{0, 1\} \} \)

\( u_k = \mathbb{P}\{X_T = 0 | X_0 = k \} \)

Discrete time Markov chains

Today:
- Random walks
- First step analysis revisited
- Branching processes
- Generating functions

Next week
- Classification of states
- Classification of chains
- Discrete time Markov chains - invariant probability distribution

Two weeks from now
- Poisson process

Solution technique for \( u_k' \)s

\( u_k = pu_{k+1} + qu_{k-1}, \quad k = 1, 2, \ldots, N - 1, \)

\( u_0 = 1, \)

\( u_N = 0 \)

Rewriting the first equation using \( p + q = 1 \) we get

\[
(p + q)u_k = pu_{k+1} + qu_{k-1} \iff 0 = p(u_{k+1} - u_k) - q(u_k - u_{k-1}) \iff x_k = (q/p)x_{k-1}
\]

with \( x_k = u_k - u_{k-1} \), such that

\[
x_k = (q/p)^{k-1}x_1
\]
Recovering \( u_k \)

\[
\begin{align*}
x_1 &= u_1 - u_0 = u_1 - 1 \\
x_2 &= u_2 - u_1 \\
& \vdots \\
x_k &= u_k - u_{k-1}
\end{align*}
\]

such that

\[
\begin{align*}
  u_1 &= x_1 + 1 \\
u_2 &= x_2 + x_1 + 1 \\
& \vdots \\
u_k &= x_k + x_{k-1} + \cdots + 1 = 1 + x_1 \sum_{i=0}^{k-1} \frac{q}{p}^i 
\end{align*}
\]

Values of absorption probabilities \( u_k \)

From \( u_N = 0 \) we get

\[
\begin{align*}
  0 &= 1 + x_1 \sum_{i=0}^{N-1} \left( \frac{q}{p} \right)^i \\
  x_1 &= -\frac{1}{\sum_{i=0}^{N-1} \left( \frac{q}{p} \right)^i}
\end{align*}
\]

Leading to

\[
\begin{align*}
  u_k &= \begin{cases} 
    1 - \left( \frac{k}{N} \right) = \left( \frac{N-k}{N} \right) & \text{when } \ p = q = \frac{1}{2} \\
    \frac{(q/p)^k - (q/p)^N}{1 - (q/p)} & \text{when } p \neq q
  \end{cases}
\end{align*}
\]

Expected number of visits to states

\[
W^{(n)}_{ij} = Q^{(0)} + Q^{(1)} + \cdots + Q^{(n)}
\]

In matrix notation we get

\[
W^{(n)} = I + Q + Q^2 + \cdots + Q^n = I + Q \left( I + Q + \cdots + Q^{n-1} \right) = I + QW^{(n-1)}
\]

Elementwise we get the “first step analysis” equations

\[
W^{(n)}_{ij} = \delta_{ij} + \sum_{k=0}^{n-1} P_{ik} W^{(n-1)}_{kj}
\]

Direct calculation as opposed to first step analysis

\[
P = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix}
\]

\[
P^2 = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix} \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix} = \begin{bmatrix} Q^2 & QR + R \\ 0 & I \end{bmatrix}
\]

\[
P^n = \begin{bmatrix} Q^n & Q^{n-1}R + Q^{n-2}R + \cdots + QR + R \\ 0 & I \end{bmatrix}
\]

\[
W^{(n)}_{ij} = E \left[ \sum_{\ell=0}^{n} \mathbb{I}(X_{\ell} = j) \mid X_0 = i \right], \quad \text{where} \quad \mathbb{I}(X_{\ell}) = \begin{cases} 
  1 & \text{if } X_{\ell} = j \\
  0 & \text{if } X_{\ell} \neq j
\end{cases}
\]
Limiting equations as $n \to \infty$

$$W = I + Q + Q^2 + \cdots = \sum_{i=0}^{\infty} Q^i$$

$$W = I + QW$$

From the latter we get

$$(I - Q)W = I$$

When all states related to $Q$ are transient (we have assumed that) we have

$$W = \sum_{i=0}^{\infty} Q^i = (I - Q)^{-1}$$

With $T = \min\{n \geq 0, r \leq X_n \leq N\}$ we have that

$$W_{ij} = E\left[\sum_{n=0}^{T-1} \mathbb{1}(X_n = j) \mid X_0 = i\right]$$

Absorption time

$$\sum_{n=0}^{T-1} \sum_{j=0}^{r} \mathbb{1}(X_n = j) = \sum_{n=0}^{T-1} 1 = T$$

Thus

$$E(T \mid X_0 = i) = E\left[\sum_{j=0}^{r} \sum_{n=0}^{T-1} \mathbb{1}(X_n = j) X_0 = i\right]$$

$$= \sum_{j=0}^{r} E\left[\sum_{n=0}^{T-1} \mathbb{1}(X_n = j) \mid X_0 = i\right]$$

$$= \sum_{j=0}^{r} W_{ij}$$

In matrix formulation

$$v = W1$$

where $v_i = E(T \mid X_0 = i)$ as last week, and $1$ is a column vector of ones.

Absorption probabilities

$$U_{ij}^{(n)} = P\{T \leq n, X_T = j \mid X_0 = i\}$$

$$U^{(1)} = R = IR$$

$$U^{(2)} = IR + QR$$

$$U^{(n)} = (I + Q + \cdots + Q^{(n-1)}) R = W^{(n-1)} R$$

Leading to

$$U = WR$$

Conditional expectation discrete case (2.1)

$$P\{Y = y \mid X = x\} = \frac{P\{X = x, Y = y\}}{P\{X = x\}}$$

$$E[Y \mid X = x] = \sum_y y P\{Y = y \mid X = x\}$$

$h(x) = E[Y \mid X = x]$ is a function of $x$, thus $h(X)$ is a random variable, which we call $E[Y \mid X]$. Now

$$E[h(X)] = \sum_x P\{X = x\} h(x) = \sum_x P\{X = x\} \sum_y y P\{Y = y \mid X = x\}$$

$$= \sum_x \sum_y y P\{X = x, Y = y\} = \sum_x \sum_y y P\{X = x, Y = y\}$$

$$= E[Y] = E\{E[Y \mid X]\}, \quad (E[g(Y)] = E\{E[g(Y) \mid X]\})$
Conditional variance discrete case

\[ \text{Var}[Y] = E[Y^2] - E[Y]^2 = E\{E[Y^2 \mid X]\} - E[Y]^2 \]

\[ = E\{\text{Var}[Y \mid X] + E[Y^2 \mid X]\} - E[E[Y \mid X]]^2 \]

\[ = E\{\text{Var}[Y \mid X]\} + E\{E[Y^2 \mid X]\} - E\{E[Y \mid X]\}^2 \]

\[ = E\{\text{Var}[Y \mid X]\} + \text{Var}[E[Y \mid X]] \]

Random sum (2.3)

\[ X = \xi_1 + \cdots + \xi_N = \sum_{i=1}^{N} \xi_i \]

where \( N \) is a random variable taking values among the non-negative integers; with

\[ E(N) = \nu, \text{Var}(N) = \tau^2, E(\xi_i) = \mu, \text{Var}(\xi_i) = \sigma^2 \]

\[ E(X) = E(E(X \mid N)) = E(N \mu) = \nu \mu \]

\[ \text{Var}(X) = E(\text{Var}(X \mid N)) + \text{Var}(E(X \mid N)) \]

\[ = E(N \sigma^2) + \text{Var}(N \mu) = \nu \sigma^2 + \tau^2 \mu^2 \]

Branching processes

\[ X_{n+1} = \xi_1 + \xi_2 + \cdots + \xi_n \]

where \( \xi_i \) are independent random variables with common propability mass function

\[ P\{\xi_i = k\} = p_k \]

From a random sum interpretation we get

\[ E(X_{n+1}) = \mu E(X_n) = \mu^{n+1} \]

\[ \text{Var}(X_{n+1}) = \sigma^2 E(X_n) + \mu \text{Var}(X_n) = \sigma^2 \mu^n + \mu^2 \text{Var}(X_n) \]

\[ = \sigma^2 \mu^n + \mu^2 (\sigma^2 \mu^{n-1} + \mu^2 \text{Var}(X_{n-1})) \]

Extinction probabilities

Define \( N \) to be the random time of extinction (\( N \) can be defective - i.e. \( P\{N = \infty\} > 0\))

\[ u_n = P\{N \leq n\} = P\{X_N = 0\} \]

And we get

\[ u_n = \sum_{k=0}^{\infty} p_k u_{n-1}^k \]
The generating function - an important analytic tool

- Manipulations with probability distributions
- Determining the distribution of a sum of random variables
- Determining the distribution of a random sum of random variables
- Calculation of moments
- Unique characterisation of the distribution
- Same information as CDF

Generating functions

\[ \phi(s) = \mathbb{E}\left(s^\xi\right) = \sum_{k=0}^{\infty} p_k s^k, \quad p_k = \frac{1}{k!} \left. \frac{d^k \phi(s)}{ds^k}\right|_{s=0} \]

Moments from generating functions

\[ \left. \frac{d\phi(s)}{ds}\right|_{s=1} = \sum_{k=1}^{\infty} p_k k s^{k-1}, \quad \mathbb{E}(\xi) \]

Similarly

\[ \left. \frac{d^2 \phi(s)}{ds^2}\right|_{s=1} = \sum_{k=2}^{\infty} p_k k(k-1) s^{k-2}, \quad \mathbb{E}(\xi - 1) \]

A factorial moment

\[ \text{Var}(\xi) = \phi''(1) + \phi'(1) - (\phi'(1))^2 \]

The sum of iid random variables

Remember Independent Identically Distributed

\[ S_n = X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i \]

With \( p_x = P\{X_i = x\}, \quad X_i \geq 0 \) we find for \( n = 2 \)

\[ S_2 = X_1 + X_2 \]

The event \( \{S_2 = x\} \) can be decomposed into the set

\[ \{(X_1 = 0, X_2 = x), (X_1 = 1, X_2 = x - 1)\}
\]

\[ \cdots (X_1 = i, X_2 = x - i), \cdots (X_1 = x, X_2 = 0)\}\]

The probability of the event \( \{S_2 = x\} \) is the sum of the probabilities of the individual outcomes.

The Probability of outcome \( (X_1 = i, X_2 = x - i) \) is \( P\{X_1 = i\} P\{X_2 = x - i\} \) by independence, which again is \( p_i p_{x-i} \).

In total we get

\[ P\{S_2 = x\} = \sum_{i=0}^{x} p_i p_{x-i} \]
Generating function - one example

Binomial distribution

\[ p_k = \binom{n}{k} p^k (1-p)^{n-k} \]

\[ \phi_{bin}(s) = \sum_{k=0}^{n} s^k p_k = \sum_{k=0}^{n} s^k \binom{n}{k} p^k (1-p)^{n-k} \]

\[ = \sum_{k=0}^{n} \binom{n}{k} (sp)^k (1-p)^{n-k} = (1-p+ps)^n \]

Generating function - another example

Poisson distribution

\[ p_k = \frac{\lambda^k}{k!} e^{-\lambda} \]

\[ \phi_{poi}(s) = \sum_{k=0}^{\infty} s^k p_k = \sum_{k=0}^{\infty} s^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(s\lambda)^k}{k!} \]

\[ = e^{-\lambda} e^{s\lambda} = e^{-\lambda(1-s)} \]

And now to the reason for all this ...

The convolution can be tough to deal with (sum of random variables)

**Theorem**

If \( X \) and \( Y \) are independent then

\[ \phi_{X+Y}(s) = \phi_X(s) \phi_Y(s) \]

where \( \phi_X(s) \) and \( \phi_Y(s) \) are the generating functions of \( X \) and \( Y \)

A probabilistic proof (which I think is instructive)

\[ \phi_{X+Y}(s) = E\left( s^{X+Y} \right) = E\left( s^X s^Y \right) = E\left( s^X \right) E\left( s^Y \right) = \phi_X(s) \phi_Y(s) \]

Sum of two Poisson distributed random variables

\( X \sim P(\lambda) \quad Y \sim P(\mu) \quad Z = X + Y \)

\[ \phi_X(s) = e^{-\lambda(1-s)} \quad \phi_Y(s) = e^{-\mu(1-s)} \quad P\{X = x\} = p_x = \lambda^x \frac{e^{-\lambda}}{x!} \]

And we get

\[ \phi_Z(s) = \phi_X(s) \phi_Y(s) = e^{-\lambda(1-s)} e^{-\mu(1-s)} = e^{-(\lambda+\mu)(1-s)} \]

Such that

\[ Z \sim P(\lambda + \mu) \]
Sum of two Binomial random variables with the same $p$

\[ X \sim B(n, p) \quad Y \sim B(m, p) \quad Z = X + Y \]

\[ \phi_X(s) = (1 - p + ps)^n \quad \phi_Y(s) = (1 - p + ps)^m \]

And we get

\[ \phi_Z(s) = \phi_X(s) \phi_Y(s) = (1 - p + ps)^n (1 - p + ps)^m = (1 - p + ps)^{n+m} \]

Such that

\[ Z \sim B(n + m, p) \]

Poisson example

\[ X \sim P(\lambda) \quad \phi_X(s) = e^{-\lambda(1-s)} \quad \left( P\{X = x\} = p_x = \frac{\lambda^x e^{-\lambda}}{x!} \right) \]

\[ \phi'(s) = -(-\lambda)e^{-\lambda(1-s)} = \lambda e^{-\lambda(1-s)} \]

And we find

\[ E(X) = \phi'(1) = \lambda e^0 = 1 \quad \phi''(s) = \lambda^2 e^{-\lambda(1-s)} \quad V(X) = \phi''(1) + \phi'(1) - (\phi'(1))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \]

Generating function - the geometric distribution

\[ \phi_{geo}(s) = \sum_{x=1}^{\infty} p_x = \left( \frac{1 - p}{1 - s} \right)^{x-1} p \]

\[ = \sum_{x=1}^{\infty} s(s(1 - p))^{x-1} p \]

A useful power series is:

\[ \sum_{i=0}^{N} a_i = \begin{cases} 1 - \frac{a^{N+1}}{1-a} & N < \infty, a \neq 1 \\ N+1 & N < \infty, a = 1 \\ \frac{1}{1-a} & N = \infty, |a| < 1 \end{cases} \]

And we get \[ \phi_{geo}(s) = \frac{sp}{1 - s(1 - p)} \]

Generating function for random sum

\[ h_X(s) = g_N(\phi(s)) \]

Applied for the branching process we get

\[ \phi_n(s) = \phi_{n-1}(\phi(s)) \]
Generating function for the sum of independent random variables

\[ X \text{ with pdf } p_x \quad Y \text{ with pdf } q_y \]

\[ Z = X + Y \text{ what is } r_z = P\{Z = z\}? \]

\[ P\{Z = z\} = r_z = \sum_{i=0}^{z} p_i q_{z-i} \]

**Theorem**

(23) page 153 If \( X \) and \( Y \) are independent then

\[ \phi_{X+Y}(s) = \phi_X(s) \phi_Y(s) \]

where \( \phi_X(s) \) and \( \phi_Y(s) \) are the generating functions of \( X \) and \( Y \)

---

**Sum of two geometric random variables with the same \( p \)**

\[ X \sim \text{geo}(p) \quad Y \sim \text{geo}(p) \quad Z = X + Y \]

\[ \phi_X(s) = \frac{sp}{1 - s(1 - p)} \quad \phi_Y(s) = \frac{sp}{1 - s(1 - p)} \]

\[ P\{X = x\} = p_x = (1 - p)^{x-1} p \]

And we get

\[ \phi_Z(s) = \phi_X(s) \phi_Y(s) = \left( \frac{sp}{1 - s(1 - p)} \right)^2 \]

The density of this distribution is

\[ P\{Z = z\} = h(z) = (z - 1)(1 - p)^{z-2} p^2 \]

Negative binomial.

---

**Sum of \( k \) geometric random variables with the same \( p \)**

More generally - sum of \( k \) geometric variables

\[ p_x = \left( \frac{x - 1}{k - 1} \right) (1 - p)^{x-k} p^k \quad \phi_X(s) = \left( \frac{sp}{1 - s(1 - p)} \right)^k \]

---

**Characteristic function and other**

- Characteristic function: \( E\left( e^{itX} \right) \)
- Moment generating function: \( E\left( e^{\theta X} \right) \)
- Laplace Stieltjes transform: \( E\left( e^{-sX} \right) \)

**EXAMPLE:** (exponential)

\[ E\left( e^{\theta X} \right) = \int_0^\infty e^{\theta x} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - \theta}, \theta < \lambda \]