Random walks and branching processess

Bo Friis Nielsen[1]

[1]DTU Informatics

02407 Stochastic Processes 2, September 8 2020

DTU

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

DTU

Simple random walk with two reflecting barriers 0 and N

$$\boldsymbol{P} = \left\| \begin{array}{ccccccc} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & \dots & 0 & 0 & 0 \\ 0 & q & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{array} \right\|$$

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

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

DTU

Solution technique for $u_k' s$

$$\begin{aligned} u_k &= pu_{k+1} + qu_{k-1}, \qquad k = 1,2,\dots,N-1, \\ u_0 &= 1, \\ u_N &= 0 \end{aligned}$$

Rewriting the first equation using $p + q = 1$ we get

$$\begin{aligned} (p+q)u_k &= pu_{k+1} + qu_{k-1} \Leftrightarrow \\ 0 &= p(u_{k+1} - u_k) - q(u_k - u_{k-1}) \Leftrightarrow \\ x_k &= (q/p)x_{k-1} \end{aligned}$$

with $x_k = u_k - u_{k-1}$, such that

$$x_k = (q/p)^{k-1} x_1$$

DTU

Recovering u_k

$$\begin{aligned} x_1 &= u_1 - u_0 = u_1 - 1 \\ x_2 &= u_2 - u_1 \\ &\vdots \\ x_k &= u_k - u_{k-1} \end{aligned}$$

such that

$$\begin{aligned} u_1 &= x_1 + 1 \\ u_2 &= x_2 + x_1 + 1 \\ &\vdots \\ u_k &= x_k + x_{k-1} + \cdots + 1 = 1 + x_1 \sum_{i=0}^{k-1} (q/p)^i \end{aligned}$$

Values of absorption probabilities u_k

From $u_N = 0$ we get

$$\begin{aligned} 0 &= 1 + x_1 \sum_{i=0}^{N-1} (q/p)^i \Leftrightarrow \\ x_1 &= -\frac{1}{\sum_{i=0}^{N-1} (q/p)^i} \end{aligned}$$

Leading to

$$u_k = \begin{cases} 1 - (k/N) = (N-k)/N & \text{when } p = q = \frac{1}{2} \\ \frac{(q/p)^k - (q/p)^N}{1-(q/p)^N} & \text{when } p \neq q \end{cases}$$

Direct calculation as opposed to first step analysis

$$\boldsymbol{P} = \left\| \begin{array}{cc} \boldsymbol{Q} & \boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I} \end{array} \right\|$$

$$\boldsymbol{P}^2 = \left\| \begin{array}{cc} \boldsymbol{Q} & \boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I} \end{array} \right\| \left\| \begin{array}{cc} \boldsymbol{Q} & \boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I} \end{array} \right\| = \left\| \begin{array}{cc} \boldsymbol{Q}^2 & \boldsymbol{QR} + \boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I} \end{array} \right\|$$

$$\boldsymbol{P}^n = \left\| \begin{array}{cc} \boldsymbol{Q}^n & \boldsymbol{Q}^{n-1}\boldsymbol{R} + \boldsymbol{Q}^{n-2}\boldsymbol{R} + \cdots + \boldsymbol{QR} + \boldsymbol{R} \\ \boldsymbol{0} & \boldsymbol{I} \end{array} \right\|$$

$$W_{ij}^{(n)} = \mathbb{E}\left[\sum_{\ell=0}^{n} \mathbb{1}(X_\ell = j) | X_0 = i \right], \quad \text{where } \mathbb{1}(X_\ell) = \begin{cases} 1 & \text{if } X_\ell = j \\ 0 & \text{if } X_\ell \neq j \end{cases}$$

Expected number of visits to states

$$W_{ij}^{(n)} = Q_{ij}^{(0)} + Q_{ij}^{(1)} + \dots Q_{ij}^{(n)}$$

In matrix notation we get

$$\begin{aligned} \boldsymbol{W}^{(n)} &= \boldsymbol{I} + \boldsymbol{Q} + \boldsymbol{Q}^2 + \cdots + \boldsymbol{Q}^n \\ &= \boldsymbol{I} + \boldsymbol{Q}\left(\boldsymbol{I} + \boldsymbol{Q} + \cdots + \boldsymbol{Q}^{n-1}\right) \\ &= \boldsymbol{I} + \boldsymbol{Q}\boldsymbol{W}^{(n-1)} \end{aligned}$$

Elementwise we get the “first step analysis” equations

$$W_{ij}^{(n)} = \delta_{ij} + \sum_{k=0}^{r-1} P_{ik} W_{kj}^{(n-1)}$$

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

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

Absorption time

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

Thus

$$\begin{aligned}\mathbb{E}(T|X_0 = i) &= \mathbb{E}\left[\sum_{j=0}^{r}\sum_{n=0}^{T-1} \mathbb{1}(X_n = j)\, X_0 = i\right] \\ &= \sum_{j=0}^{r} \mathbb{E}\left[\sum_{n=0}^{T-1} \mathbb{1}(X_n = j|X_0 = i\right] \\ &= \sum_{j=0}^{r} W_{ij}\end{aligned}$$

In matrix formulation

$$v = W\mathbf{1}$$

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

Absorption probabilities

$$U_{ij}^{(n)} = \mathbb{P}\{T \leq n, X_T = j|X_0 = i\}$$
$$U^{(1)} = R = IR$$
$$U^{(2)} = IR + QR$$
$$U^{(n)} = (I + Q + \cdots + Q^{(n-1)}R = W^{(n-1)}R$$

Leading to

$$U = WR$$

Conditional expectation discrete case (2.1)

$$\mathbb{P}\{Y = y|X = x\} = \frac{\mathbb{P}\{X = x, Y = y\}}{\mathbb{P}\{X = x\}}$$
$$\mathbb{E}[Y = y|X = x] = \sum_y y\mathbb{P}\{Y = y|X = x\}$$

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

$$\mathbb{E}[h(X)] = \sum_x \mathbb{P}\{X = x\}h(x) = \sum_x \mathbb{P}\{X = x\}\sum_y y\mathbb{P}\{Y = y|X = x\}$$
$$= \sum_x\sum_y y\mathbb{P}\{X = x\}\frac{\mathbb{P}\{X = x, Y = y\}}{\mathbb{P}\{X = x\}} = \sum_x\sum_y y\mathbb{P}\{X = x, Y = y\}$$
$$= \mathbb{E}[Y] = \mathbb{E}\{\mathbb{E}[Y|X]\}, \quad (\mathbb{E}[g(Y)] = \mathbb{E}\{\mathbb{E}[g(Y)|X]\})$$

Conditional variance discrete case

$$\mathbb{V}\text{ar}[Y] = \text{E}\left[Y^2\right] - \text{E}\left[Y\right]^2 = \mathbb{E}\{\mathbb{E}[Y^2|X]\} - \text{E}\left[Y\right]^2$$
$$= \mathbb{E}\{\mathbb{V}\text{ar}[Y|X] + \mathbb{E}[Y|X]^2\} - \mathbb{E}\{\mathbb{E}[Y|X]\}^2$$
$$= \mathbb{E}\{\mathbb{V}\text{ar}[Y|X]\} + \mathbb{E}\{\mathbb{E}[Y|X]^2\} - \mathbb{E}\{\mathbb{E}[Y|X]\}^2\}$$
$$\mathbb{E}\{\mathbb{V}\text{ar}[Y|X]\} + \mathbb{V}\text{ar}\{\mathbb{E}[Y|X]\}$$

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

$$\mathbb{E}(N) = \nu, \mathbb{V}\text{ar}(N) = \tau^2, \mathbb{E}(\xi_i) = \mu, \mathbb{V}\text{ar}(\xi_i) = \sigma^2$$

$$\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}(\mathbb{E}(X|N)) = \mathbb{E}(N\mu) = \nu\mu \\
\mathbb{V}\text{ar}(X) &= \mathbb{E}(\mathbb{V}\text{ar}(X|N)) + \mathbb{V}\text{ar}(\mathbb{E}(X|N)) \\
&= \mathbb{E}(N\sigma^2) + \mathbb{V}\text{ar}(N\mu) = \nu\sigma^2 + \tau^2\mu^2
\end{aligned}$$

Branching processes

$$X_{n+1} = \xi_1 + \xi_2 + \cdots + \xi_{X_n}$$

where ξ_i are independent random variables with common propability mass function

$$\mathbb{P}\{\xi_i = k\} = p_k$$

From a random sum interpretation we get

$$\begin{aligned}
\mathbb{E}(X_{n+1}) &= \mu\mathbb{E}(X_n) = \mu^{n+1} \\
\mathbb{V}\text{ar}(X_{n+1}) &= \sigma^2\mathbb{E}(X_n) + \mu\mathbb{V}\text{ar}(X_n) = \sigma^2\mu^n + \mu^2\mathbb{V}\text{ar}(X_n) \\
&= \sigma^2\mu^n + \mu^2(\sigma^2\mu^{n-1} + \mu^2\mathbb{V}\text{ar}(X_{n-1}))
\end{aligned}$$

Extinction probabilities

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

$$u_n = \mathbb{P}\{N \leq n\} = \mathbb{P}\{X_N = 0\}$$

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

$$\phi(s) = \mathbb{E}\left(s^{\xi}\right) = \sum_{k=0}^{\infty} p_k s^k, \qquad p_k = \frac{1}{k!}\frac{d^k\phi(s)}{ds^k}\bigg|_{s=0}$$

Moments from generating functions

$$\frac{d\phi(s)}{ds}\bigg|_{s=1} = \sum_{k=1}^{\infty} p_k k s^{k-1}\bigg|_{s=1} = \mathbb{E}(\xi)$$

Similarly

$$\frac{d^2\phi(s)}{ds^2}\bigg|_{s=1} = \sum_{k=2}^{\infty} p_k k(k-1)s^{k-2}\bigg|_{s=1} = \mathbb{E}(\xi(\xi-1))$$

a factorial moment

$$\mathbb{V}\text{ar}(\xi) = \phi''(1) + \phi'(1) - (\phi'(1))^2$$

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 $p_x = P\{X_i = x\}$, $X_i \geq 0$ we find for $n = 2$ $S_2 = X_1 + X_2$
The event $\{S_2 = x\}$ can be decomposed into the set
$\{(X_1 = 0, X_2 = x), (X_1 = 1, X_2 = x - 1)$
$\ldots (X_1 = i, X_2 = x - i), \ldots (X_1 = x, X_2 = 0)\}$
The probability of the event $\{S_2 = x\}$ is the sum of the probabilities of the individual outcomes.

Sum of iid random variables - continued

The Probability of outcome $(X_1 = i, X_2 = x - i)$ is $P\{X_1 = i, X_2 = x - i\} = P\{X_1 = i\}P\{X_2 = x - i\}$ 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
$\square$

A probabilistic proof (which I think is instructive)

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

Sum of two Poisson distributed random variables

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

$$\phi_X(s) = e^{-\lambda(1-s)} \quad \phi_Y(s) = e^{-\mu(1-s)} \quad \left(\mathbb{P}\{X = x\} = p_x = \frac{\lambda^x}{x!} e^{-\lambda}\right)$$

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

$$X \sim B(n,p) \qquad Y \sim B(m,p) \qquad Z = X + Y$$

$$\phi_X(s) = (1-p+ps)^n \qquad \left(\mathbb{P}\{X = x\} = p_x = \binom{n}{x} p^x (1-p)^{n-x}\right)$$
$$\phi_Y(s) = (1-p+ps)^m$$

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

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

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

And we find

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

Generating function - the geometric distribution

$$p_x = (1-p)^{x-1}p$$

$$\phi_{geo}(s) = \sum_{x=1}^{\infty} s^x p_x = \sum_{x=1}^{\infty} s^x (1-p)^{x-1} p$$
$$= \sum_{x=1}^{\infty} s(s(1-p))^{x-1} p$$

A useful power series is:

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

And we get $\phi_{geo}(s) = \dfrac{sp}{1-s(1-p)}$

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 with pdf p_x $\qquad$ Y with pdf q_y

$Z = X + Y$ 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

$X \sim geo(p) \qquad Y \sim geo(p) \qquad Z = X + Y$

$\phi_X(s) = \frac{sp}{1-s(1-p)}$

$\phi_Y(s) = \frac{sp}{1-s(1-p)}$ $\qquad \left(P\{X = x\} = p_x = (1-p)^{x-1}p\right)$

And we get

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

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

$$p_x = \binom{x-1}{k-1}(1-p)^{x-k}p^k \qquad \phi_X(s) = \left(\frac{sp}{1-s(1-p)}\right)^k$$

Characteristic function and other

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

EXAMPLE: (exponential)

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