Multi-type branching process approximation to CTMC

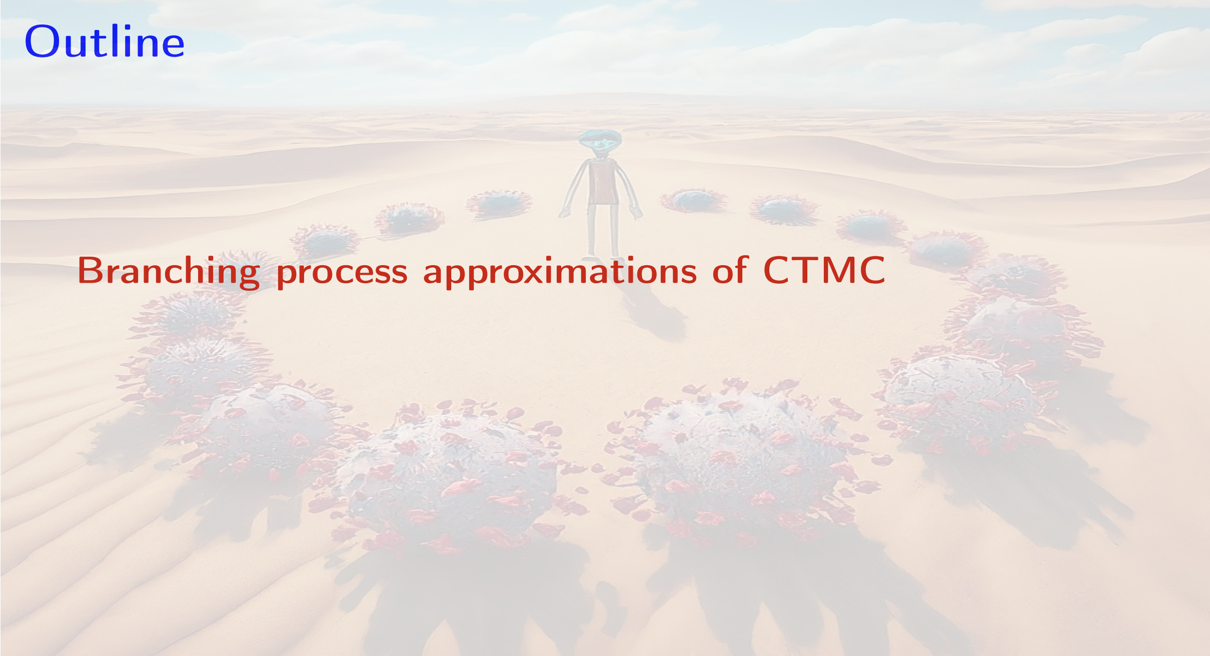
MATH 8xyz – Lecture 24

Julien Arino
Department of Mathematics @ University of Manitoba
Maud Menten Institute @ PIMS
julien.arino@umanitoba.ca

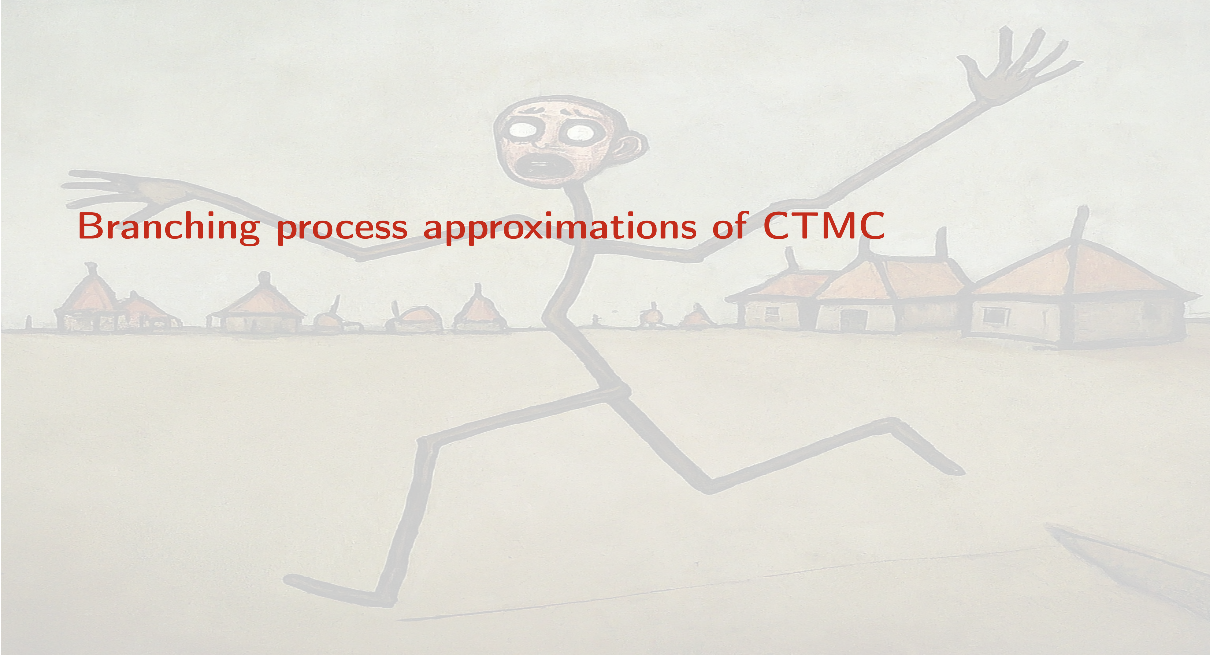
Winter 20XX

The University of Manitoba campuses are located on original lands of Anishinaabeg, Ininew, Anisininew, Dakota and Dene peoples, and on the National Homeland of the Red River Métis.
We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of Reconciliation and collaboration.

Outline



Branching process approximations of CTMC



Branching process approximations of CTMC

What is a Branching Process?

The Core Idea

A branching process is a mathematical model for a population where individuals produce a random number of offspring and then die.

- Think of bacteria splitting, a virus spreading, or even the survival of family surnames.
- We start with an initial population, Z_0.
- Each individual in generation n produces a number of offspring for generation $n+1$.
- This "number of offspring" is a random variable. All individuals produce offspring according to the same probability distribution, independently of each other.

The Galton-Watson Process

Let Z_n be the size of the population in generation n. We typically start with $Z_0 = 1$. The population evolves according to the rule:

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i}$$

- The term $X_{n,i}$ represents the number of offspring produced by the i-th individual in generation n.
- The variables $\{X_{n,i}\}$ are assumed to be **independent and identically distributed (i.i.d.)** integer-valued random variables.
- We call their common distribution $\{p_k\}_{k=0}^{\infty}$ the **offspring distribution**, where $p_k = P(X = k)$.

The Fundamental Questions

1. What is the long-term expected size of the population?
2. What is the probability that the population eventually dies out?

Mean Offspring

The fate of the population hinges on a single parameter: the mean of the offspring distribution

$$\mu = E[X] = \sum_{k=0}^{\infty} k \cdot p_k$$

Expected Population Size

Using the law of total expectation, we find the expected size of the next generation:

$$E[Z_{n+1} | Z_n] = E\left[\sum_{i=1}^{Z_n} X_{n,i} \middle| Z_n \right] = Z_n E[X] = Z_n \mu$$

Taking the expectation again, we get a simple recurrence:

$$E[Z_{n+1}] = \mu E[Z_n]$$

This implies:

The Three Regimes of Population Growth

The behavior of $E[Z_n] = Z_0\mu^n$ suggests three distinct cases:

Subcritical ($\mu < 1$)

$E[Z_n] \to 0$. The population is expected to shrink. It goes extinct with probability 1.

Critical ($\mu = 1$)

$E[Z_n] = Z_0$. The population is expected to remain stable. Surprisingly, it still goes extinct with probability 1.

Supercritical ($\mu > 1$)

$E[Z_n] \to \infty$. The population is expected to grow exponentially. It has a non-zero probability of surviving forever.

Tool: The Probability Generating Function

To find the extinction probability, we need a powerful tool: the **probability generating function (PGF)** of the offspring distribution X.

$$G(s) = E[s^X] = \sum_{k=0}^{\infty} p_k s^k \quad \text{for } |s| \leq 1$$

Key Properties

- $G(1) = \sum p_k = 1$
- The mean can be found from the derivative: $G'(1) = \sum k p_k = \mu$.
- The PGF of Z_n is the n-th iterate of $G(s)$ with itself. If $G_n(s)$ is the PGF of Z_n, then $G_{n+1}(s) = G(G_n(s))$.

The Extinction Probability Equation

Let π_0 be the probability of eventual extinction, starting with $Z_0 = 1$.

$$\pi_0 = P(\text{population dies out}) = \lim_{n\to\infty} P(Z_n = 0)$$

Since $P(Z_n = 0) = G_n(0)$, and $G_{n+1}(0) = G(G_n(0))$, in the limit the extinction probability π_0 must satisfy the equation:

$$\pi_0 = G(\pi_0)$$

Theorem 1

The extinction probability π_0 is the ***smallest non-negative solution*** *to the equation $s = G(s)$.*

- If $\mu \leq 1$, the only solution in $[0, 1]$ is $s = 1$. So $\pi_0 = 1$.
- If $\mu > 1$, there is a unique solution in $[0, 1)$, which is the extinction probability $\pi_0 < 1$.

From Discrete to Continuous Time

Limitation of Galton-Watson

Generations don't happen in synchronized steps in the real world. Individuals give birth and die at random times.

This leads us to **Continuous-Time Markov Chains (CTMCs)**.

- The state of the system is the population size, $k \in \{0, 1, 2, \dots\}$.
- Instead of generations, we have transition rates:
  - λ_k: rate of birth when population is size k (moves to $k+1$).
  - δ_k: rate of death when population is size k (moves to $k-1$).
- Often, we assume these rates are linear: $\lambda_k = k\lambda$ and $\delta_k = k\delta$. This means individuals act independently.

Branching Process Approximation of a CTMC

The Key Insight

At the beginning of an outbreak (or for a very large population), the dynamics caused by a single individual are largely independent of others.

This allows us to approximate the start of a CTMC population process with a branching process.

Example: A Simple Epidemic (SIR Model)

- S: Susceptible, I: Infected, R: Recovered.
- An infected person meets others at a certain rate. If they meet a susceptible, a new infection may occur (an "offspring").
- The infected person recovers (or dies) at another rate, ending their infectious period.
- **Question:** How many new infections does a single infected person cause on average?

Case Study: The Basic Reproduction Number $\mathcal{R}_0$

Consider a single infected individual in a large population of susceptibles.

- Let β be the infection rate (rate of producing "offspring").
- Let γ be the recovery rate (rate of "dying").

The individual's infectious lifetime is an exponential random variable with mean $1/\gamma$. The average number of secondary infections they cause is:

$$\mathcal{R}_0 = (\text{rate of infection}) \times (\text{average infectious period}) = \beta \times \frac{1}{\gamma} = \frac{\beta}{\gamma}$$

The Connection

$\mathcal{R}_0$ is precisely the **mean offspring number** μ for the embedded branching process that approximates the start of the epidemic.

Applying Branching Theory to Epidemics

The fate of the epidemic's initial phase is determined by $\mathcal{R}_0$:

- If $\mathcal{R}_0 \leq 1$ ($\mu \leq 1$): The number of infected individuals is a subcritical or critical process. The epidemic will die out with probability 1.
- If $\mathcal{R}_0 > 1$ ($\mu > 1$): The process is supercritical. There is a positive probability that the epidemic takes off and causes a major outbreak.

We can even calculate the probability of a major outbreak! It is $1 - \pi_0$, where π_0 is the extinction probability.
For this simple birth-death infection process, the PGF is $G(s) = \frac{\gamma}{\beta+\gamma} + \frac{\beta}{\beta+\gamma} s$. Solving $s = G(s)$ gives the extinction probability:

$$\pi_0 = \frac{\gamma}{\beta} = \frac{1}{\mathcal{R}_0}$$

The probability of a major outbreak is $1 - 1/\mathcal{R}_0$.

Summary

- **Branching Processes** model populations with i.i.d. offspring generation.
- The fate of the population is determined by the **mean offspring number** μ. Extinction is certain if $\mu \leq 1$.
- The **extinction probability** π_0 can be calculated as the smallest non-negative fixed point of the probability generating function $G(s)$.
- The initial stages of many large-scale **Continuous-Time Markov Chains** can be approximated by a branching process.
- This allows us to apply the theory to real-world problems, like calculating an epidemic's **basic reproduction number** $\mathcal{R}_0$ and its probability of causing a major outbreak.

Bibliography I