

The Most Recent Common Ancestor of a Randomly Chosen Sample in a Galton-Watson Process

A thesis submitted in partial fulfilment of the
requirements for the degree of Master of
Mathematical Sciences (Advanced)

by

Albert Christian Soewongsono

Supervisor

A/Prof. Conrad Burden

Mathematical Sciences Institute
ANU College of Science
The Australian National University
October 2018



Declaration

This thesis is an account of research conducted between February 2018 and and October 2018 at the Mathematical Science Institute, The Australian National University, Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Albert Christian Soewongsono
25 October 2018

Acknowledgements

I would first like to thank my thesis supervisor A/Prof. Conrad Burden to guide me towards the completion of my thesis. His guidance throughout the year has helped me to build a good understanding of the subject. His office door was always open whenever I ran into a trouble spot or had questions about my research. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it.

I would also like to thank Prof. Steve Roberts and Prof. Ross Maller for being the examiners for this thesis and validating the results. Last but not least, I would also to express my thank you to the Lord and my family for supporting me throughout my studies at the Australian National University.

Abstract

One of the interesting topics to be examined in the field of theoretical population genetics is to find the most recent common ancestor of the current population. In this thesis, firstly we proceed by finding the common ancestor of two randomly chosen individuals at the current time s , thus we will build a mathematical model to illustrate this situation using a Galton-Watson process which allows the population to grow stochastically over time, and only genetic drift affects the model. We then generalise the model into modelling the common ancestor for n randomly chosen individuals at the current time step. Later, we derive the joint estimate of having a common ancestor from n chosen individuals and the current population at time s does not exceed the observed population. Hence, this will be used to determine the estimate of the initial population size and the time of the most recent common ancestor for the case of $n = 2$. Lastly, we will compare the coalescent time with result found by Slatkin and Hudson (1991) using the Wright-Fisher model.

Keywords: *Population Genetics, Most Recent Common Ancestor, Galton-Watson Process, Wright-Fisher Model, Coalescent Time*

List of Notations

- m_0 = The initial population size.
- κ_0 = The scaled initial population size.
- t = The continuous time step.
- s = The scaled continuous time step.
- τ = The discrete time step.
- $\mathcal{E}_2$ = Event that two randomly chosen individuals at the current time are descended from a common ancestor at time $s = 0$.
- $\mathcal{E}_n$ = Event that n randomly chosen individuals from a fraction x of the current population at time s descended from a common ancestor at time $s = 0$.
- κ = The observed scaled current population size.
- $Z(s)$ = The population size at time s .
- $\kappa(s)$ = The scaled population size at time s .
- κ_1 = Sub-population of κ which is direct descendants from one individual at time 0.
- κ_2 = Sub-population of κ which is descendants from $m_0 - 1$ individuals at time 0.
- $\log \lambda$ = Per generation growth rate.
- $\Pi_0(s, \kappa_0, x_0)$ = The extinction probability of the fraction x_0 of the initial population before time s .
- $u(s, \kappa_0)$ = Probability of having single ancestor from a sample of size n from the current population at time s .
- $v(s, \kappa_0)$ = Probability that the population at time s will not exceed the observed current population given the starting population κ_0 .
- $L(s, \kappa_0 | \kappa)$ = The likelihood function that the population at time s will not exceed the observed population κ for a given starting population κ_0 .
- $J_{U,V}(s, \kappa_0)$ = The Jacobian matrix of $u(s, \kappa_0)$ and $v(s, \kappa_0)$.

Contents

Declaration	i
Acknowledgements	ii
Abstract	iii
Table of Contents	vi
1 Population Genetics: A Quick Look	1
2 The Wright-Fisher Model: A Brief Introduction	3
2.1 Allele Fixation Probability in Wright-Fisher Model with No Mutations	5
3 Galton-Watson Process in Population Genetics	7
3.1 A Brief Historical Review	7
3.2 The Derivation of The Model	8
3.3 The Continuum Galton-Watson Process Via the Forward Kolmogorov Equation	10
4 Earlier Results in the Application to Mitochondrial Eve	13
4.1 The Time of mtE with Galton-Watson Branching Process Imposed	14
4.2 The Coalescent Time of Two Individuals with Wright-Fisher Model Imposed	20
5 Novel Results and Discussion	23
5.1 The Most Recent Common Ancestor of a Sample of Two Individuals in a Galton-Watson Process	23
5.2 The Most Recent Common Ancestor of a Sample of n Individuals in a Galton-Watson Process	31

5.3	The Mean Coalescent Time of Two Randomly Chosen Individuals with a Galton-Watson Process Imposed	37
6	Conclusion	46
7	Appendix	48
7.1	The Dirac-Delta Function and its Properties	48
7.2	The Modified Bessel Function of the First Kind	49
7.3	Derivation of Forward-Kolmogorov Equation	50
7.4	The Solution of Forward-Kolmogorov Equation Via Bartlett's Equa- tion	54
7.5	R Codes for this Study	58

Chapter 1

Population Genetics: A Quick Look

First introduced in 1920's and 1930's by Fisher, Haldane, and Wright, Population Genetics occupies a major role in understanding of evolutionary processes of plants and animals important to mankind (Ewens (2004)). As the main concerns of this subject are the genetic differences within populations and the dynamics of how populations evolve throughout time as a result of genetic mutations happening in the germlines¹ of individuals (Gillespie (2004)), the study itself cannot be separated of what is known as "Genetic Drift". The genetic drift occurs when a single individual is the ancestor of a significant portion of the population in a subsequent population (See fig.4.1). Therefore, to understand the importance of genetic drift in determining the make up of a current population, it is useful to understand the statistical properties of the "Most Recent Common Ancestor" (MRCA) of a random sample of the population.

More specifically, one may want to estimate the time since the MRCA and the size of the population at that time. This is where mathematical modelling is required. As pioneers on this subject, the Wright-Fisher (W-F) model is the first to set foot on the study of the subject, and has ever since become the most commonly used model on researches in this area. The W-F model is popular because it is relatively easy to study the ancestry of a sample of individuals, since it naturally 'looks backwards' in time i.e. each individual randomly chooses their parent.

However, this model has disadvantages of not being a good physical descrip-

¹refers to the lineage of cells which, as cells segregate, might pass to the next generations.

tion of how populations propagate and it does not allow for a stochastically variable population size. On the other hand, the Galton-Watson (G-W) model is a closer description of reality since each individual randomly chooses the number of offspring. For a G-W model, populations evolve according to a common probability distribution for the number of offspring of each parent on the assumption that this distribution is identical but independent for the whole population and throughout generations (O'Connell (1995); Cyran and Kimmel (2010)). In a later chapter, we will discuss this model more thoroughly and its features, including its diffusion limit via Forward-Kolmogorov equation.

Unlike the W-F model that 'looks backward' in time, this model 'looks forward' in time, which makes studying the coalescence of the ancestry of a sample of the population difficult. Exploring the statistical properties of the MRCA of a random sample in a G-W process is the aim of this thesis.

Chapter 2

The Wright-Fisher Model: A Brief Introduction

Introduced by Sewall Wright and Ronald Fisher (Neither of whom published the statement for the model, but did publish series of publications regarding the model) in 1931, the W-F model is the earliest model to deeply study Population Genetics and has been long considered important as the starting point for many researches in the field (Ewens (2004)). Although being the earliest model in the field, its assumption for constant population size i.e. no stochastic influence on the size of populations makes it biologically unrealistic, as for example, it cannot capture a growing population such as human populations (Burden and Simon (2016)). Based on the Binomial distribution, we shall discuss the model more deeply in later section. Since the main focus of this paper is on genetic drift, we will focus mainly on models without mutation or selection, including the W-F Model.

The Wright-Fisher model is constructed under the following assumptions:

- **Non-Overlapping Generations:** Imagine an annual plant where in each season, new seeds replace the old ones. This assumes each individual in a population in a given generation reproduce within discrete time step τ and $\tau + 1$, where $\tau \in \{1, 2, 3, \dots\}$.
- **Fixed Population Size:** This assumes the population of size M remains constant throughout generations.
- **Diploid Population:** This assumes that each individual in a population has two copies of genome. Therefore, it is more convenient to think the

population size $M = 2N$ copies of genome, where N is the number of diploid individuals. An example of a diploid population is Clarkia plants.

- **Monoecious Population:** This assumes that each individual in a population carries both male and female reproductive organs. Hence, a mating can occur between two individuals or within an individual. In fact, many plants are monoecious.
- **Random Mating:** This assumes that each individual in generation $\tau + 1$ is offspring of two randomly chosen and independent individuals in generation τ . Since the population is monoecious, both parents might be the same individual, therefore the probability of having one parent in previous generation is $\frac{1}{M}$.

To start the construction of the model, given an observed SNP¹, we shall examine the model in which two distinct alleles² $A_1, A_2 \in \{A, C, G, T\}$ ³ are present. Next, we define the random variable $Y(\tau)$ as the number of copies of allele A_1 within the population at the time step or generation τ . It follows that its range is given by $\Omega_{Y(\tau)} = \{0, 1, \dots, M\}$. Since the population is constant over generations and random mating is allowed, the number of copies of allele i in the population at the generation $\tau + 1$ is dependent only on the previous generation. Thus, it follows what is known as “Markov Processes”. Hence, we can define the transition probability for number of A_1 allele copies of the population going from generation τ to generation $\tau + 1$ as follows:

$$p_{ij} = \text{Prob}(Y(\tau + 1) = j | Y(\tau) = i), \quad (2.1)$$

where $Y(\tau + 1) = j | Y(\tau) = i \sim \text{Bin}(M, \frac{i}{M})$. Following the assumptions, the probability that any j individuals with allele type A_1 at generation $\tau + 1$ being the offspring from i parents in the previous generation is $\frac{i}{M}$. Consequently, the probability that the rest of current population $M - Y_\alpha\{\tau\}$ inherit the other allele type is $1 - \frac{i}{M}$. Thus,

$$p_{ij} = \binom{M}{j} \left(\frac{i}{M}\right)^j \left(1 - \frac{i}{M}\right)^{M-j}. \quad (2.2)$$

¹Single Nucleotide Polymorphisms or SNPs are the most common type of genetic variation.

²Common term used in Population Genetics to describe different variants at a genetic locus.

³Acronyms for 4 different types of Nucleobases: A =Adenine, C =Cytosine, G =Guanine, T =Thymine.

It follows immediately from Eq. (2.2) that;

$$p_{0j} = \begin{cases} 0 & \text{if } j \neq 0 \\ 1 & \text{if } j = 0 \end{cases}, \quad p_{Mj} = \begin{cases} 0 & \text{if } j \neq M, \\ 1 & \text{if } j = M. \end{cases} \quad (2.3)$$

This implies that 0 and M are the absorbing states of the model. In other words, these states correspond to situation where the population has either allele A_1 or allele A_2 fixed. Hence, this gives rise to a question regarding the fixation probabilities of both alleles in a long run.

2.1 Allele Fixation Probability in Wright-Fisher Model with No Mutations

To begin with, let's assume there are i number of copies of allele A_1 within the population at time $\tau = 0$ i.e. $Y(0) = i$, where $i \in \{0, 1, \dots, M\}$. Then, first we would like to seek the probability of fixation of allele A_1 which can be interpreted as;

$$\text{Prob}(\text{Allele } A_1 \text{ is fixed} | Y_0 = i) = \text{Prob}(Y(\infty) = M | Y_0 = i). \quad (2.4)$$

Since this whole thing is a Markov Process, we can impose “the law of total probability” onto Eq. (2.4) to break it into sum of different cases as follows;

$$\begin{aligned} & \text{Prob}(Y(\infty) = M | Y(0) = i) \\ &= \sum_{j=1}^M \text{Prob}(Y(\infty) = M | Y(1) = j) \text{Prob}(Y(1) = j | Y(0) = i). \end{aligned} \quad (2.5)$$

From Eq. (2.3), we know that 0 and M are the absorbing states. It follows immediately that the boundary conditions of the model are as follows;

$$\text{Prob}(Y(\infty) = M | Y(0) = 0) = 0, \quad \text{Prob}(Y(\infty) = M | Y(0) = M) = 1. \quad (2.6)$$

By defining $\Psi_i = \text{Prob}(Y(\infty) = M | Y(\tau) = i)$, we can rewrite Eq. (2.5) and Eq. (2.6) as follows;

$$\Psi_i = \sum_{j=1}^M p_{ij} \Psi_j, \quad \Psi_0 = 0, \Psi_M = 1. \quad (2.7)$$

It is easy to check that the solution to Eq. (2.7) is;

$$\Psi_i = \frac{i}{M}. \quad (2.8)$$

Similarly, by replacing the boundary conditions to $\Psi_0 = 1$ and $\Psi_M = 0$, the fixation probability for allele A_2 is $1 - \frac{i}{M}$.

Chapter 3

Galton-Watson Process in Population Genetics

We will discuss briefly the G-W process, and subsequently build a story that will lead us to the main discussion of this thesis, which is finding the MRCA of a sample of a growing population. Since no mutations are involved in this thesis, we will proceed this discussion for the no-mutation model of G-W process.

3.1 A Brief Historical Review

Named after British mathematician and statistician Francis Galton and fellow British mathematician Henry William Watson, this term refers to Francis Galton's statistical investigation on the extinction of family names, particularly concerning the extinction of the Victorians' aristocratic surnames. This problem was posed by Galton through the 1873 series of *The Educational Times* regarding the probability of such an event. At the same year, just a few months later, Watson replied to the problem issued by Galton also through *The Educational Times* with a solution, in which Galton described as "totally erroneous". A year later, they published a joint paper (Galton and Watson, 1875) with the correct solution.

The main foundation of the model the two developed on their paper is by assuming that surnames would only be passed onto male children by their father, then by supposing the number of a man's sons as a discrete random variable say, Y_α where $Y_\alpha \in \{0, 1, 2, \dots\}$ they again assumed that different men's sons are all random variables with same distribution where each one of them is independent to each other. Thus, the number of sons population from all different male parents

with a particular surname at the next generation $\tau + 1$, $M(\tau + 1)$, can be defined as

$$M(\tau + 1) = \sum_{\alpha=1}^{M(\tau)} Y_{\alpha}.$$

From here we can see that this forms a stochastic process which only depends on exactly one previous generation. The conclusion drawn from this is that if the expected number of a man's sons is less than one, i.e $E(Y_{\alpha}) \leq 1$ then their surname will go extinct "almost surely"¹, while if its expectation is more than one, then there is probability that the inherited surname will survive in any given number of generations (Galton and Watson, 1875). Their work was actually independent of an earlier similar result found but unpublished by French Statistician I.J. Bienaymé.

The branching model has many applications in various research fields, such as understanding human Y-chromosome DNA haplogroups² (Kendall, 1966), understanding Nuclear Chain Reaction (Byers, 2002), and in the field of Population Genetics, was used by Burden and Simon (2016) to study Mitochondrial Eve³. The latter will also be discussed on this paper in later section as a building block to lead us to the main discussion of this paper.

3.2 The Derivation of The Model

We shall start by stating all the assumptions used for developing G-W Branching Process. Those assumptions are as follows:

- The population size is not fixed, but rather varying stochastically.
- Consider a population of $M(\tau)$ haploid individuals at generation τ in which every individual in the population reproduce in discrete, non-overlapping, independent generations $\tau = 1, 2, 3, \dots$.
- The population is grouped into n different allele types such that the number

¹It is terminology in Probability Theory to describe that an event which occurs with probability 1, even though the complementary event is, in principle, possible.

²It is a group of genes inherited from a single parent in which mutations occurred in the non-recombining portions of the DNA from the Y-chromosome.

³See also Mode et al. (2013).

of sub-population with allele type η is $Y_\eta(\tau)$. Thus,

$$M(\tau) = \sum_{i=1}^n Y_i(\tau). \quad (3.1)$$

- No mutations between allele types nor selection occur.
- The number of offspring per individuals with allele type η in any generation is a set of *independent and identically distributed* (i.i.d.) random variables $S_i^{(\eta)}$, $i = 1, \dots, Y_\eta(\tau)$. Since they are identical, we will denote from now on a non-negative integer valued random variable S having mean and variance denoted by;

$$E(S) = \lambda, \quad \text{Var}(S) = \sigma^2. \quad (3.2)$$

Since λ and σ^2 are common to all allele types, the process can be thought of as modelling a population subject to genetic drift but not mutation or selection. It is clear that;

$$Y_\eta(\tau + 1) = \sum_{i=1}^{Y_\eta(\tau)} S_i^{(\eta)}. \quad (3.3)$$

Hence, using the formula for the mean of the sum of a random number of i.i.d. random variables,

$$\begin{aligned} E(Y_\eta(\tau + 1)) &= E\left(\sum_{i=1}^{Y_\eta(\tau)} S_i^{(\eta)}\right) \\ &= E(Y_\eta(\tau)) E(S_i^{(\eta)}) \\ &= E(Y_\eta(\tau)) E(S) \\ &= \lambda E(Y_\eta(\tau)). \end{aligned} \quad (3.4)$$

Suppose we have initial conditions given as follows;

$$Y_\eta(0) = y_\eta, \quad M(0) = m_0 = \sum_{i=1}^n Y_i(0). \quad (3.5)$$

It, then, follows immediately from Eq. (3.1) and Eq. (3.4) that;

$$E(Y_\eta(\tau)) = y_\eta \lambda^\tau, \quad \text{Var}(Y_\eta(\tau + 1)|Y_\eta(\tau)) = Y_\eta(\tau) \sigma^2, \quad E(M(\tau)) = m_0 \lambda^\tau. \quad (3.6)$$

3.3 The Continuum Galton-Watson Process Via the Forward Kolmogorov Equation

We will examine the diffusion limit of the G-W branching model using the Forward-Kolmogorov equation ⁴.

The main reason we need to look into its continuum model is because, naturally, the population size is generally large, especially human population. This continuum approach of the G-W model has been discussed in detail by Feller (1951) and is summarised by Bailey (1964) and Cox and Miller (1978). We will specify only for the case of 2 alleles, therefore we will start by stating that $Y(\tau)$ to be the number of copies or sub-population of allele type A_1 . These following changes are made to accommodate the model:

$$t = \frac{\tau}{m_0} \quad X(t) = \frac{1}{M} Y(\tau) \quad x_0 = \frac{y_0}{m_0} = p \text{ "Scaled initial population"} \quad (3.7)$$

$$\delta t = \frac{1}{M} \quad \delta X(t) = \frac{1}{M} [Y(\tau + 1) - Y(\tau)]. \quad (3.8)$$

In Appendix 7.3 we assume the existence of the functions $a(x)$ and $b(x)$ defined by Eq. (7.8), thus we will find what $a(x)$ and $b(x)$ correspond to in the Galton-Watson process. As noted in Burden and Simon (2016), the continuum limit is obtained under assumptions $m_0 \rightarrow \infty$ and $\lambda \rightarrow 1$ such that the growth rate $\alpha = m_0 \log \lambda$ remains constant.

⁴Also known as Fokker-Planck equation in Physics, it is a partial differential equation that describes the change in time of a probability density function under an influence of random forces, in this case, it would be the genetic drift in the population.

Then,

$$\begin{aligned}
& E(\delta X(t)|X(t) = x) \\
&= \frac{1}{m_0} E(Y(\tau + 1) - Y(\tau)|Y(\tau) = y) \\
&= \frac{1}{m_0} [E(Y(\tau + 1)|Y(\tau) = y) - E(Y(\tau)|Y(\tau) = y)] \\
&= \frac{1}{m_0} [\lambda E(Y(\tau)|Y(\tau) = y) - y] \\
&= \frac{1}{m_0} (\lambda y - y) \\
&= \frac{1}{m_0} (\lambda - 1)y \\
&= \frac{1}{m_0} (\lambda - 1)xm_0 \\
&= \alpha x \delta t + O(\delta t^2).
\end{aligned} \tag{3.9}$$

Therefore,

$$a(x) = \alpha x. \tag{3.10}$$

Then, to determine $b(x)$, let's examine the following;

$$\begin{aligned}
& E(\delta X(t)^2|X(t) = x) \\
&= \text{Var}(\delta X(t)|X(t) = x) + E(\delta X(t)|X(t) = x)^2 \\
&= \text{Var}(\delta X(t)|X(t) = x) + (\alpha x \delta t + O(\delta t))^2 \\
&= \text{Var}(\delta X(t)|X(t) = x) + O(\delta t^2) \\
&= \frac{1}{m_0^2} \text{Var}(Y(\tau + 1) - Y(\tau)|Y(\tau) = y) + O(\delta t^2) \\
&= \frac{1}{m_0^2} [\text{Var}(Y(\tau + 1) - y|Y(\tau) = y)] \\
&= \frac{1}{m_0^2} \text{Var}(Y(\tau + 1)|Y(\tau) = y) + O(\delta t^2) \\
&= \frac{1}{m_0^2} y \sigma^2 + O(\delta t^2) \\
&= \frac{1}{m_0} x \sigma^2 + O(\delta t^2) \\
&= \sigma^2 x \delta t + O(\delta t^2).
\end{aligned} \tag{3.11}$$

Therefore,

$$b(x) = \sigma^2 x. \tag{3.12}$$

Substitute these back to Eq. (7.18) in Appendix 7.3, we get the corresponding Forward-Kolmogorov equation for the Galton-Watson process as follows:

$$\frac{\partial f}{\partial t} = -\alpha \frac{\partial}{\partial x} (xf(x, t)) + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} (xf(x, t)). \quad (3.13)$$

The solution of the above equation for $\alpha \neq 0$ and the initial condition $f(x, 0) = \delta(x - x_0)$, as shown in Bailey (1964) is the following; (See Appendix 7.4 for a derivation)

$$f(x, t) = \frac{2\alpha}{\sigma^2 (e^{\alpha t} - 1)} \left(\frac{x_0 e^{\alpha t}}{x} \right)^{\frac{1}{2}} \exp \left[\frac{-2\alpha (x_0 e^{\alpha t} + x)}{\sigma^2 (e^{\alpha t} - 1)} \right] I_1 \left(\frac{4\alpha (x_0 x e^{\alpha t})^{\frac{1}{2}}}{\sigma^2 (e^{\alpha t} - 1)} \right) + \delta(x) p_0(t), \quad (3.14)$$

where I_1 is the *Modified Bessel Function* (See Appendix 7.2) of the first kind, $\delta(x)$ is the *Dirac Delta function* (See Appendix 7.1), and the extinction probability given by (Bailey, 1964),

$$p_0(t) = \exp \left\{ \frac{-2\alpha x_0 e^{\alpha t}}{\sigma^2 (e^{\alpha t} - 1)} \right\}. \quad (3.15)$$

Note that the term involving the delta function in Eq. (3.14) corresponds to a point mass at zero which indicates that the given allele type becomes extinct at or before time t (Burden and Simon, 2016). Furthermore, the limiting cases for the extinction probability are given by;

$$\lim_{t \rightarrow \infty} p_0(t) = \begin{cases} 1 & \text{if } \alpha \leq 0 \\ e^{\frac{-2\alpha x_0}{\sigma^2}} & \text{if } \alpha > 0 \end{cases}. \quad (3.16)$$

This implies that for sub-critical and critical growth rate α , the given allele type becomes extinct almost surely, while for the super-critical growth rate case, there is a possibility that the given allele type will survive.

Chapter 4

Earlier Results in the Application to Mitochondrial Eve

Since centuries ago, many researches have been undertaken among scientists to determine the origin of humankind. One of the most discussed one is to find the *mitochondrial Eve*¹ or the most recent common ancestor (MRCA) of all currently living humans following the maternal line (See Fig.4.1).

The male counterparts to describe the most recent common ancestor of all currently living human being following the paternal line is known as *Y-chromosomal Adam* i.e Y-MRCA. It is worth noting that both Y-MRCA and mtE did not live at the same time period (Francalacci et al., 2013), although they both were not far apart, as suggested by Cann (2013) which is roughly around 120,000 and 156,000 years ago.

There have been extensive researches carried out for the study of finding the Mitochondrial eve, some of them include O’Connell (1995) which estimated the age of mtE from a near-critical branching process for the case of supercritical growth (later being compared by Cyran and Kimmel (2010)), and research with an aim to estimate the population size at the time of mtE such as the one conducted by Zimmerman (2001). On this chapter, we will also give the brief review to the research conducted by Burden and Simon (2016) which examined the time of mtE in a non-mutating population with two allele types using the Galton-Watson branching process, and also the one from Slatkin and Hudson (1991) which examined the same problem but using the Wright-Fisher model.

¹We will refer it later simply as "mtE".

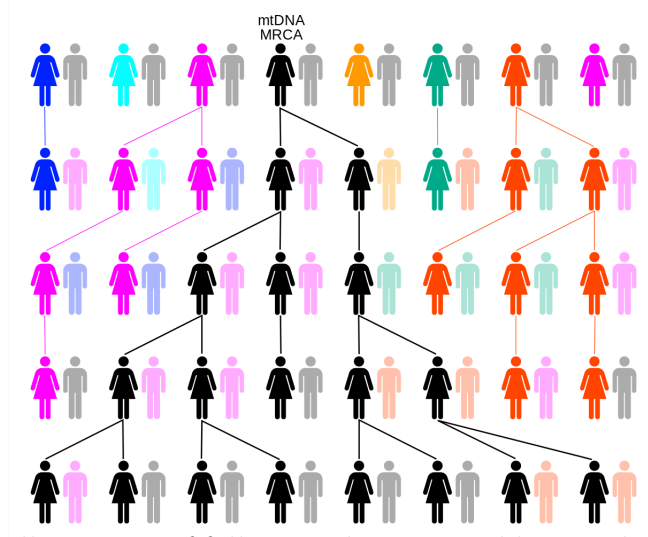


Figure 4.1: The illustration of following the maternal line to the time of mtE over five generations

Source:Wikipedia

4.1 The Time of mtE with Galton-Watson Branching Process Imposed

In this section, we will follow through the discussion of the result from Burden and Simon (2016). Recall from chapter 3 that we have different cases for α . The case $\alpha = 0$, is analogous to a Wright-Fisher model for which the population size is assumed constant. On the other hand, when imposing with the Galton-Watson branching process, we take the case for $\alpha > 0$, i.e. supercritical growth, since in this branching process the population size is assumed to grow stochastically over time. Then, let's make the following changes to the following variables as suggested in Burden and Simon (2016);

$$s = \alpha t = \tau \log \lambda, \quad Z(s) = X(t)e^{-\alpha t} = \frac{Y(\tau)}{m_0 \lambda^\tau}, \quad \kappa_0 = \frac{2\alpha}{\sigma^2} = \frac{2m_0 \log \lambda}{\sigma^2}. \quad (4.1)$$

Using the solution of the Forward-Kolmogorov equation derived in Appendix 7.4, one can easily see that we have the following density function with respect to Z ;

$$f_Z(z|s, \kappa_0, x_0) = \frac{\kappa_0}{1 - e^{-s}} \left(\frac{x_0}{z} \right)^{\frac{1}{2}} \exp \left[\frac{-\kappa_0(x_0 + z)}{1 - e^{-s}} \right] I_1 \left(\frac{2\kappa_0(x_0 z)^{\frac{1}{2}}}{1 - e^{-s}} \right) + \delta(z)\pi_0(s, \kappa_0, x_0), \quad (4.2)$$

where

$$\pi_0(s, \kappa_0, x_0) = \exp\left(\frac{-\kappa_0 x_0}{1 - e^{-s}}\right) \quad (4.3)$$

is the extinction probability of one of the alleles at time s . By the definition, the total scaled population size Z_{tot} is the sum of all sub-populations of each allele type given as follows;

$$Z_{\text{tot}}(s) = \sum_{i=1}^n Z_i(s) = \frac{M(\tau)}{m_0 \lambda^\tau}, \quad (4.4)$$

where Z_{tot} is also the Galton-Watson process by the independence of each $Z_i(s)$ and the assumption that each allele type has common mean and variance. Therefore, we can define the density function of Z_{tot} in terms of Eq. (4.2) as,

$$f_{Z_{\text{tot}}}(z|s, \kappa_0) = f_Z(z|s, \kappa_0, 1). \quad (4.5)$$

As argued in Burden and Simon (2016), the time of the mtE with the initial population size m_0 is likely around when $\kappa_0 \approx 1$. We will proceed the following heuristic arguments of why $\kappa_0 \gg 0$ and $\kappa_0 \ll 0$ were not the case.

- Given the initial population m_0 , and assume we divide the population into two distinct sub-populations. If κ_0 is very large, it is most likely from Eq. (4.3) that neither sub-population becomes extinct so, with high probability that both sub-populations will have descendants in present time. Such case implies that the initial population is ‘too large’ to be the population when the mtE lived. In other words, since the population is growing stochastically in time, we have not traced back in time far enough to reach the time of the mtE.
- Again by dividing the initial population into two distinct sub-populations, since κ_0 is very small, Eq. (4.3) tells us that extinction occurs with high probability, so it is very likely that only one of the sub-populations will have descendants that survive until the present time. However, since the time between generations is ”*infinitesimal*” under the continuum limit, that ancestor will most likely have a daughter who is also the common ancestor of the entire population at the present time, hence that ancestor cannot be the mtE (Burden and Simon, 2016). In other words, the initial population is ‘too small’ to support the mtE i.e we have gone too far back in time that we actually have passed the time of mtE.

To start with the quantitative analysis, we define a ‘scaled continuous population size’ by;

$$K(s) = \frac{2M(\tau)\log\lambda}{\sigma^2}. \quad (4.6)$$

Then Z_{tot} and $K(s)$ are related via;

$$Z_{tot}(s) = \frac{K(s)}{\kappa_0 e^s}. \quad (4.7)$$

Define the following event;

$$\begin{aligned} \epsilon(\kappa_0, s) = & \text{The event that the population at time } s = 0 \text{ contains an individual} \\ & \text{who is the common ancestor of the entire population at time } s. \end{aligned} \quad (4.8)$$

Consider the joint probability of the two events $\epsilon(\kappa_0, s)$ and $K(s) \in [\kappa, \kappa + d\kappa]$. Imagine we have an initial population size m_0 , in which $m_0 - 1$ lineages become extinct, while the one remaining lineage (this lineage represents a fraction $\frac{1}{m_0}$ of the initial population) survives and is the common ancestor of all of the population at time s . Since there are m_0 possibilities for the MRCA, we then have;

$$\begin{aligned} & \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \epsilon(\kappa_0, s) | K(0) = \kappa_0) \\ &= m_0 f_Z \left(\frac{\kappa}{\kappa_0 e^s} \middle| s, \kappa_0, \frac{1}{m_0} \right) d \left(\frac{\kappa}{\kappa_0 e^s} \right) \times \pi_0 \left(s, \kappa_0, 1 - \frac{1}{m_0} \right). \end{aligned} \quad (4.9)$$

Using Eq. (4.2) with $z > 0$, so that the δ -function does not contribute, we know that

$$m_0 f_Z \left(z \middle| s, \kappa_0, \frac{1}{m_0} \right) = m_0^{\frac{1}{2}} \frac{\kappa_0}{z^{\frac{1}{2}} (1 - e^{-s})} \exp \left[\frac{-\kappa_0 \left(z + \frac{1}{m_0} \right)}{1 - e^{-s}} \right] I_1 \left(\frac{2\kappa_0 \left(\frac{z}{m_0} \right)^{\frac{1}{2}}}{1 - e^{-s}} \right) \quad (4.10)$$

Since $\frac{\kappa_0 \left(\frac{z}{m_0} \right)^{\frac{1}{2}}}{1 - e^{-s}} \rightarrow 0$ as $m_0 \rightarrow \infty$, we have;

$$I_1 \left(\frac{2\kappa_0 \left(\frac{z}{m_0} \right)^{\frac{1}{2}}}{1 - e^{-s}} \right) = \frac{\kappa_0 \left(\frac{z}{m_0} \right)^{\frac{1}{2}}}{1 - e^{-s}} + O \left(\frac{1}{m_0} \right). \quad (4.11)$$

Substitute this back to Eq. (4.10) and ignore all terms involving $\frac{1}{m_0}$ (as they eventually tend to 0), we get;

$$m_0 f_Z \left(z \left| s, \kappa_0, \frac{1}{m_0} \right. \right) = \frac{\kappa_0^2}{(1 - e^{-s})^2} \exp \left(\frac{-\kappa_0 z}{1 - e^{-s}} \right). \quad (4.12)$$

Thus, substitute this back to Eq. (4.9) for $z = \frac{\kappa}{\kappa_0 e^s}$, and using Eq. (4.3) yields;

$$\text{Prob} (K(s) \in [\kappa, \kappa + d\kappa], \epsilon(\kappa_0, s) | K(0) = \kappa_0) = \frac{\kappa_0 e^{-s}}{(1 - e^{-s})^2} \exp \left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}} \right) d\kappa. \quad (4.13)$$

Again by using the solution to the Forward-Kolmogorov equation (See Eq. (4.2)), we can determine the probability density that the population size grows to size κ at time s given we start with a initial population of size κ_0 as the following;

$$\begin{aligned} \text{Prob} (K(s) \in [\kappa, \kappa + d\kappa] | K(0) = \kappa_0) &= f_Z \left(\frac{\kappa}{\kappa_0 e^s} \left| s, \kappa_0, 1 \right. \right) d \left(\frac{\kappa}{\kappa_0 e^s} \right) \\ &= \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{\kappa} \right)^{\frac{1}{2}} \exp \left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}} \right) I_1 \left(\frac{2(\kappa \kappa_0 e^s)^{\frac{1}{2}}}{e^s - 1} \right) d\kappa. \end{aligned} \quad (4.14)$$

Thus, using Eq. (4.13) and Eq. (4.14), the probability that the current population is descended from one individual at time 0 given that we start with a scaled population of size κ_0 and the scaled population grows to a size κ , can be calculated as follows;

$$\begin{aligned} &\text{Prob} (\epsilon(\kappa_0, s) | K(0) = \kappa_0, K(s) = \kappa) \\ &= \frac{\text{Prob} (K(s) \in [\kappa, \kappa + d\kappa], \epsilon(\kappa_0, s) | K(0) = \kappa_0)}{\text{Prob} (K(s) \in [\kappa, \kappa + d\kappa] | K(0) = \kappa_0)} = \frac{(\kappa \kappa_0 e^s)^{\frac{1}{2}} (e^s - 1)^{-1}}{I_1 \left(2(\kappa \kappa_0 e^s)^{\frac{1}{2}} (e^s - 1)^{-1} \right)} \\ &= \frac{\omega}{I_1(2\omega)}, \end{aligned} \quad (4.15)$$

where

$$\omega = \frac{(\kappa \kappa_0 e^s)^{\frac{1}{2}}}{e^s - 1}. \quad (4.16)$$

As described by Cairn et al. (1987), all the current human population is originally descended from a single population in Africa dated around 140,000 to

²Here we have imposed the property of the modified Bessel Function of order 1 that $I_1(2w) = w + O(w^2)$ as $w \rightarrow 0$

200,000 years ago. We define the “current” time to be the end of the upper paleolithic period, which is around 10,000 BC. The reason is because, after that time, the growth rate λ increased dramatically following the invention of agriculture, whereas λ was (very approximately) constant for the $\approx 150,000$ years during the upper paleolithic period (Biraben, 1979). Goldewijk et al. (2010) estimates the size of the human population around the time to be around 1 million to 20 million. Thus, for the sake of the simulation, we will set $M(t) = 3 \times 10^6$ to be the female population at the end of the paleolithic. As estimated by Biraben (1979), the annual growth rate is 0.008% per generation. Assuming one generation takes around 20 years, this corresponds with the annual growth $\log \lambda = 0.0015$. In addition, as argued in Burden and Simon (2016), we will be using $\sigma^2 = 2$. These parameters correspond to a ‘current’ scaled population of $\kappa = 4500$. Last thing, we will define the likelihood function defined as the probability that the population at time s won’t exceed the observed population κ given as follows;

$$\begin{aligned} L(s, \kappa_0 | \kappa) &= \text{Prob}(\kappa(s) \leq \kappa | \kappa(0) = \kappa_0) \\ &= \int_0^{\kappa/(\kappa_0 e^s)} f_{Z_{tot}}(u | s, \kappa_0) du. \end{aligned} \quad (4.17)$$

As seen in fig.4.2, the contour map of Eq. (4.15) grades from a probability near 1 on the left edge of the plot corresponding to earlier initial time to probability near 0 at recent times. This means, if the stochastic system starts near the left-hand edge of the plot, the probability that the current population has a common ancestor amongst the initial population of κ_0 at $s = 0$ is nearly 1. On the other hand, if the system starts on the right-hand edge of the plot, the current population is almost certainly not descended from a single ancestor at the initial time. Therefore, it implies that the most likely occurrence of mtE was at a time and population in the intermediate rapid changeover region as marked by the dot on the plot. Having said that, Burden and Simon (2016) found that in the original parameters used, the estimated population at the time of mtE was approximated to be around 1030 and occurred at 5610 generations before 10,000 BCE which is around 125,000 BCE.

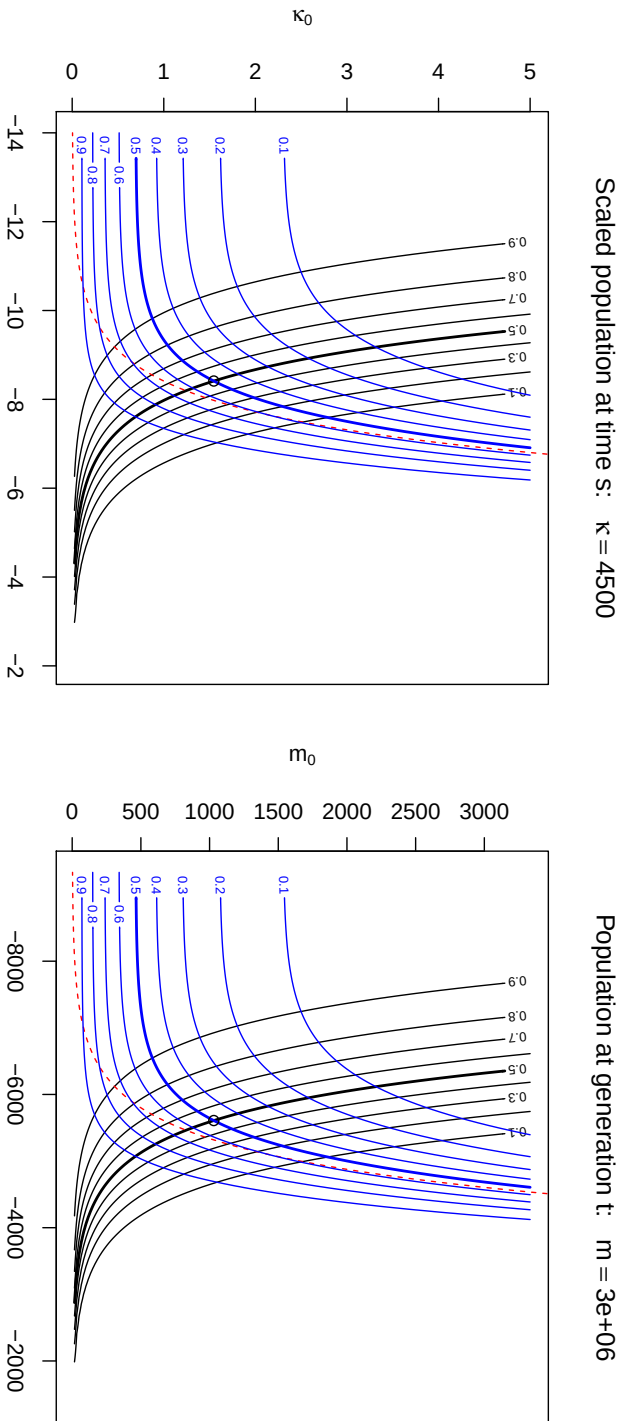


Figure 4.2: Contour maps (black lines) of the probability density described in Eq. 4.15 that the “current” population is descended from a single individual in the past with the input parameters described above. Also shown in blue lines, the likelihood function defined as the probability that the population at time s won’t exceed the observed population κ for a given starting position shown on the graphs. Superimposed in red line is the plot for an exponential growth function. The left graph corresponds to the scaled parameters, while the right graph corresponds to the original parameters.

4.2 The Coalescent Time of Two Individuals with Wright-Fisher Model Imposed

In the previous section, the main assumption is the population grows stochastically as time progresses, hence it follows directly as a G-W process. Here, we assume that the the population size grows deterministically with variance $\sigma^2 = 1$, which corresponds with the W-F model. In the section, we will discuss the result from Slatkin and Hudson (1991) to find the the coalescent time i.e the time to the MRCA of a sample of two individuals at "current" time with W-F model imposed. From page 558 of Slatkin and Hudson (1991), the probability density function for the coalescent time of two individuals (written in this thesis' notations) is given by (See fig.4.3);

$$f_S(s) = \frac{2e^s}{\kappa} \exp\left(\frac{2(1-e^s)}{\kappa}\right) \quad (4.18)$$

with the mean coalescent time given as follows;

$$\begin{aligned} E(S) &= \int_0^\infty \frac{2se^s}{\kappa} \exp\left(\frac{2(1-e^s)}{\kappa}\right) ds \\ &\quad u = e^s \text{ and } du = e^s ds \\ &= \int_1^\infty \frac{2\ln(u)}{\kappa} \exp\left(\frac{2(1-u)}{\kappa}\right) du \\ &= \frac{2e^{\frac{2}{\kappa}}}{\kappa} \int_1^\infty \ln(u) \exp\left(\frac{-2u}{\kappa}\right) du \\ &= \frac{2e^{\frac{2}{\kappa}}}{\kappa} \int_1^\infty \frac{\kappa}{2} \frac{e^{\frac{-2u}{\kappa}}}{u} du \\ &= e^{\frac{2}{\kappa}} \int_1^\infty \frac{e^{\frac{-2u}{\kappa}}}{u} du \\ &\quad t = \frac{2u}{\kappa} \text{ and } du = \frac{\kappa}{2} dt = \frac{u}{t} dt \\ &= e^{\frac{2}{\kappa}} \int_{\frac{2}{\kappa}}^\infty \frac{e^{-t}}{t} dt \\ &= e^{\frac{2}{\kappa}} E_1\left(\frac{2}{\kappa}\right), \end{aligned} \quad (4.19)$$

where $E_1(z) = \int_z^\infty \frac{e^{-t}}{t} dt$ is the exponential integral as defined in Abramowitz and Stegun (1964) chapter 5. Slatkin and Hudson (1991) notes that for $\alpha \gg 1$ we have $E_1\left(\frac{2}{\kappa}\right) \approx \ln\left(\frac{\kappa}{2}\right) - \gamma$ where γ is the Euler-Mascheroni constant which is

approximated to be around 0.577. Therefore, from Eq. (4.19) we have the mean coalescent time of the a sample of size two,

$$E(S) = e^{\frac{2}{\kappa}} \left(\ln \left(\frac{\kappa}{2} \right) - \gamma \right), \quad (4.20)$$

then under the same input parameters, we get that the time of the MRCA of two randomly chosen individuals is around $s = 7.144645$, which in original parameter corresponds to $\tau \approx 4763$ generations, hence given that one generation takes around 20 years, we get the time of the MRCA is around 105,260 BCE.

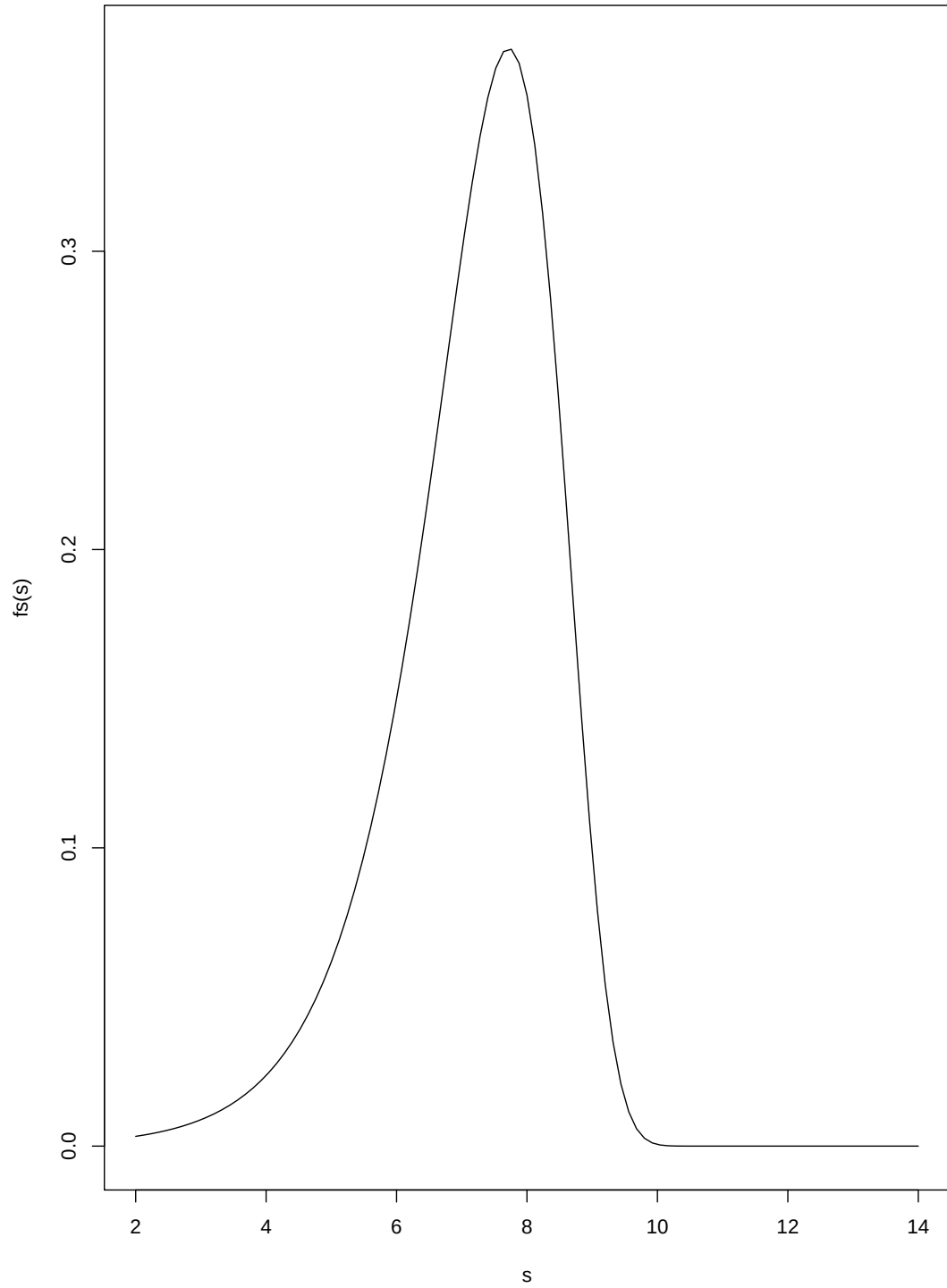


Figure 4.3: The plot of the density function defined in Eq. (4.18) .

Chapter 5

Novel Results and Discussion

We will derive new results in this chapter by formulating the density function of the MRCA of a sample of two randomly chosen individuals using a G-W process, and later generalize for the case of any sample size. Then, we will find their mean coalescent time and estimated population size at the time of the MRCA.

5.1 The Most Recent Common Ancestor of a Sample of Two Individuals in a Galton-Watson Process

In the previous section, Burden and Simon (2016) discussed the common ancestor of the entire population at the current time. Here we are interested in the following question: Suppose we pick two random individuals from the current population, how long do we need to trace back in time until we find the common ancestor of these two individuals ?. Under G-W process and the same parameters used as before, we will start examining this case by first defining an event;

$$\begin{aligned} \mathcal{E}_2 = & \text{Event that two randomly chosen individuals at the "current" time } s \\ & \text{are descended from a common ancestor at time } 0. \end{aligned} \tag{5.1}$$

Recall that we have previously defined the scaled population size as $K(s) = \frac{2M(\tau)\log\lambda}{\sigma^2}$. Then, if we divide the current population as two sub-populations, say $K_1 = xK$ and $K_2 = (1 - x)K$ where K_1 are the direct descendants from that common ancestor in the past, while K_2 are descended from the remaining $m_0 - 1$ individuals at time 0. Hence, we are interested in finding the joint density

function of K_1 and K_2 given that we have an initial population size κ_0 and an initial fraction of population $\frac{1}{m_0}$ whose descendants constitute a fraction x of the current population and an initial fraction $1 - \frac{1}{m_0}$ whose descendants constitute a fraction $1 - x$ of the current population. Denote this density function by $f_{K_1, K_2}(\kappa_1, \kappa_2 | s, \kappa_0, m_0)$. By the independence of parallel G-W processes, we then have;

$$\begin{aligned} f_{K_1, K_2}(\kappa_1, \kappa_2 | s, \kappa_0, m_0) d\kappa_1 d\kappa_2 &= f_{K_1} \left(\kappa_1 \middle| s, \kappa_0, \frac{1}{m_0} \right) d\kappa_1 f_{K_2} \left(\kappa_2 \middle| s, \kappa_0, 1 - \frac{1}{m_0} \right) d\kappa_2 \\ &= f_K \left(x\kappa \middle| s, \kappa_0, \frac{1}{m_0} \right) d\kappa_1 \\ &\quad \times f_K \left((1-x)\kappa \middle| s, \kappa_0, 1 - \frac{1}{m_0} \right) d\kappa_2, \end{aligned} \quad (5.2)$$

where the density f_K will be defined below by Eq. (5.5). From Eq. (4.7), we know that Z and K are related via

$$Z = \frac{K(s)}{\kappa_0 e^s}. \quad (5.3)$$

Using Eq. (5.3) we can rewrite the solution to Forward-Kolmogorov equation in terms of K ;

$$\begin{aligned} f_K(\kappa | s, \kappa_0, x_0) &= f_Z \left(\frac{\kappa}{\kappa_0 e^s} \middle| s, \kappa_0, x_0 \right) \frac{dz}{d\kappa} \\ &= \frac{1}{\kappa_0 e^s} f_Z \left(\frac{\kappa}{\kappa_0 e^s} \middle| s, \kappa_0, x_0 \right). \end{aligned} \quad (5.4)$$

Hence,

$$\begin{aligned} f_K(\kappa | s, \kappa_0, x_0) &= \frac{1}{e^s - 1} \left(\frac{x_0 \kappa_0 e^s}{\kappa} \right)^{\frac{1}{2}} \exp \left[\frac{-\kappa_0 \left(x_0 + \frac{\kappa}{\kappa_0} e^{-s} \right)}{1 - e^{-s}} \right] I_1 \left(\frac{2\kappa_0 \left(x_0 \frac{\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}}}{1 - e^{-s}} \right) \\ &\quad + \frac{1}{\kappa_0 e^s} \delta \left(\frac{\kappa}{\kappa_0 e^s} \right) \Pi_0(s, \kappa_0, x_0) \end{aligned} \quad (5.5)$$

where $\Pi_0(s, \kappa_0, x_0)$ is the probability that the fraction x_0 of the initial population eventually became extinct before time s with initial population κ_0 .

Then, under the similar argument as in Burden and Simon (2016) for the common ancestor of the entire population in the current time, we can apply the

same thing to find the probability density that two randomly chosen individuals from population size $K(s)$ are descended from a single individual at time $s = 0$ given we have the initial population size κ_0 as follows ;

$$\begin{aligned} \text{Prob} (K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0) \\ = m_0 \int_0^1 x^2 f_{X,K} (x, \kappa | s, \kappa_0) dx d\kappa. \end{aligned} \quad (5.6)$$

Note that m_0 appears since we have m_0 possibilities for the common ancestor of the two individuals, and x^2 appears since both individuals come from a fraction x of the current total population size. In order to get the joint probability density function in X and K , we can apply the following using Jacobian transformation.

$$\begin{aligned} f_{X,K} (x, \kappa | s, \kappa_0) &= f_{K_1, K_2} (\kappa_1, \kappa_2 | s, \kappa_0, m_0) \left| \frac{\partial(\kappa_1, \kappa_2)}{\partial(x, \kappa)} \right| \\ &= f_{K_1, K_2} (\kappa_1, \kappa_2 | s, \kappa_0, m_0) \kappa \\ &= \kappa f_K \left(x\kappa \middle| s, \kappa_0, \frac{1}{m_0} \right) f_K \left((1-x)\kappa \middle| s, \kappa_0, 1 - \frac{1}{m_0} \right). \end{aligned} \quad (5.7)$$

Substitute this back to Eq. (5.6) yields;

$$\begin{aligned} \text{Prob} (K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0) \\ = m_0 \kappa d\kappa \int_0^1 x^2 f_K \left(x\kappa \middle| s, \kappa_0, \frac{1}{m_0} \right) f_K \left((1-x)\kappa \middle| s, \kappa_0, 1 - \frac{1}{m_0} \right) dx \end{aligned} \quad (5.8)$$

where, from Eq. (5.5),

$$\begin{aligned} f_K \left(x\kappa \middle| s, \kappa_0, \frac{1}{m_0} \right) &= \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{m_0 x \kappa} \right)^{\frac{1}{2}} \exp \left(\frac{\frac{-\kappa_0}{m_0} - x \kappa e^{-s}}{1 - e^{-s}} \right) I_1 \left(\frac{2\kappa_0 \left(\frac{x\kappa}{m_0 \kappa_0 e^s} \right)^{\frac{1}{2}}}{1 - e^{-s}} \right) \\ &\quad + \delta(x\kappa) \Pi_0 \left(s, \kappa_0, \frac{1}{m_0} \right). \end{aligned} \quad (5.9)$$

Note that since m_0 is arbitrarily large, it is very high in probability that the fraction $\frac{1}{m_0}$ of the initial population size will eventually become extinct over

generations i.e. $\Pi_0\left(s, \kappa_0, \frac{1}{m_0}\right) \rightarrow 1$ as $m_0 \rightarrow \infty$. Therefore,

$$\begin{aligned}
f_K\left(x\kappa\left|s, \kappa_0, \frac{1}{m_0}\right.\right) &= \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{m_0 x \kappa}\right)^{\frac{1}{2}} \exp\left(\frac{\frac{-\kappa_0}{m_0} - x\kappa e^{-s}}{1 - e^{-s}}\right) I_1\left(\frac{2\kappa_0 \left(\frac{x\kappa}{m_0 \kappa_0 e^s}\right)^{\frac{1}{2}}}{1 - e^{-s}}\right) + \delta(x\kappa) \\
&= \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{m_0 x \kappa}\right)^{\frac{1}{2}} \exp\left(\frac{\frac{-\kappa_0}{m_0} - x\kappa e^{-s}}{1 - e^{-s}}\right) I_1\left(\frac{2\kappa_0 \left(\frac{x\kappa}{m_0 \kappa_0 e^s}\right)^{\frac{1}{2}}}{1 - e^{-s}}\right) + \frac{1}{k}\delta(x)
\end{aligned} \tag{5.10}$$

But, by the property of delta function, we have $\int_0^1 x^2 \delta(x) dx = 0$. Hence, the contribution from the delta function in Eq. (5.10) is zero when being evaluated by Eq. (5.8). Therefore, we only have effectively

$$f_K\left(x\kappa\left|s, \kappa_0, \frac{1}{m_0}\right.\right) = \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{m_0 x \kappa}\right)^{\frac{1}{2}} \exp\left(\frac{\frac{-\kappa_0}{m_0} - x\kappa e^{-s}}{1 - e^{-s}}\right) I_1\left(\frac{2\kappa_0 \left(\frac{x\kappa}{m_0 \kappa_0 e^s}\right)^{\frac{1}{2}}}{1 - e^{-s}}\right). \tag{5.11}$$

It is easily observed that $\frac{\kappa_0 \left(\frac{x\kappa}{m_0 \kappa_0 e^s}\right)^{\frac{1}{2}}}{1 - e^{-s}} \rightarrow 0$ as $\frac{1}{m_0} \rightarrow 0$, hence by applying the property of modified Bessel function of first kind¹, we can rewrite Eq. (5.11) as follows;

$$\begin{aligned}
f_K\left(x\kappa\left|s, \kappa_0, \frac{1}{m_0}\right.\right) &= \frac{\kappa_0}{m_0(e^s - 1)(1 - e^{-s})} \exp\left(\frac{\frac{-\kappa_0}{m_0} - x\kappa e^{-s}}{1 - e^{-s}}\right) \\
&= \frac{\kappa_0}{m_0(e^s - 1)(1 - e^{-s})} \exp\left(\frac{-x\kappa e^{-s}}{1 - e^{-s}}\right) + O\left(\frac{1}{m_0^2}\right) \text{ as } m_0 \rightarrow \infty.
\end{aligned} \tag{5.12}$$

Similarly, applying the solution of Forward-Kolmogorov equation for the other

¹That is $I_1(2\omega) = \omega + O(\omega^2)$ as $\omega \rightarrow 0$.

sub-population and collecting all the terms involving $\frac{1}{m_0}$ gives;

$$\begin{aligned}
& f_K \left((1-x)\kappa \middle| s, \kappa_0, 1 - \frac{1}{m_0} \right) \\
&= \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s - \frac{\kappa_0 e^s}{m_0}}{(1-x)\kappa} \right)^{\frac{1}{2}} \exp \left(\frac{-\kappa_0 + \frac{\kappa_0}{m_0} - (1-x)\kappa e^{-s}}{1 - e^{-s}} \right) \\
& \quad I_1 \left(\frac{2\kappa_0 \left(\frac{(m_0-1)(1-x)\kappa}{m_0 \kappa_0 e^s} \right)^{\frac{1}{2}}}{1 - e^{-s}} \right) + \frac{1}{\kappa} \delta(1-x) \Pi_0 \left(s, \kappa_0, 1 - \frac{1}{m_0} \right) \\
&= \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{(1-x)\kappa} \right)^{\frac{1}{2}} \exp \left(\frac{-\kappa_0 - (1-x)\kappa e^{-s}}{1 - e^{-s}} \right) I_1 \left(\frac{2\kappa_0 \left(\frac{(1-x)\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}}}{1 - e^{-s}} \right) \\
& \quad + \frac{1}{\kappa} \delta(1-x) \Pi_0(s, \kappa_0, 1) + O \left(\frac{1}{m_0} \right) \text{ as } m_0 \rightarrow \infty. \tag{5.13}
\end{aligned}$$

We can ignore $O \left(\frac{1}{m_0} \right)$ terms as $m_0 \rightarrow \infty$. Note that in this case we need to consider the extinction probability as $1 - \frac{1}{m_0}$ fraction of population in the past might have gone extinct with finite probability. Substitute these back to Eq. (5.8) yields;

$$\begin{aligned}
& \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0) \\
&= \kappa d\kappa \int_0^1 x^2 \frac{\kappa_0}{(e^s - 1)(1 - e^{-s})} \exp \left(\frac{-x\kappa e^{-s}}{1 - e^{-s}} \right) \left[\frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{(1-x)\kappa} \right)^{\frac{1}{2}} \right. \\
& \quad \exp \left(\frac{-\kappa_0 - (1-x)\kappa e^{-s}}{1 - e^{-s}} \right) I_1 \left(\frac{2\kappa_0}{1 - e^{-s}} \left(\frac{(1-x)\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}} \right) \\
& \quad \left. + \frac{1}{\kappa} \delta(1-x) \Pi_0(s, \kappa_0, 1) \right] dx. \tag{5.14}
\end{aligned}$$

Setting

$$g(x) = \frac{x^2 \kappa_0}{(e^s - 1)(1 - e^{-s})} \exp \left(\frac{-x\kappa e^{-s}}{1 - e^{-s}} \right) \frac{1}{\kappa} \Pi_0(s, \kappa_0, 1) \tag{5.15}$$

we can rewrite Eq. (5.14) as,

$$\begin{aligned}
& \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0) \\
&= \kappa d\kappa \left[\int_0^1 \frac{x^2}{(1-x)^{\frac{1}{2}} (e^s - 1)^2 (1 - e^{-s})} \left(\frac{\kappa_0^3 e^s}{\kappa} \right)^{\frac{1}{2}} \exp \left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}} \right) \right. \\
& \quad \left. \times I_1 \left(\frac{2\kappa_0}{1 - e^{-s}} \left(\frac{(1-x)\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}} \right) dx + \int_0^1 g(x) \delta(1-x) dx \right]. \tag{5.16}
\end{aligned}$$

Noting that $\int_0^1 g(x)\delta(1-x)dx = g(1)$ we then have

$$\begin{aligned} & \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0) \\ &= \frac{(\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)^2 (1 - e^{-s})} \exp\left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}}\right) d\kappa \int_0^1 \frac{x^2}{(1-x)^{\frac{1}{2}}} \\ & \quad I_1\left(\frac{2\kappa_0}{1 - e^{-s}} \left(\frac{(1-x)\kappa}{\kappa_0 e^s}\right)^{\frac{1}{2}}\right) dx + \frac{\kappa_0}{(e^s - 1)(1 - e^{-s})} \exp\left(\frac{-\kappa e^{-s}}{1 - e^{-s}}\right) \Pi_0(s, \kappa_0, 1) d\kappa. \end{aligned} \quad (5.17)$$

Using the integral ²

$$\int_0^1 \frac{x^2}{(1-x)^{\frac{1}{2}}} I_1\left(c(1-x)^{\frac{1}{2}}\right) dx = \frac{16I_2(c) - 2c^2}{c^3}, \quad (5.18)$$

where

$$c = \frac{2\kappa_0}{1 - e^{-s}} \left(\frac{\kappa}{\kappa_0 e^s}\right)^{\frac{1}{2}} = \frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}, \quad (5.19)$$

we have

$$\begin{aligned} & \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0) \\ &= \frac{(\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)^2 (1 - e^{-s})} \exp\left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}}\right) \left(\frac{16I_2(c)}{(e^s - 1)c^3}\right) d\kappa \\ &= \frac{16(\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)^2 (1 - e^{-s}) c^3} I_2\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) \exp\left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}}\right) d\kappa. \end{aligned} \quad (5.20)$$

By doing the algebra, it is clear that we have

$$\frac{16(\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(1 - e^{-s}) c^3} = \frac{2}{\kappa} (e^s - 1)^2, \quad (5.21)$$

giving Eq. (5.20) as

$$\begin{aligned} & \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0) \\ &= \frac{2}{\kappa} I_2\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) \exp\left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}}\right) d\kappa. \end{aligned} \quad (5.22)$$

²The integral is found with help of online calculator application, Wolfram-Alpha.

Recall from Eq. (4.14) that Burden and Simon (2016) found that;

$$\begin{aligned} \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa] | s, \kappa_0) &= \frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{\kappa} \right)^{\frac{1}{2}} \exp \left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}} \right) \\ &\quad I_1 \left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1} \right) d\kappa. \end{aligned} \quad (5.23)$$

Hence, in a similar manner as in Eq. (4.15), the probability the two randomly chosen individuals at the "current" time are descended from one individual at $s = 0$ is;

$$\begin{aligned} \text{Prob}(\mathcal{E}_2 | K(s) = \kappa, K(0) = \kappa_0, s) &= \frac{\text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_2 | s, \kappa_0)}{\text{Prob}(K(s) \in [\kappa, \kappa + d\kappa] | s, \kappa_0)} \\ &= \frac{\frac{2}{\kappa} I_2 \left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1} \right)}{\frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{\kappa} \right)^{\frac{1}{2}} I_1 \left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1} \right)}. \end{aligned} \quad (5.24)$$

Recall that we have previously defined $\omega = \frac{(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}$, then we can write Eq. (5.24) as;

$$\text{Prob}(\mathcal{E}_2 | K(s) = \kappa, K(0) = \kappa_0, s) = \frac{\frac{2}{\kappa} I_2(2\omega)}{\frac{1}{\kappa} \omega I_1(2\omega)} = \frac{2I_2(2\omega)}{\omega I_1(2\omega)}. \quad (5.25)$$

This probability is plotted in fig.5.1.

Next we will conduct an analytic check on Eq. (5.25) to see if the result is indeed a probability. We will check for both limiting cases i.e for $\omega \rightarrow 0$ and for $\omega \rightarrow \infty$.

Check: Note that for $\omega \rightarrow 0$, we have $I_2(2\omega) \approx \frac{\omega^2}{2}$ and $I_1(2\omega) \approx \omega$ (Abramowitz and Stegun, 1964). Hence, substitute these back into Eq. (5.25) gives;

$$\text{Prob}(\mathcal{E}_2 | \kappa(s) = \kappa, \kappa(0) = \kappa_0, s) \approx \frac{\omega^2}{(\omega)(\omega)} = 1 \text{ as } \omega \rightarrow 0. \quad (5.26)$$

For $\omega \rightarrow \infty$, we have;

$$\begin{aligned} I_2(2\omega) &\approx \frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left[1 - \frac{1}{\omega} + O\left(\frac{1}{\omega^2}\right) \right] \\ I_1(2\omega) &\approx \frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left[1 - \frac{1}{4\omega} + O\left(\frac{1}{\omega^2}\right) \right]. \end{aligned} \quad (5.27)$$

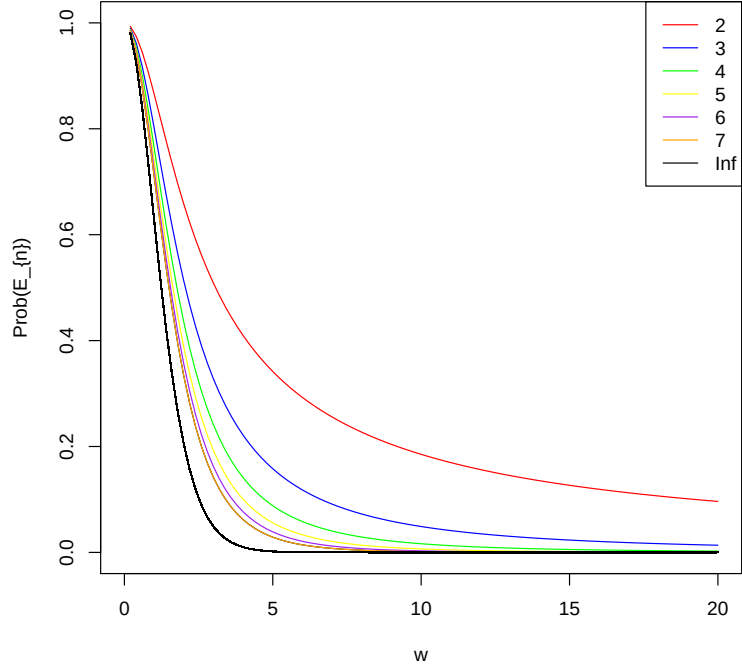


Figure 5.1: The plot of the probability that n randomly chosen individuals at time s had a common ancestor at time 0 (See Eq. (5.25) for $n = 2$, and Eq. (5.39) for the general case) given that the current, scaled population size is K for $n = 2, \dots, 7$. The black curve is the limiting case as $n \rightarrow \infty$ (See Eq. (4.15)). The parameter ω is given by Eq. (4.16).

(Abramowitz and Stegun, 1964)

Substitute these back to Eq. (5.25) gives;

$$\text{Prob}(\mathcal{E}_2 | \kappa(s) = \kappa, \kappa(0) = \kappa_0, s) \approx \frac{2 \frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left[1 - \frac{1}{\omega} + O\left(\frac{1}{\omega^2}\right)\right]}{\omega \frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left[1 - \frac{1}{4\omega} + O\left(\frac{1}{\omega^2}\right)\right]} \approx 0. \quad (5.28)$$

From the checking, it can be seen that the probability defined in Eq. (5.25) has range between 0 and 1, hence it is indeed a probability.

The plot also tells us that when $\omega \rightarrow 0$ (corresponding to $s \rightarrow \infty$), the probability that two individuals have a common ancestor at time 0 is almost certain because we have gone far enough back in time. On the other hand, when $\omega \rightarrow \infty$ (corresponds to $s \rightarrow 0$), it is very unlikely that the two individuals have the MRCA in the past because we have not gone far enough back in time.

5.2 The Most Recent Common Ancestor of a Sample of n Individuals in a Galton-Watson Process

In the previous section we discussed the case for sample of size two having a common ancestor in the past. The more interesting case is by choosing a sample of size n at present time and determine the likely parameters of their MRCA. This is what we will be trying to achieve in this section. In order to do get its probability density function, let's define the following event;

$$\begin{aligned} \mathcal{E}_n = & \text{Event that a random sample of size } n \text{ of the "current"} \\ & \text{population at time } s \text{ descended from a common ancestor at time } 0. \end{aligned} \quad (5.29)$$

Again by dividing the current population size into two sub-populations K_1 taking a fraction x of the current population, and K_2 taking a fraction $1 - x$ of the current population, and following the similar procedure as with the case of two individuals, we can define the following;

$$\begin{aligned} & \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_n | s, \kappa_0) \\ &= m_0 \kappa d\kappa \int_0^1 x^n f_K \left(x\kappa \middle| s, \kappa_0, \frac{1}{m_0} \right) f_K \left((1-x)\kappa \middle| s, \kappa_0, 1 - \frac{1}{m_0} \right) dx. \end{aligned} \quad (5.30)$$

Note that the term x^n appears since n individuals come from the fraction x of the current population size. Using Eq. (5.12) and Eq. (5.13) for Eq. (5.30) gives;

$$\begin{aligned} & \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_n | s, \kappa_0) \\ &= m_0 \kappa d\kappa \int_0^1 x^n \left[\frac{\kappa_0}{m_0(e^s - 1)(1 - e^{-s})} \exp \left(\frac{-x\kappa e^{-s}}{1 - e^{-s}} \right) \right] \\ & \quad \left[\left(\frac{1}{e^s - 1} \left(\frac{\kappa_0 e^s}{(1-x)\kappa} \right)^{\frac{1}{2}} \exp \left(\frac{-\kappa_0 - (1-x)\kappa e^{-s}}{1 - e^{-s}} \right) I_1 \left(\frac{2\kappa_0 \left(\frac{(1-x)\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}}}{1 - e^{-s}} \right) \right) \right. \\ & \quad \left. + \left(\frac{1}{\kappa} \delta(1-x) \Pi_0(s, \kappa_0, 1) \right) \right] \\ &= \kappa d\kappa \left[\int_0^1 \frac{x^n}{(1-x)^{\frac{1}{2}} (e^s - 1)^2 (1 - e^{-s})} \left(\frac{\kappa_0^3 e^s}{\kappa} \right)^{\frac{1}{2}} \exp \left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}} \right) \right. \\ & \quad \left. \times I_1 \left(\frac{2\kappa_0}{1 - e^{-s}} \left(\frac{(1-x)\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}} \right) dx \right] + \int_0^1 h(x) \delta(1-x) dx \end{aligned} \quad (5.31)$$

where;

$$h(x) = \frac{x^n \kappa_0}{\kappa (e^s - 1) (1 - e^{-s})} \exp \left(\frac{-x \kappa e^{-s}}{1 - e^{-s}} \right) \Pi_0(s, \kappa_0, 1). \quad (5.32)$$

Using the property of delta function, we have;

$$\int_0^1 h(x) \delta(1 - x) dx = h(1) = \frac{\kappa_0}{\kappa (e^s - 1) (1 - e^{-s})} \exp \left(\frac{-\kappa e^{-s}}{1 - e^{-s}} \right) \Pi_0(s, \kappa_0, 1). \quad (5.33)$$

Substitute back to Eq. (5.31) we have;

$$\begin{aligned} & \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_n | s, \kappa_0) \\ &= \frac{(\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)^2 (1 - e^{-s})} \exp \left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}} \right) d\kappa \int_0^1 \frac{x^n}{(1 - x)^{\frac{1}{2}}} \\ & \quad \times I_1 \left(\frac{2\kappa_0}{1 - e^{-s}} \left(\frac{(1 - x)\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}} \right) dx + \frac{\kappa_0}{(e^s - 1) (1 - e^{-s})} \exp \left(\frac{-\kappa e^{-s}}{1 - e^{-s}} \right) \\ & \quad \times \Pi_0(s, \kappa_0, 1) d\kappa. \end{aligned} \quad (5.34)$$

Using the integral ³,

$$\int_0^1 \frac{x^n}{(1 - x)^{\frac{1}{2}}} I_1 \left(c (1 - x)^{\frac{1}{2}} \right) dx = 2^{n+1} c^{-n-1} \Gamma(n + 1) I_n(c) - \frac{2}{c}. \quad (5.35)$$

Setting

$$c = \frac{2\kappa_0}{1 - e^{-s}} \left(\frac{\kappa}{\kappa_0 e^s} \right)^{\frac{1}{2}} = \frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1} \quad (5.36)$$

and by definition;

$$\Pi_0(s, \kappa_0, 1) = \exp \left(\frac{-\kappa_0}{1 - e^{-s}} \right), \quad (5.37)$$

³This integral is found using help of online calculator application, Wolfram-Alpha.

then Eq. (5.34) can be written as ;

$$\begin{aligned}
& \text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_n | s, \kappa_0) \\
&= \frac{(\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)(1 - e^{-s})} \exp\left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}}\right) \\
&\quad \times \left[\frac{2^{n+1} c^{-n} \Gamma(n+1) I_n(c) - 2}{(e^s - 1) c} + \left(\frac{e^{-s}}{\kappa_0 \kappa}\right)^{\frac{1}{2}} \right] d\kappa \\
&= \frac{(\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)(1 - e^{-s})} \exp\left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}}\right) \left[\frac{2^{n+1} c^{-n} \Gamma(n+1) I_n(c)}{(e^s - 1) c} \right] d\kappa \\
&= \frac{2^{n+1} \Gamma(n+1) (\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)(1 - e^{-s})} \exp\left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}}\right) \left[\frac{I_n(c)}{(e^s - 1) c^{n+1}} \right] d\kappa \\
&= \frac{2^{n+1} \Gamma(n+1) (\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)^2 (1 - e^{-s}) c^{n+1}} \exp\left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}}\right) I_n\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) d\kappa.
\end{aligned}$$

Noting that $\Gamma(n+1) = n!$, we then have;

$$\begin{aligned}
&= \frac{2^{n+1} n! (\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(e^s - 1)^2 (1 - e^{-s}) c^{n+1}} \exp\left(\frac{-\kappa_0 - \kappa e^{-s}}{1 - e^{-s}}\right) I_n\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) d\kappa \\
&= \frac{n! (\kappa_0^3 e^s \kappa)^{\frac{1}{2}}}{(1 - e^{-s}) (\kappa_0 \kappa e^s)^{\frac{1}{2}(n+1)}} (e^s - 1)^{n-1} \exp\left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}}\right) I_n\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) d\kappa \\
&= \frac{n! \kappa_0^{1-\frac{1}{2}n}}{(1 - e^{-s}) (\kappa e^s)^{\frac{n}{2}}} (e^s - 1)^{n-1} \exp\left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}}\right) I_n\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) d\kappa \\
&= \frac{n! \kappa_0^{1-\frac{1}{2}n} e^s}{(e^s - 1) \kappa^{\frac{n}{2}} e^{s\frac{1}{2}n}} (e^s - 1)^{n-1} \exp\left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}}\right) I_n\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) d\kappa \\
&= \frac{n! \kappa_0^{1-\frac{1}{2}n} e^{(1-\frac{1}{2}n)s}}{\kappa^{\frac{1}{2}n}} (e^s - 1)^{n-2} \exp\left(-\frac{\kappa_0 + \kappa e^{-s}}{1 - e^{-s}}\right) I_n\left(\frac{2(\kappa_0 \kappa e^s)^{\frac{1}{2}}}{e^s - 1}\right) d\kappa.
\end{aligned} \tag{5.38}$$

Using Eq. (5.23) again, the probability density that a random sample of size n at the current time are descended from a common individual at time 0 can be

derived as follows;

$$\begin{aligned}
\text{Prob}(\mathcal{E}_n|s, \kappa_0) &= \frac{\text{Prob}(K(s) \in [\kappa, \kappa + d\kappa], \mathcal{E}_n|s, \kappa_0)}{\text{Prob}(K(s) \in [\kappa, \kappa + d\kappa]|s, \kappa_0)} \\
&= \frac{n! \kappa_0^{\left(1-\frac{n}{2}\right)} e^{\left(1-\frac{n}{2}\right)s} (e^s - 1)^{n-2} I_n(2\omega)}{\kappa^{\frac{n}{2}-1} \omega I_1(2\omega)} \\
&= \frac{n! \kappa_0^{\left(1-\frac{n}{2}\right)} e^{\left(1-\frac{n}{2}\right)s} (\kappa \kappa_0 e^s)^{\frac{n}{2}-1} I_n(2\omega)}{\kappa^{\frac{n}{2}-1} \omega \omega^{n-2} I_1(2\omega)} \\
&= \frac{n!}{\omega^{n-1}} \frac{I_n(2\omega)}{I_1(2\omega)} \\
&= \frac{n!}{\omega^{n-1}} \frac{I_n(2\omega)}{I_1(2\omega)}. \tag{5.39}
\end{aligned}$$

Lastly, as in the case of a sample of size two, we will also do the check to confirm that Eq. (5.39) is indeed a probability.

Check:

As $\omega \rightarrow \infty$, we have;

$$\begin{aligned}
I_n(2\omega) &\approx \frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left(1 - \frac{4n^2 - 1}{16\omega} + O\left(\frac{1}{\omega^2}\right)\right) \\
I_1(2\omega) &\approx \frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left(1 - \frac{3}{16\omega} + O\left(\frac{1}{\omega^2}\right)\right). \tag{5.40}
\end{aligned}$$

(Abramowitz and Stegun, 1964)

Thus,

$$\begin{aligned}
\text{Prob}(\mathcal{E}_n|s, \kappa_0) &\approx \frac{n!}{\kappa^{\frac{1}{2}n-1} \omega^{n-1}} \frac{\frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left(1 - \frac{4n^2-1}{16\omega} + O\left(\frac{1}{\omega^2}\right)\right)}{\frac{e^{2\omega}}{\sqrt{4\pi\omega}} \left(1 - \frac{3}{16\omega} + O\left(\frac{1}{\omega^2}\right)\right)} \\
&\approx 0, \quad \text{as } \omega \rightarrow \infty. \tag{5.41}
\end{aligned}$$

As $\omega \rightarrow 0$, we have;

$$\begin{aligned}
I_n(2\omega) &\approx \frac{\omega^n}{n!} \\
I_1(2\omega) &\approx \omega. \tag{5.42}
\end{aligned}$$

(Abramowitz and Stegun, 1964)

Thus,

$$\text{Prob}(\mathcal{E}_n|s, \kappa_0) \approx 1, \quad \text{as } \omega \rightarrow 0. \tag{5.43}$$

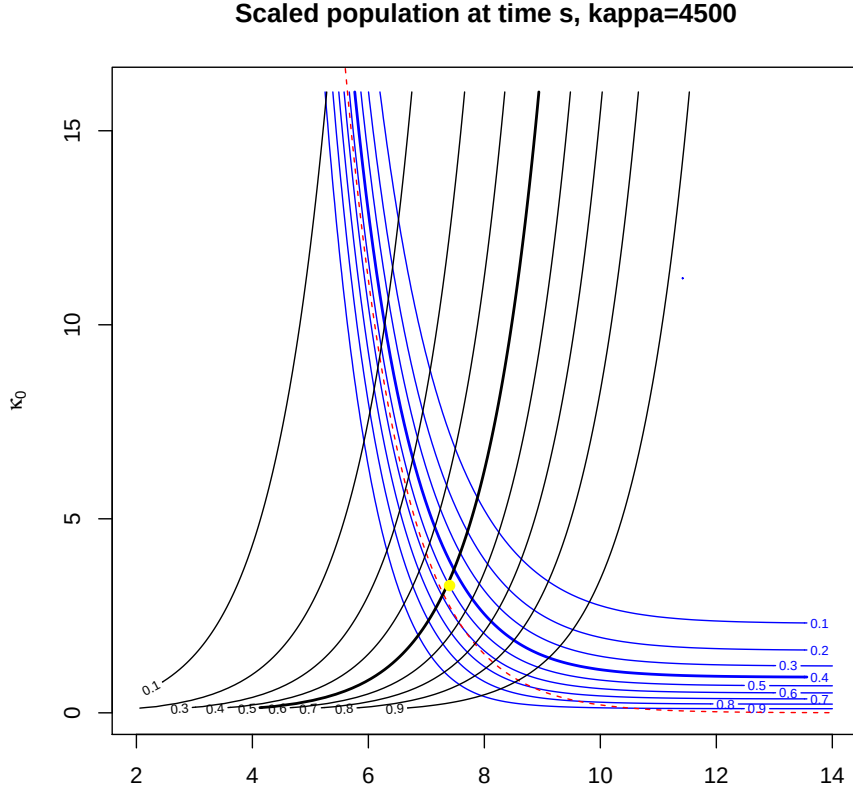


Figure 5.2: The black lines describe Eq. (5.39) that the $n=2$ randomly chosen individual at scaled time s are descended from a single ancestor in the past with scaled initial population κ_0 . Superimposed in blue are the likelihood function that the “current” population won’t exceed the given population $\kappa = 4500$ as seen in Eq. (4.17). The red-dotted line is the exponential growth function for population.

Therefore, as seen above, for both limiting cases we have that $\text{Prob}(\mathcal{E}_n|s, \kappa_0) \in [0, 1]$, hence it is indeed a probability. Again, these two limiting cases tell us that when $\omega \rightarrow \infty$ (Corresponds to $s \rightarrow 0$), the probability that a random sample of size n are descended from a common ancestor in the past is almost 0 because we have not gone back far enough in time. On the other hand, when $\omega \rightarrow 0$ (Corresponds to $s \rightarrow \infty$), it is very high in probability that a random sample of size n have a common ancestor in the past because we have gone back far enough in time. Also, from fig.5.1, it can be seen there that Eq. (4.15) is the $n \rightarrow \infty$ limit of Eq. (5.39)

Looking at fig.5.2, the yellow dot on the plot marks the point where 50% prob-

ability crossover-This is an estimate of the time and population size correspond to the MRCA. In other words, there is a 50% chance this starting point will give a sample of size two a common ancestor at time 0 and also a 50% chance that the population will grow to at least the current size.

5.3 The Mean Coalescent Time of Two Randomly Chosen Individuals with a Galton-Watson Process Imposed

In the previous sections we have derived the probability for both two and n randomly chosen individuals having a common ancestor. Another interesting question is to estimate the expected time we need to trace back in time until we find the MRCA. As described in Section 4.2, this problem is related to the coalescent. Slatkin and Hudson (1991) answered this question using the Wright-Fisher model; in our case we will assume a Galton-Watson. We will begin examining for the case of n randomly chosen individuals at time s , then specifically for the case of two randomly chosen individuals in order to be able to compare with the result found by Slatkin and Hudson (1991). Let's define the following;

$u(s, \kappa_0)$ = Probability of having single ancestor at time zero from the sample of size n taken from the current population at time s , given a starting population κ_0 .

$v(s, \kappa_0)$ = Probability that the population at time s will not exceed the observed current population given a starting population κ_0 .

As defined in previous section, $u(s, \kappa_0)$ and $v(s, \kappa_0)$ have same definitions with the following;

$$\begin{aligned} u(s, \kappa_0) &= \text{Prob}(\mathcal{E}_n | K(0) = \kappa_0, K(s) = \kappa) \\ v(s, \kappa_0) &= \text{Prob}(K(s) \leq \kappa | K(0) = \kappa_0). \end{aligned} \quad (5.44)$$

The functions u and v are both probabilities with $u, v \in [0, 1]$. We will define corresponding continuous random variables U and V , and postulate a uniform prior distribution on these random variables, leading to the joint probability density

$$f_{U,V}(u, v) = 1, \quad 0 \leq u, v \leq 1 \quad (5.45)$$

In order to locate the MRCA in the (s, κ_0) -plane, we need to apply a transformation using the Jacobian matrix to find the joint density function in terms of s and κ_0 , as follows,

$$f_{S,\kappa_0}(s, \kappa_0) = f_{U,V}(u, v) \left| \frac{\partial(u, v)}{\partial(s, \kappa_0)} \right| = |J_{U,V}(s, \kappa_0)|. \quad (5.46)$$

Recall that from Eq. (5.39) we have;

$$u(s, \kappa_0) = \frac{n!}{\omega^{n-1}} \frac{I_n(2\omega)}{I_1(2\omega)} \quad \text{where} \quad \omega = \frac{(\kappa\kappa_0 e^s)^{\frac{1}{2}}}{e^s - 1}. \quad (5.47)$$

Note that ω as defined above can be written as follows;

$$\omega = \frac{(\kappa\kappa_0 e^s)^{\frac{1}{2}}}{e^s - 1} = \frac{(\kappa\kappa_0)^{\frac{1}{2}}}{e^{\frac{s}{2}} - e^{-\frac{s}{2}}} = \frac{(\kappa\kappa_0)^{\frac{1}{2}}}{2 \sinh\left(\frac{s}{2}\right)}. \quad (5.48)$$

Hence,

$$\begin{aligned} \frac{\partial \omega}{\partial s} &= -\frac{(\kappa\kappa_0)^{\frac{1}{2}} \cosh\left(\frac{s}{2}\right)}{4 \sinh^2\left(\frac{s}{2}\right)} = -\frac{(\kappa\kappa_0)^{\frac{1}{2}} \coth\left(\frac{s}{2}\right)}{4 \sinh\left(\frac{s}{2}\right)} \\ &= -\frac{1}{2} \omega \coth\left(\frac{s}{2}\right) \end{aligned} \quad (5.49)$$

$$\begin{aligned} \frac{\partial \omega}{\partial \kappa_0} &= \frac{\frac{1}{2} \kappa^{\frac{1}{2}} \kappa_0^{-\frac{1}{2}} (2 \sinh\left(\frac{s}{2}\right))}{4 \sinh^2\left(\frac{s}{2}\right)} = \frac{\kappa^{\frac{1}{2}} \kappa_0^{-\frac{1}{2}}}{4 \sinh\left(\frac{s}{2}\right)} \\ &= \frac{(\kappa\kappa_0)^{\frac{1}{2}}}{2 \kappa_0 (2 \sinh\left(\frac{s}{2}\right))} \\ &= \frac{\omega}{2 \kappa_0} \end{aligned} \quad (5.50)$$

Note: $\frac{d}{dz} \left(\frac{I_n(z)}{z^n} \right) = \frac{I_{n+1}(z)}{z^n}$. Hence, for $z = 2\omega$ we have;

$$\begin{aligned} \frac{\partial u}{\partial \omega} &= n! \frac{d}{d\omega} \left(\frac{I_n(2\omega)}{\omega^n} \frac{\omega}{I_1(2\omega)} \right) \\ &= 2^{n-1} n! \frac{d}{d\omega} \left(\frac{\frac{I_n(2\omega)}{(2\omega)^n}}{\frac{I_1(2\omega)}{2\omega}} \right) \\ &= 2^n n! \left[\frac{\frac{I_1(2\omega)}{2\omega} \frac{I_{n+1}(2\omega)}{(2\omega)^n} - \frac{I_n(2\omega)}{(2\omega)^n} \frac{I_2(2\omega)}{2\omega}}{\left(\frac{I_1(2\omega)}{2\omega} \right)^2} \right] \\ &= \frac{2^n n!}{(2\omega)^{n-1}} \left[\frac{I_1(2\omega) I_{n+1}(2\omega) - I_2(2\omega) I_n(2\omega)}{I_1(2\omega)^2} \right] \\ &= \frac{2n!}{\omega^{n-1}} \left[\frac{I_1(2\omega) I_{n+1}(2\omega) - I_2(2\omega) I_n(2\omega)}{I_1(2\omega)^2} \right]. \end{aligned} \quad (5.51)$$

Thus,

$$\frac{du}{d\omega} = \frac{2n!}{\omega^{n-1}} (\Phi(\omega)), \quad (5.52)$$

where

$$\Phi(\omega) = \frac{I_1(2\omega)I_{n+1}(2\omega) - I_2(2\omega)I_n(2\omega)}{I_1(2\omega)^2} \quad (5.53)$$

Using chain rule, we can find $\frac{\partial u}{\partial s}$ and $\frac{\partial u}{\partial \kappa_0}$ as follows;

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial \omega} \frac{\partial \omega}{\partial s} = -\frac{n!}{\omega^{n-2}} \coth\left(\frac{s}{2}\right) \Phi(\omega) \quad (5.54)$$

$$\frac{\partial u}{\partial \kappa_0} = \frac{\partial u}{\partial \omega} \frac{\partial \omega}{\partial \kappa_0} = \frac{n!}{\kappa_0 \omega^{n-2}} \Phi(\omega). \quad (5.55)$$

We now find a convenient representation of $v(s, \kappa_0)$ that will enable us to calculate $\frac{\partial v}{\partial s}$ and $\frac{\partial v}{\partial \kappa_0}$.

Using Eq. (4.7) and Eq. (4.1), we have,

$$K(s) = \kappa_0 X(s). \quad (5.56)$$

Then, Eq. (5.44) implies;

$$\begin{aligned} v(s, \kappa_0) &= L(s, \kappa_0 | \kappa) \\ &= \text{Prob}(K(s) \leq \kappa | K(0) = \kappa_0) \\ &= \text{Prob}\left(X(s) \leq \frac{\kappa}{\kappa_0} | X(0) = 1\right) \\ &= 1 - \int_{\frac{\kappa}{\kappa_0}}^{\infty} f_X(x|s, \kappa_0) dx. \end{aligned} \quad (5.57)$$

In the last line of Eq. (5.57), using Eq. (3.14) with $x_0 = 1$ and Eq. (4.1), the density function f_X corresponding to the random variable $X(s)$ is,

$$\begin{aligned} f_X(x|s, \kappa_0) &= \frac{\kappa_0}{e^s - 1} \left(\frac{e^s}{x}\right)^{\frac{1}{2}} \exp\left[\frac{-\kappa_0(e^s + x)}{e^s - 1}\right] I_1\left[\frac{2\kappa_0(xe^s)^{\frac{1}{2}}}{e^s - 1}\right] \\ &\quad + \delta(x) \exp\left(\frac{-\kappa_0}{1 - e^{-s}}\right). \end{aligned} \quad (5.58)$$

Note that $f_X(x|s, \kappa_0)$ satisfies the Forward-Kolmogorov equation (See Eq. (3.13)),

$$\frac{\partial f_X(x|s, \kappa_0)}{\partial s} = -\frac{\partial}{\partial x}(x f_X(x|s, \kappa_0)) + \frac{1}{\kappa_0} \frac{\partial^2}{\partial x^2}(x f_X(x|s, \kappa_0)). \quad (5.59)$$

Using Eq. (5.59) we have;

$$\begin{aligned}
\frac{\partial v(s, \kappa_0)}{\partial s} &= - \int_{\frac{\kappa}{\kappa_0}}^{\infty} \left[-\frac{\partial}{\partial x} (x f_X(x|s, \kappa_0)) + \frac{1}{\kappa_0} \frac{\partial^2}{\partial x^2} (x f_X(x|s, \kappa_0)) \right] dx \\
&= \int_{\frac{\kappa}{\kappa_0}}^{\infty} \left[\frac{\partial}{\partial x} (x f_X(x|s, \kappa_0)) - \frac{1}{\kappa_0} \frac{\partial^2}{\partial x^2} (x f_X(x|s, \kappa_0)) \right] dx \\
&= \left[x f_X(x|s, \kappa_0) - \frac{1}{\kappa_0} \frac{\partial}{\partial x} (x f_X(x|s, \kappa_0)) \right]_{\frac{\kappa}{\kappa_0}}^{\infty} \\
&= -\frac{\kappa-1}{\kappa_0} f_X \left(\frac{\kappa}{\kappa_0} \middle| s, \kappa_0 \right) + \frac{\kappa}{\kappa_0^2} \frac{\partial}{\partial x} f_X \left(\frac{\kappa}{\kappa_0} \middle| s, \kappa_0 \right). \tag{5.60}
\end{aligned}$$

Since $\kappa > 0$, the δ -function in Eq. (5.58) does not contribute, hence we have;

$$f_X \left(\frac{\kappa}{\kappa_0} \middle| s, \kappa_0 \right) = \frac{\kappa_0}{e^s - 1} \left(\frac{e^s \kappa_0}{\kappa} \right)^{\frac{1}{2}} \exp \left[-\frac{\kappa + \kappa_0 e^s}{e^s - 1} \right] I_1(2\omega). \tag{5.61}$$

Taking the derivative of Eq. (5.58) with respect to x , and then setting $x = \frac{\kappa}{\kappa_0}$ yields;

$$\begin{aligned}
\frac{\partial}{\partial x} f_X \left(\frac{\kappa}{\kappa_0} \middle| s, 1, \kappa_0 \right) &= -\frac{\kappa_0^{\frac{5}{2}}}{2(e^s - 1)^2 \kappa^{\frac{3}{2}}} e^{\frac{s}{2}} e^{-\frac{\kappa + \kappa_0 e^s}{e^s - 1}} \left[-(\kappa_0 e^s \kappa)^{\frac{1}{2}} I_0(2\omega) \right. \\
&\quad \left. + (2\kappa + e^s) I_1(2\omega) + I_1(-2\omega) - (\kappa \kappa_0 e^s)^{\frac{1}{2}} I_2(2\omega) \right] \tag{5.62}
\end{aligned}$$

Note that;

$$I_v(-z) = (-z)^v z^{-v} I_v(z) \tag{5.63}$$

(Abramowitz and Stegun, 1964)

therefore, for $z = 2\omega$ and $v = 1$ we have;

$$I_1(-2\omega) = -I_1(2\omega). \tag{5.64}$$

Substitute this back to Eq. (5.62) we have;

$$\begin{aligned}
\frac{\partial}{\partial x} f_x \left(\frac{\kappa}{\kappa_0} \middle| s, 1, \kappa_0 \right) &= -\frac{\kappa_0^{\frac{5}{2}}}{2(e^s - 1)^2 \kappa^{\frac{3}{2}}} e^{\frac{s}{2}} e^{-\frac{\kappa + \kappa_0 e^s}{e^s - 1}} \left[-(\kappa_0 e^s \kappa)^{\frac{1}{2}} I_0(2\omega) \right. \\
&\quad \left. + (2\kappa + e^s) I_1(2\omega) - I_1(2\omega) - (\kappa \kappa_0 e^s)^{\frac{1}{2}} I_2(2\omega) \right]. \tag{5.65}
\end{aligned}$$

Substitute Eq. (5.61) and Eq. (5.65) into Eq. (5.60) gives us the explicit formula $\frac{\partial v(s, \kappa_0)}{\partial s}$ as follows;

$$\begin{aligned} \frac{\partial v(s, \kappa_0)}{\partial s} = & -\frac{\kappa - 1}{e^s - 1} \left(\frac{e^s \kappa_0}{\kappa} \right)^{\frac{1}{2}} \exp \left(-\frac{\kappa + \kappa_0 e^s}{e^s - 1} \right) \\ & \times I_1(2\omega) + \frac{\kappa_0}{2\kappa (e^s - 1)^2} \exp \left(-\frac{e^s \kappa_0 + \kappa}{e^s - 1} \right) \\ & \times \left[\kappa e^s I_0(2\omega) - \left(\frac{e^s \kappa}{\kappa_0} \right)^{\frac{1}{2}} (2\kappa + e^s - 1) I_1(2\omega) + e^s \kappa I_2(2\omega) \right]. \end{aligned} \quad (5.66)$$

Recall that we have;

$$f_{S, \kappa_0}(s, \kappa_0) = |J_{U, V}(s, \kappa_0)| \quad (5.67)$$

where $J_{U, V}(s, \kappa_0)$ is the Jacobian matrix.

Hence;

$$f_{S, \kappa_0}(s, \kappa_0) = \left| \frac{\partial u}{\partial s} \frac{\partial v}{\partial \kappa_0} - \frac{\partial u}{\partial \kappa_0} \frac{\partial v}{\partial s} \right|. \quad (5.68)$$

At this point, we have everything to evaluate Eq. (5.68), except $\frac{\partial v}{\partial \kappa_0}$. In order to find $\frac{\partial v}{\partial \kappa_0}$, we will do it numerically using a forward finite difference defined as follows;

$$\frac{\partial v(s, \kappa_0)}{\partial \kappa_0} = \frac{L(s, \kappa_0 + \epsilon | \kappa) - L(s, \kappa_0 | \kappa)}{\epsilon} \quad (5.69)$$

where $L(s, \kappa_0 | \kappa)$ is defined in Eq. (5.57). Substitute all of these back to Eq. (5.68) and by using R program to plot the result, we get the following contour map for the joint probability of s and κ_0 as shown in fig.5.3.

Finally, taking the integral over κ_0 of $f_{s, \kappa_0}(s, \kappa_0)$ gives the marginal distribution $f_S(s)$ as shown in fig.5.4. After finding the marginal distribution $f_S(s)$, we still need to check that it is indeed a density function. In order to check, we take the integral over s of f_S . Using the current input parameters and R program, we get;

$$\int_s f_S(s) ds \approx 0.9498109 \quad (5.70)$$

Since it's not exactly equal to 1 but close to 1, we need to make sure whether the cause is because of the rounding errors from using the Finite Difference for $\frac{\partial v}{\partial \kappa_0}$ or the way we developed the model. To check this, we take an extreme value

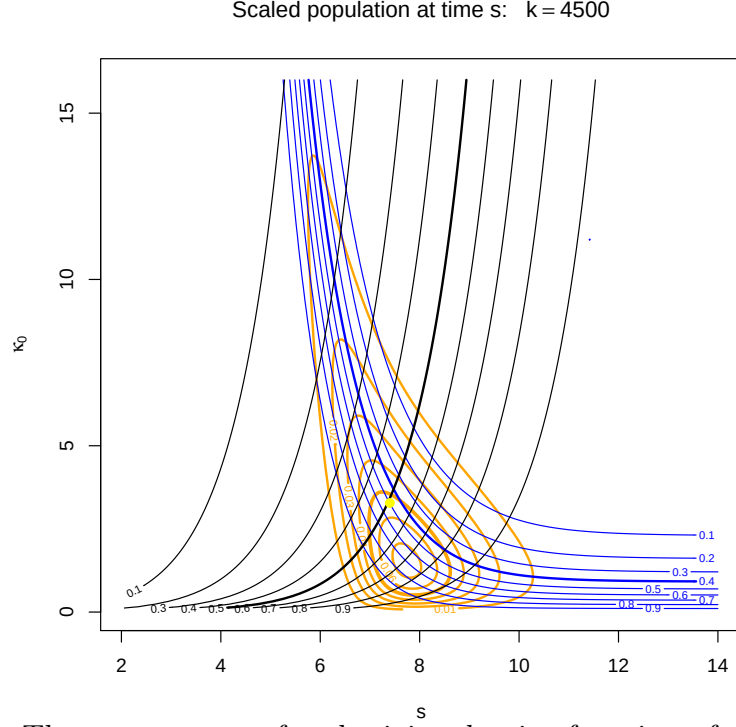


Figure 5.3: The contour map for the joint density function of s and κ_0 of two randomly chosen samples (See Eq. (5.68)) with $\kappa = 4500$, $\sigma^2 = 2$ and the growth rate $\log \lambda = 0.0015$ given the initial population size of 3×10^6 . Also, we used $\epsilon = 10^{-9}$ for the finite difference. Superimposed in the contour is fig.5.2.

for the scaled current population size $\kappa = 3$, taking the integration for this input gives a result which is very close to 1. Hence, we can be assured that it is most likely that for larger κ , we missed some part of the left tail of fig.5.4 (which do not behave properly for a very small value of s , but behaves properly if we use small value of κ) during the calculation. Thus, we can conclude that it is indeed a density function since the test with small value of κ gives integral which is very close to 1 (around 0.99).

There is another way of finding f_S without using finite difference for $\frac{\partial v}{\partial \kappa_0}$ as given below;

Taking the integral of Eq. (5.68) over κ_0 gives the marginal distribution function with respect to the scaled continuous time steps given by;

$$f_S(s) = \int_0^\infty \left| \frac{\partial u}{\partial s} \frac{\partial v}{\partial \kappa_0} - \frac{\partial u}{\partial \kappa_0} \frac{\partial v}{\partial s} \right| d\kappa_0. \quad (5.71)$$

By assuming that $\frac{\partial u}{\partial s} \frac{\partial v}{\partial \kappa_0} - \frac{\partial u}{\partial \kappa_0} \frac{\partial v}{\partial s} \geq 0$ we then have;

$$\begin{aligned}
f_S(s) &= \int_0^\infty \left(\frac{\partial u}{\partial s} \frac{\partial v}{\partial \kappa_0} - \frac{\partial u}{\partial \kappa_0} \frac{\partial v}{\partial s} \right) d\kappa_0 \\
&= \int_0^\infty \frac{\partial u}{\partial s} \frac{\partial v}{\partial \kappa_0} d\kappa_0 - \int_0^\infty \frac{\partial u}{\partial \kappa_0} \frac{\partial v}{\partial s} d\kappa_0 \\
&= \frac{\partial u}{\partial s} v(s, \kappa_0) \Big|_0^\infty - \int_0^\infty \frac{\partial^2 u}{\partial \kappa_0 \partial s} v(s, \kappa_0) d\kappa_0 - \int_0^\infty \frac{\partial u}{\partial \kappa_0} \frac{\partial v}{\partial s} d\kappa_0. \tag{5.72}
\end{aligned}$$

Then, we claim that $\frac{\partial u}{\partial s} v(s, \kappa_0) \Big|_0^\infty = 0$.

Check:

Recall that $\frac{\partial u}{\partial s} = -\frac{n!}{\omega^{n-2}} \coth\left(\frac{s}{2}\right) \Phi(\omega)$ where $\omega = \frac{(\kappa \kappa_0)^{\frac{1}{2}}}{2 \sinh(\frac{s}{2})}$. We can see here that when $\kappa_0 \rightarrow \infty$, $\frac{\partial u}{\partial s} \rightarrow 0$. While, when $\kappa_0 \rightarrow 0$ i.e when $\omega \rightarrow 0$ we have;

$$\begin{aligned}
\frac{\partial u}{\partial s} &= -\frac{n!}{\omega^{n-2}} \coth\left(\frac{s}{2}\right) \frac{I_1(2\omega)I_{n+1}(2\omega) - I_2(2\omega)I_n(2\omega)}{I_1(2\omega)^2} \\
&\approx -\frac{n!}{\omega^{n-2}} \coth\left(\frac{s}{2}\right) \frac{\frac{\omega}{2} \frac{(2\omega)^{n+1}}{(n+1)!} - \frac{1}{2} \frac{(2\omega)^2}{2} \frac{1}{2} \frac{(2\omega)^n}{n!}}{\omega^2} \\
&\approx -\frac{n!}{\omega^n} \coth\left(\frac{s}{2}\right) \left[\frac{2^n}{(n+1)!} \omega^{n+2} - \frac{2^{n-1}}{n!} \omega^{n+2} \right] \\
&\approx -n! \coth\left(\frac{s}{2}\right) \left(\frac{2^n}{(n+1)!} - \frac{2^{n-1}}{n!} \right) \omega^2 \\
&\rightarrow 0. \tag{5.73}
\end{aligned}$$

4

Hence;

$$\frac{\partial u}{\partial s} v(s, \kappa_0) \Big|_0^\infty = 0 \blacksquare \tag{5.74}$$

Thus, Eq. (5.72) becomes;

$$f_S(s) = - \int_0^\infty \frac{\partial^2 u}{\partial \kappa_0 \partial s} v(s, \kappa_0) d\kappa_0 - \int_0^\infty \frac{\partial u}{\partial \kappa_0} \frac{\partial v}{\partial s} d\kappa_0 \tag{5.75}$$

However, solving the above integral is not easy.

⁴where we have imposed the property of the modified Bessel function of the first kind that $I_v(z) \approx \frac{1}{2} \frac{z^v}{v!}$ as $z \rightarrow 0$ (Abramowitz and Stegun, 1964).

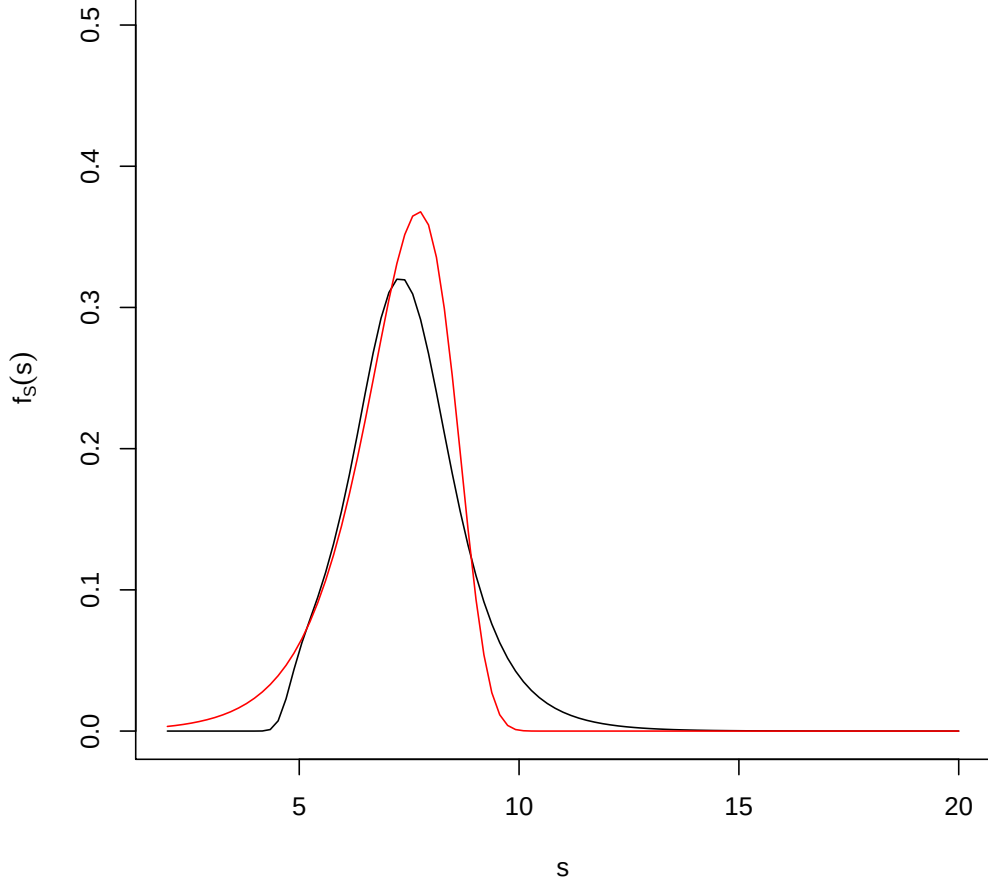


Figure 5.4: The marginal density $f_s(s)$ with respect to the scaled continuous time step s . The black curve using a Galton-Watson process, while the red curve is based on Slatkin and Hudson (1991) with the Wright-Fisher model (See Eq. (4.18)).

The numerical computation using $f_s(s)$ shows that the mean coalescent time of the sample of two individuals at current time is around $s \approx 7.451628$ and from the calculation done in fig.5.2, we found that the scaled initial population size during the coalescent event is approximated around $\kappa_0 \approx 3.28$. In the original variables, these correspond to $t \approx 4968$ generations before 10,000 BCE which is around 109,360 BCE (Assuming that one generation takes around 20 years before the next generation) with the initial population size around 2187 people. Recall from previous section that under the same input parameters, using the Wright-Fisher model imposed, the result from Slatkin and Hudson (1991) found that the mean coalescent time is around $s = 7.144645$ which corresponds to around

105,260 BCE (See fig.5.4 for the comparison). This is not far away from our result with a Galton-Watson process imposed. However, we can not make an exact comparison with result from Slatkin, since there is a slight numerical error when computing $f_S(s)$ (around 5×10^{-2}) due to applying Finite Difference for computing $\frac{\partial v}{\partial \kappa_0}$. However, assuming the error is fixed, in a real life situation, since we are talking about human population, it will be more correct to say that the population size keeps changes, stochastically, over generations i.e based on Galton-Watson process rather than a deterministically changing population size i.e based on Wright-Fisher model with an imposed exponential growth as in Slatkin and Hudson (1991).

Chapter 6

Conclusion

In this thesis, we take further the result from Burden and Simon (2016) to develop a model under a Galton-Watson process to determine the probability density function of two randomly chosen individuals at the current time have common ancestor at $s = 0$. Doing the analysis step as shown in the previous chapter, we found that;

$$\text{Prob}(\mathcal{E}_2 | \kappa(s) = \kappa, \kappa(0) = \kappa_0, s) = \frac{2I_2(2\omega)}{\omega I_1(2\omega)}. \quad (6.1)$$

Then, we generalised the problem into developing a model for n randomly chosen individuals at the current time have common ancestor at $s = 0$. The analysis shows that the density function of such event to happen is given by;

$$\text{Prob}(\mathcal{E}_n | \kappa(s) = \kappa, \kappa(0) = \kappa_0, s) = \frac{n!}{\omega^{n-1}} \frac{I_n(2\omega)}{I_1(2\omega)}. \quad (6.2)$$

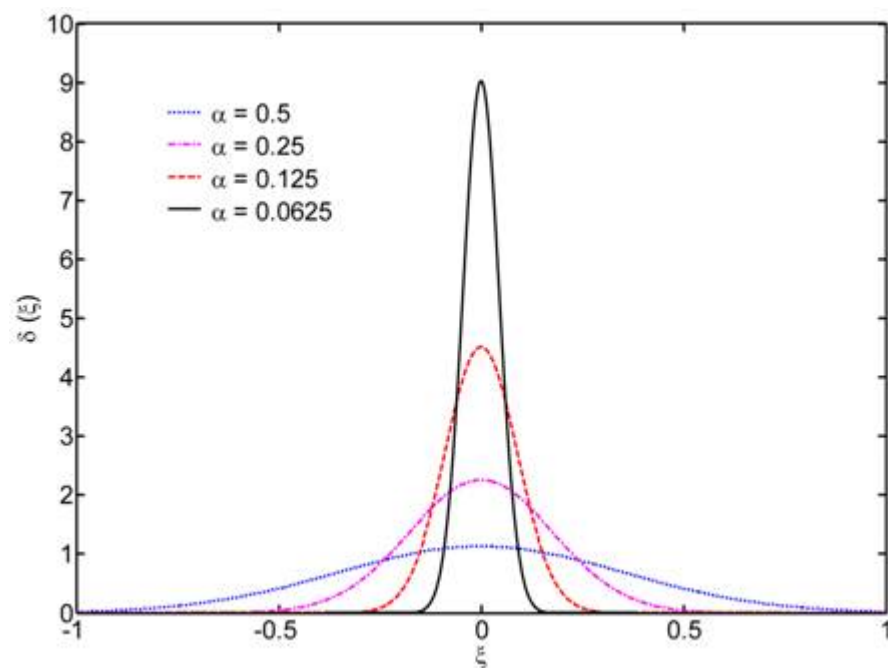
After finding these, the ultimate goal is to try making a comparison of the mean coalescent time of two randomly chosen individuals with a Galton-Watson process imposed and with the Wright-Fisher model imposed as described in Slatkin and Hudson (1991). After developing the model for this case and doing the numerical analysis, we found that the mean coalescent time of two randomly chosen individuals with a Galton-Watson process imposed is approximated to be around 109,360 BCE with the estimated population size at the time is around 2187 people, while with the Wright-Fisher model imposed we found that the mean coalescent time is around 105,260 BCE. However, in real life situation, we can be assured that the human population size keeps changes stochastically, hence using a Galton-Watson process should give better prediction for the dating of their common ancestor compared to that using the Wright-Fisher model.

There are some future works that can be done following the result derived in this paper. One of them is to find the marginal probability $f_S(s)$ directly as described in Eq. (5.75) without the need to find $\frac{\partial v(s, \kappa_0)}{\partial \kappa_0}$ first. Some other possible future works that can be explored including estimating the times at which a random sample of n individuals had l ancestors, for any $1 \leq l \leq n$; and also considering other factors such as selection and mutation, instead of only including genetic drift, in the model.

Chapter 7

Appendix

7.1 The Dirac-Delta Function and its Properties



The Dirac-Delta function, named after the physicist Paul Dirac, is a generalised function or distribution that models the density of a point mass that is equal to zero everywhere on the real number line except at zero, which its integral

over the entire real line is equal to 1. In other words;

$$\delta(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases} \quad (7.1)$$

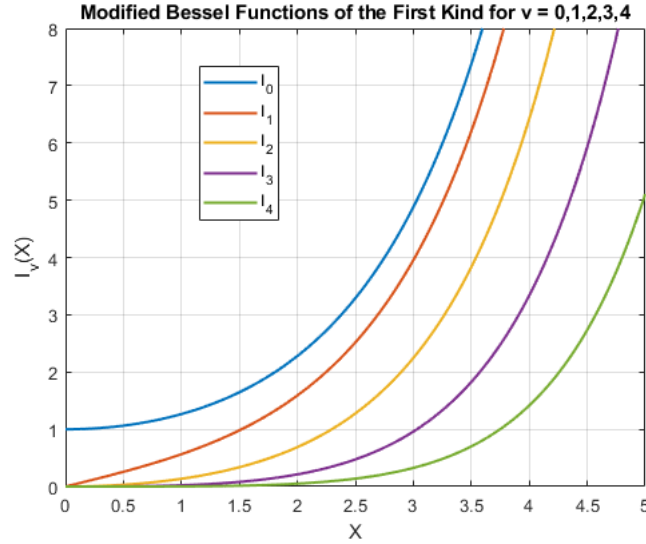
and

$$\int_{-\infty}^{\infty} \delta(x) dx = 1. \quad (7.2)$$

For any continuous function $\varphi(x)$ at $x = a$, the integrals involving the Dirac-Delta function have some properties as follows;

$$\begin{aligned} \int_{-\infty}^{\infty} \varphi(x) \delta(x - a) dx &= \varphi(a) \\ \int_{-\infty}^{\infty} \varphi(x) \delta'(x - a) dx &= -\varphi'(a) \\ \int_{-\infty}^{\infty} \varphi(x) \delta''(x - a) dx &= \varphi''(a). \end{aligned} \quad (7.3)$$

7.2 The Modified Bessel Function of the First Kind



The modified Bessel function of the first kind $I_n(x)$ is one of the solutions to the modified Bessel differential equation given by;

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} - (x^2 + n^2) y = 0. \quad (7.4)$$

The function is closely related the Bessel function of the first kind $J_n(x)$, and can be represented in the following form;

$$I_v(z) = \left(\frac{1}{2}z\right)^v \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(v+k+1)} \quad (7.5)$$

where v is an integer and $\Gamma(z)$ is the Gamma function. In integral form, the function can be represented as;

$$I_v(z) = \frac{1}{\pi} \int_0^{\pi} e^{z \cos \theta} \cos(v\theta) d\theta - \frac{\sin(v\pi)}{\pi} \int_0^{\infty} e^{-z \cosh t - vt} dt. \quad (7.6)$$

When v is an integer it gets simplified to;

$$I_v(z) = \frac{1}{\pi} \int_0^{\pi} e^{z \cos \theta} \cos(v\theta) d\theta \quad (7.7)$$

7.3 Derivation of Forward-Kolmogorov Equation

In this section, we will take a brief review on the derivation of Forward-Kolmogorov equation, beginning with Eq. (3.7) and Eq. (3.8).

It will prove useful when deriving the solution to the Forward-Kolmogorov equation to assume the existence of functions $a(u)$ and $b(u)$ satisfying

$$\begin{aligned} E(\delta X(t) | X(t) = u) &= a(u)\delta t + O(\delta t^2) \\ E(\delta X(t)^2 | X(t) = u) &= b(u)\delta t + O(\delta t^2) \\ &\vdots \\ E(\delta X(t)^k | X(t) = u) &= O(\delta t^2), k = 3, 4, \dots \end{aligned} \quad (7.8)$$

(Ewens, 2004)

From fig.7.1, it is clearly seen that

$$\begin{aligned} f(x; p, t + \delta t) dx &= \text{Prob}(x \leq X(t + \delta t) \leq x + dx | X(0) = p) \\ &= \sum_{u \in [0,1]} \text{Prob}(u \leq X(t) \leq u + du | X(0) = p) \\ &\quad \times \text{Prob}(x \leq X(t + \delta t) \leq x + dx | X(t) = u). \end{aligned} \quad (7.9)$$

Eq. (7.9) can also be written as follows;

$$f(x; p, t + \delta t) = \int_0^1 f(u; p, t) f(x; u, \delta t) du. \quad (7.10)$$

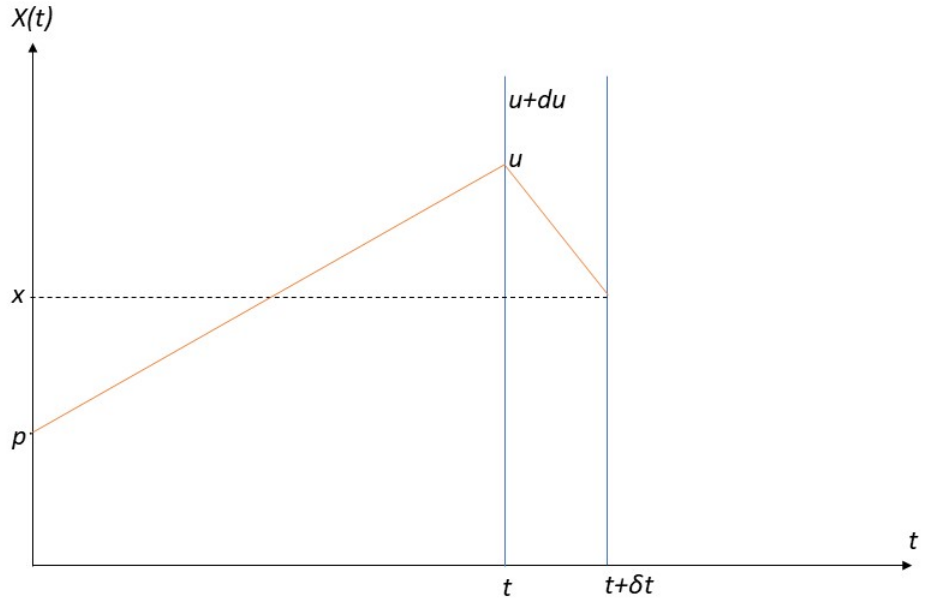


Figure 7.1: The probability distribution of $X(t)$ can be viewed as the continuum limit of a Markov process.

We claim that $f(x; u, \delta t)$ in Eq. (7.10) can be represented as

$$f(x; u, \delta t) = \delta(x - u) - a(u)\delta'(x - u)\delta t + \frac{1}{2}b(u)\delta''(x - u)\delta t + O(\delta t^2) \quad (7.11)$$

In order to prove the above claim indeed represents the distribution above, we check assumptions in Eq. (7.8).

$$\begin{aligned}
E(\delta X(t)|X(t) = u) &= E(X(t + \delta t) - X(t)|X(t) = u) \\
&= \int_0^1 (x - u) f(x; u, \delta t) dx \\
&= \int_0^1 (x - u) \left[\delta(x - u) - a(u) \delta'(x - u) \delta t + \frac{1}{2} b(u) \delta''(x - u) \delta t + O(\delta t^2) \right] dx \\
&= \int_0^1 (x - u) \delta(x - u) dx \\
&\quad + \left[- \int_0^1 (x - u) a(u) \delta'(x - u) dx + \frac{1}{2} \int_0^1 (x - u) b(u) \delta''(x - u) dx \right] \delta t \\
&\quad + O(\delta t^2) \\
&= (u - u) + \left[\int_0^1 a(u) \delta(x - u) dx + \frac{1}{2} \int_0^1 0 \cdot b(u) \delta(x - u) dx \right] \delta t + O(\delta t^2) \\
&= 0 + a(u) \delta t + 0 + O(\delta t^2) \\
&= a(u) \delta t + O(\delta t^2) \blacksquare
\end{aligned} \tag{7.12}$$

$$\begin{aligned}
E(\delta X(t)^2|X(t) = u) &= E((X(t + \delta t) - X(t))^2|X(t) = u) \\
&= \int_0^1 (x - u)^2 f(x; u, \delta t) dx \\
&= \int_0^1 (x - u)^2 \delta(x - u) dx \\
&\quad + \left[- \int_0^1 (x - u)^2 a(u) \delta'(x - u) dx + \frac{1}{2} \int_0^1 (x - u)^2 b(u) \delta''(x - u) dx \right] \delta t \\
&\quad + O(\delta t^2) \\
&= (u - u)^2 + \left[- \int_0^1 2(x - u) a(u) \delta(x - u) dx + \frac{1}{2} \int_0^1 2b(u) \delta(x - u) dx \right] \delta t \\
&\quad + O(\delta t^2) \\
&= 0 + [-2(u - u)a(u) + b(u)] \delta t + O(\delta t^2) \\
&= b(u) \delta t + O(\delta t^2) \blacksquare
\end{aligned} \tag{7.13}$$

An analogous proof follows for all higher moments, i.e. when $k > 2$.

Now, taking the Taylor expansion for the left-hand side of Eq. (7.10) at $t + \delta t$ near t gives;

$$f(x; p, t + \delta t) = f(x; p, t) + \frac{\partial f(x; p, t)}{\partial t} \delta t + O(\delta t^2). \tag{7.14}$$

Substitute Eq. (7.14) and Eq. (7.11) to Eq. (7.10) gives

$$f(x; p, t) + \frac{\partial f(x; p, t)}{\partial t} \delta t + O(\delta t^2) = \int_0^1 \left(\delta(x-u) - a(u) \delta'(x-u) \delta t + \frac{1}{2} b(u) \delta''(x-u) \delta t + O(\delta t^2) \right) f(u; p, t). \quad (7.15)$$

Let's examine the right-hand side of Eq. (7.15);

$$\begin{aligned} & \int_0^1 \left(\delta(x-u) - a(u) \delta'(x-u) \delta t + \frac{1}{2} b(u) \delta''(x-u) \delta t + O(\delta t^2) \right) f(u; p, t) \\ &= \int_0^1 \delta(x-u) f(u; p, t) du - \int_0^1 a(u) \left[-\frac{\partial \delta(x-u)}{\partial u} \right] f(u; p, t) \delta t du \\ & \quad + \frac{1}{2} \int_0^1 b(u) \left[\frac{\partial^2 \delta(x-u)}{\partial u^2} \right] f(u; p, t) \delta t du + O(\delta t^2) \\ &= f(x; p, t) - \int_0^1 \frac{\partial [a(u) f(u; p, t)]}{\partial u} \delta(x-u) \delta t du \\ & \quad + \frac{1}{2} \int_0^1 \frac{\partial^2 [b(u) f(u; p, t)]}{\partial u^2} \delta(x-u) \delta t du + O(\delta t^2) \\ &= f(x; p, t) - \frac{\partial [a(x) f(x; p, t)]}{\partial x} \delta t + \frac{1}{2} \frac{\partial^2 [b(x) f(x; p, t)]}{\partial x^2} \delta t + O(\delta t^2). \end{aligned} \quad (7.16)$$

Substitute Eq. (7.16) to Eq. (7.15) we have;

$$\begin{aligned} f(x; p, t) + \frac{\partial f(x; p, t)}{\partial t} \delta t + O(\delta t^2) &= f(x; p, t) - \frac{\partial [a(x) f(x; p, t)]}{\partial x} \delta t \\ & \quad + \frac{1}{2} \frac{\partial^2 [b(x) f(x; p, t)]}{\partial x^2} \delta t + O(\delta t^2) \\ \frac{\partial f(x; p, t)}{\partial t} \delta t &= - \frac{\partial [a(x) f(x; p, t)]}{\partial x} \delta t + \frac{1}{2} \frac{\partial^2 [b(x) f(x; p, t)]}{\partial x^2} \delta t. \end{aligned} \quad (7.17)$$

Finally, divide both sides with δt and take the limit $\delta t \rightarrow 0$ gives us the Forward-Kolmogorov equation;

$$\frac{\partial f(x; p, t)}{\partial t} = - \frac{\partial [a(x) f(x; p, t)]}{\partial x} + \frac{1}{2} \frac{\partial^2 [b(x) f(x; p, t)]}{\partial x^2} \blacksquare \quad (7.18)$$

7.4 The Solution of Forward-Kolmogorov Equation Via Bartlett's Equation

In this section, we will solve the Forward-Kolmogorov equation for a G-W diffusion, namely Eq. (3.13). It will prove more convenient to work in terms of the random variable $Z(s)$ defined by;

$$Z(s) = X(t)e^{-\alpha t} = X(t)e^{-s}, \quad (7.19)$$

and derive the corresponding density Eq. (4.2).

From Eq. (7.19) we have;

$$\begin{aligned} \frac{\partial}{\partial t} &= \frac{\partial}{\partial s} \frac{\partial s}{\partial t} + \frac{\partial}{\partial z} \frac{\partial z}{\partial s} = \alpha \frac{\partial}{\partial s} - \alpha z \frac{\partial}{\partial z} \\ \frac{\partial}{\partial x} &= \frac{\partial}{\partial s} \frac{\partial s}{\partial x} + \frac{\partial}{\partial z} \frac{\partial z}{\partial x} = 0 + e^{-s} \frac{\partial}{\partial z} \end{aligned} \quad (7.20)$$

and at fixed t , we have;

$$\begin{aligned} f_X(x, t) dx &= f_Z(z, s) dz \\ f_X(x, t) &= e^{-s} f_Z(z, s). \end{aligned} \quad (7.21)$$

Substitute Eq. (7.20) and Eq. (7.21) to Eq. (3.13) yields;

$$\begin{aligned} \alpha \frac{\partial (e^{-s} f_Z(z, s))}{\partial s} - \alpha z \frac{\partial (e^{-s} f_Z(z, s))}{\partial z} &= -\alpha e^{-s} \frac{\partial (ze^s e^{-s} f_Z(z, s))}{\partial z} + \frac{1}{2} \sigma^2 e^{-s} \\ &\quad \frac{\partial}{\partial z} \left(e^{-s} \frac{\partial (ze^s e^{-s} f_Z(z, s))}{\partial z} \right) \\ e^{-s} \frac{\partial (f_Z(z, s))}{\partial s} - e^{-s} f_Z(z, s) - ze^{-s} \frac{\partial f_Z(z, s)}{\partial z} &= -e^{-s} \frac{\partial (zf_Z(z, s))}{\partial z} + \frac{\sigma^2}{2\alpha} \\ &\quad \times e^{-2s} \frac{\partial^2 (zf_Z(z, s))}{\partial z^2} \\ \frac{\partial f_Z(z, s)}{\partial s} &= f_Z(z, s) + z \frac{\partial f_Z(z, s)}{\partial z} - \frac{\partial (zf_Z(z, s))}{\partial z} \\ &\quad + \frac{e^{-s}}{\kappa_0} \frac{\partial^2 (zf_Z(z, s))}{\partial z^2} \\ \frac{\partial f_Z(z, s)}{\partial s} &= f_Z(z, s) + z \frac{\partial f_Z(z, s)}{\partial z} - f_Z(z, s) \\ &\quad - z \frac{\partial f_Z(z, s)}{\partial z} + \frac{e^{-s}}{\kappa_0} \frac{\partial^2 (zf_Z(z, s))}{\partial z^2}. \end{aligned} \quad (7.22)$$

This leads to the following equation,

$$\frac{\partial f_Z(z, s)}{\partial s} = \frac{e^{-s}}{\kappa_0} \frac{\partial^2 (z f_Z(z, s))}{\partial z^2} \quad (7.23)$$

with the corresponding initial condition;

$$f_Z(z, 0) = \delta(z - 1). \quad (7.24)$$

Next, we will use Laplace transform to change the variable z to θ (Cox and Miller, 1978) (p.218 and p.236), that is;

$$\begin{aligned} \varphi(\theta, s) &= \mathcal{L}(f_Z(z, s)) \\ &= \int_{-\infty}^{\infty} e^{-\kappa_0 \theta z} f_Z(z, s) dz \\ &= E(e^{-\kappa_0 \theta Z(s)}). \end{aligned} \quad (7.25)$$

Hence, from Eq. (7.25) we have;

$$\frac{\partial \varphi(\theta, s)}{\partial s} = \int_{-\infty}^{\infty} e^{-\kappa_0 \theta z} \frac{\partial f_Z(z, s)}{\partial s} dz. \quad (7.26)$$

Substitute into Eq. (7.23) to Eq. (7.26) gives;

$$\begin{aligned} \frac{\partial \varphi(\theta, s)}{\partial s} &= \int_{-\infty}^{\infty} e^{-\kappa_0 \theta z} \left[\frac{e^{-s}}{\kappa_0} \frac{\partial^2 (z f_Z(z, s))}{\partial z^2} \right] dz \\ &= \frac{e^{-s}}{\kappa_0} \kappa_0^2 \theta^2 \int_{-\infty}^{\infty} e^{-\kappa_0 \theta z} z f_Z(z, s) dz \\ &= \kappa_0 e^{-s} \theta^2 \int_{-\infty}^{\infty} \left(\frac{-1}{\kappa_0} \frac{\partial e^{-\kappa_0 \theta z}}{\partial \theta} \right) f_Z(z, s) dz \\ &= -e^{-s} \theta^2 \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} e^{-\kappa_0 \theta z} f_Z(z, s) dz \\ &= -e^{-s} \theta^2 \frac{\partial \varphi(\theta, s)}{\partial \theta}. \end{aligned} \quad (7.27)$$

Thus, after the change of variables we have,

$$\frac{\partial \varphi(\theta, s)}{\partial s} + e^{-s} \theta^2 \frac{\partial \varphi(\theta, s)}{\partial \theta} = 0 \quad (7.28)$$

with the initial condition $\varphi(\theta, 0) = e^{-\kappa_0 \theta}$.

Note that φ is constant along the characteristic curve in the (θ, s) -plane given by

$$\frac{d\varphi}{ds} = \frac{\partial \varphi(\theta, s)}{\partial s} + \frac{d\theta}{ds} \frac{\partial \varphi(\theta, s)}{\partial \theta} = 0. \quad (7.29)$$

Comparing Eq. (7.28) and Eq. (7.29) yields;

$$\begin{aligned}
\frac{d\theta}{ds} &= e^{-s} \theta^2 \\
-\frac{d\theta}{\theta^2} &= -e^{-s} ds \\
\int_0^s -\frac{d\theta}{\theta^2} &= \int_0^s -e^{-s} ds \\
\frac{1}{\theta(s)} - \frac{1}{\theta(0)} &= e^{-s} - 1
\end{aligned} \tag{7.30}$$

which implies;

$$\theta(s) = \left(\frac{1}{\theta(0)} + e^{-s} - 1 \right)^{-1} = \frac{\theta(0)}{1 - (1 - e^{-s}) \theta(0)}. \tag{7.31}$$

Therefore, given the boundary at;

$$\theta(0) = \left(\frac{1}{\theta} + (1 - e^{-s}) \right)^{-1} = \frac{\theta}{1 + (1 - e^{-s}) \theta} \tag{7.32}$$

we have;

$$\begin{aligned}
\varphi(\theta, s) &= \varphi(\theta(0), 0) \\
&= e^{-\kappa_0 \theta(0)} \\
&= \exp \left[-\frac{\kappa_0 \theta}{1 + (1 - e^{-s}) \theta} \right].
\end{aligned} \tag{7.33}$$

Let's examine Eq. (7.33). Setting $A = \kappa_0$ and $B = 1 - e^{-s}$ we can see that Eq. (7.33) is of the form;

$$\begin{aligned}
\varphi(\theta, s) &= \exp \left[-\frac{A\theta}{1 + B\theta} \right] \\
&= e^{-\frac{A}{B}} e^{\frac{A}{B} - \frac{A\theta}{1+B\theta}} \\
&= e^{-\frac{A}{B}} e^{\frac{A+AB\theta-AB\theta}{B(1+B\theta)}} \\
&= e^{-\frac{A}{B}} e^{\frac{A}{B} \frac{1}{1+B\theta}}.
\end{aligned} \tag{7.34}$$

Note that the series representation for $e^{\frac{A}{B} \frac{1}{1+B\theta}}$ is,

$$e^{\frac{A}{B} \frac{1}{1+B\theta}} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{A}{B} \right)^n \frac{1}{(1+B\theta)^n}. \tag{7.35}$$

Therefore;

$$\begin{aligned}\varphi(\theta, s) &= e^{-\frac{A}{B}} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{A}{B}\right)^n \frac{1}{(1+B\theta)^n} \\ &= e^{-\frac{A}{B}} + e^{-\frac{A}{B}} \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{A}{B}\right)^n \frac{1}{(1+B\theta)^n}.\end{aligned}\quad (7.36)$$

Notice that $e^{-\frac{A}{B}}$ is the Laplace transform of $e^{-\frac{A}{B}}\delta(z)$ i.e;

$$\int_{-\infty}^{\infty} e^{-\kappa_0\theta z} e^{-\frac{A}{B}}\delta(z)dz = e^{-\frac{A}{B}}. \quad (7.37)$$

Next, we claim that $\frac{1}{(1+B\theta)^n}$ is the Laplace transform of

$$\frac{z^{n-1}\kappa_0^n}{B^n} \frac{e^{-\frac{\kappa_0 z}{B}}}{(n-1)!}, \quad (7.38)$$

since,

$$\begin{aligned}\int_0^{\infty} \left[\frac{z^{n-1}\kappa_0^n}{B^n} \frac{e^{-\frac{\kappa_0 z}{B}}}{(n-1)!} \right] e^{-\kappa_0\theta z} dz &= \frac{\kappa_0^n}{B^n(n-1)!} \int_0^{\infty} z^{n-1} e^{-\kappa_0(\frac{1}{B}+\theta)z} dz \\ &= \frac{\kappa_0^n}{B^n(n-1)!} \frac{1}{\kappa_0^n(\frac{1}{B}+\theta)^n} \int_0^{\infty} u^{n-1} e^{-u} du \\ &= \frac{1}{(n-1)!} \frac{1}{(1+B\theta)^n} \Gamma(n) \\ &= \frac{1}{(n-1)!} \frac{1}{(1+B\theta)^n} (n-1)! \\ &= \frac{1}{(1+b\theta)^n}. \blacksquare\end{aligned}\quad (7.39)$$

Thus, the inverse Laplace transform of Eq. (7.36) gives;

$$\begin{aligned}\mathcal{L}^{-1}(\varphi)(z, s) &= e^{-\frac{A}{B}}\delta(z) + e^{-\frac{A}{B}} \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{A}{B}\right)^n \frac{z^{n-1}\kappa_0^n}{B^n} \frac{e^{-\frac{\kappa_0 z}{B}}}{(n-1)!} \\ &= e^{-\frac{\kappa_0}{1-e^{-s}}}\delta(z) + e^{-\frac{\kappa_0}{1-e^{-s}}} \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{\kappa_0}{1-e^{-s}}\right)^n \frac{z^{n-1}\kappa_0^n}{(1-e^{-s})^n} \frac{e^{-\frac{\kappa_0 z}{1-e^{-s}}}}{(n-1)!} \\ &= e^{-\frac{\kappa_0}{1-e^{-s}}}\delta(z) + e^{-\frac{\kappa_0}{1-e^{-s}}} \sum_{m=0}^{\infty} \frac{1}{m!(m+1)!} \left(\frac{\kappa_0}{1-e^{-s}}\right)^{2m+2} z^m e^{-\frac{\kappa_0 z}{1-e^{-s}}}.\end{aligned}\quad (7.40)$$

Then, we claim that;

$$e^{-\frac{\kappa_0}{1-e^{-s}}} \sum_{m=0}^{\infty} \frac{1}{m!(m+1)!} \left(\frac{\kappa_0}{1-e^{-s}} \right)^{2m+2} z^m e^{-\frac{\kappa_0 z}{1-e^{-s}}} = \frac{\kappa_0}{1-e^{-s}} \frac{1}{z^{\frac{1}{2}}} e^{-\frac{\kappa_0}{1-e^{-s}}(1+z)} \times I_1 \left(\frac{2\kappa_0 z^{\frac{1}{2}}}{1-e^{-s}} \right). \quad (7.41)$$

Check: Suppose that $\hat{\kappa}_0 = \frac{\kappa_0}{1-e^{-s}}$ then the right-hand side of Eq. (7.41) becomes, using Eq. (7.5),

$$\begin{aligned} \hat{\kappa}_0 \frac{1}{z^{\frac{1}{2}}} e^{-\hat{\kappa}_0(1+z)} I_1 \left(2\hat{\kappa}_0 z^{\frac{1}{2}} \right) &= \hat{\kappa}_0 \frac{1}{z^{\frac{1}{2}}} e^{-\hat{\kappa}_0(1+z)} \sum_{m=0}^{\infty} \frac{1}{m!(m+1)!} \left(\hat{\kappa}_0 z^{\frac{1}{2}} \right)^{2m+1} \\ &= e^{-\hat{\kappa}_0} \sum_{m=0}^{\infty} \frac{1}{m!(m+1)!} \hat{\kappa}_0^{2m+2} z^m e^{-\hat{\kappa}_0 z} \\ &= e^{-\frac{\kappa_0}{1-e^{-s}}} \sum_{m=0}^{\infty} \frac{1}{m!(m+1)!} \left(\frac{\kappa_0}{1-e^{-s}} \right)^{2m+2} z^m e^{-\frac{\kappa_0 z}{1-e^{-s}}}. \blacksquare \end{aligned} \quad (7.42)$$

Therefore, Eq. (7.40) becomes;

$$\mathcal{L}^{-1}(\varphi)(z, s) = f_z(z) = e^{-\frac{\kappa_0}{1-e^{-s}}} \delta(z) + \frac{\kappa_0}{1-e^{-s}} \frac{1}{z^{\frac{1}{2}}} e^{-\frac{\kappa_0}{1-e^{-s}}(1+z)} I_1 \left(\frac{2\kappa_0 z^{\frac{1}{2}}}{1-e^{-s}} \right) \quad (7.43)$$

This agrees with Eq. (4.2) with $x_0 = 1$.

7.5 R Codes for this Study

1. The Contour Map of Probability Density Function that the Entire Population Descended from a Single Individual (fig.4.2)

```
m <- 3000000
logLambda <- 0.0015
sigmaSq <- 2
#
kappa <- 2*m*logLambda/sigmaSq
#
windows(height=5.5, width=11)
oldpar <- par(mfrow=c(1,2))
```

```

gridSize <- 200
kappa0 <- seq(0, 5, length.out= gridSize+1)[-1]
s <- seq(2, 14, length.out=gridSize+1)[-1]
#
t <- s/logLambda
m0 <- kappa0*sigmaSq/2/logLambda
#
probFunction <- function(w){
  probFunction <- w/besselI(2*w, nu=1)
  return(probFunction)
}
lKappa <- length(kappa0)
lS <- length(s)
probArray <- array(dim=c(lS, lKappa))
#
for(iS in 1:lS){
  for(jKappa in 1: lKappa){
    w <- sqrt(kappa0[jKappa]*kappa*exp(-s[iS]))
    /(1 - exp(-s[iS]))
    probArray[iS, jKappa] <- probFunction(w)
  }
}
#
contour(-s[lS:1], kappa0, probArray[lS:1,], method="edge",
lwd=c(rep(1, 5), 2, rep(1, 5)), labcex=0.6,xlab="-s",
ylab=expression(kappa[0]),
main=substitute("Scaled \u2013population \u2013at \u2013time \u2013s: \u2013\u2013\u2013")
*kappa == kk, list(kk = kappa)))
#
#Superimpose curve kappa0 = kappa*exp(-s)
#
kappa0Curve <- kappa*exp(-s)
points(-s, kappa0Curve, type = "l", lty=2, col="red")
#
#Calculate 'confidence levels
#
fZBasic <- function(z, alphaTau, kappa0, x0){

```

```

#
tauKappaFactor <- kappa0/(1 - exp(-alphaTau))

#
fZBasic <- tauKappaFactor * sqrt(x0/z) *
exp(-tauKappaFactor*(x0 + z - 2*sqrt(x0*z))) *
besseli(2*tauKappaFactor*(sqrt(x0*z)), 1,
expon.scaled=TRUE)
return(fZBasic)
}

FZOfZ <- function(z, alphaTau, kappa0, x0){
  if(length(alphaTau)>1)stop('function _FZOfZ
requires _length(alphaTau) _==_1 ')
  if(length(kappa0)>1)stop('function _FZOfZ_requires
length(kappa0) _==_1 ')
  if(length(x0)>1)stop('function _FZOfZ_requires
length(x0) _==_1 ')
#
tauKappaFactor <- kappa0/(1 - exp(-alphaTau))

#
whichZZero <- z==0
whichZPos <- (1:length(z))[z>0]

#
FZOfZ <- rep(NaN, dim=length(z))
FZOfZ[whichZZero] <- exp(-tauKappaFactor*x0)
for(zIndex in whichZPos){
  FZOfZ[zIndex] <- 1 -
  integrate(fZBasic, lower = z[zIndex],
upper= Inf, alphaTau=alphaTau, kappa0=kappa0,
x0=x0)$value
}
return(FZOfZ)
}

#
#Now sweep through the kappa0-s grid calculating
#confidence levels
#

```

```

confLevel <- array(dim=c(lS , lKappa))
for(iS in 1:lS){
  thisS <- s[iS]
  for(jKappa in 1:lKappa){
    thisKappa0 <- kappa0[jKappa]
    thisZ <- kappa*exp(-thisS)/thisKappa0
    cL <- FZOfZ(thisZ , thisS , thisKappa0 , 1)
    if(abs(cL - 1) < 1.e-10){cL <- 1 - 1.e-10}
    confLevel[iS , jKappa] <- cL
  }
}

#
contour(-s[lS:1] , kappa0 , confLevel[lS:1 , ] , add=TRUE,
        col="blue" , lwd=c(rep(1 , 5) , 2 ,
        rep(1 , 5)))

#
diffFromHalf <- abs(confLevel - 0.5) + abs(probArray - 0.5)
whichIndices <- which(diffFromHalf == min(diffFromHalf) ,
  arr.ind=TRUE)
sHatRough <- s[whichIndices[1]]
kappa0HatRough <- kappa0[whichIndices[2]]

#
func <- function(x){
  SS <- x[1]
  KKappa0 <- x[2]
  ZZ <- kappa*exp(-SS)/KKappa0
  confL <- FZOfZ(ZZ , SS , KKappa0 , 1)
  w <- sqrt(KKappa0*kappa*exp(-SS))/(1 - exp(-SS))
  probA <- probFunction(w)
  func <- (confL - 0.5)^2 + (probA - 0.5)^2
  return(func)
}

#
estimates <- optim(c(sHatRough , kappa0HatRough) , func ,
method ="Nelder-Mead")$par
sHat <- estimates[1]
kappa0Hat <- estimates[2]

```

```

#
cat("\nAt m=", m, " Kappa=", kappa,
    " estimates are kappa0=", kappa0Hat, " s=",
    sHat, " m0=", m0Hat <- kappa0Hat*sigmaSq/2/logLambda,
    " t=", tHat <- sHat/logLambda)
#
points(-sHat, kappa0Hat)
#
#Second plot in terms of t and m_0
#
contour(-t[1S:1], m0, probArray[1S:1,], method="edge",
        lwd=c(rep(1, 5), 2, rep(1, 5)),
        xlab="-t",
        ylab=expression(m[0]),
        main=substitute("Population at
        ..... generation t:"*m == mm, list(mm = m)))
#
m0Curve <- kappa0Curve*sigmaSq/2/logLambda
points(-t, m0Curve, type = "l", lty=2, col="red")
#
contour(-t[1S:1], m0, confLevel[1S:1,], add=TRUE,
        col="blue", lwd=c(rep(1, 5), 2, rep(1, 5)))
#
points(-tHat, m0Hat)
#
par(oldpar)
#

```

2. The plot that n randomly chosen individuals at time s had a common ancestor at time 0 (fig.5.1).

```

windows(height=11, width=11)

```

```

addCurve <- TRUE
for(n in 2:100){
  ProbE <- function(w){
    ProbE <- (factorial(i)/w^(i-1)) *
    bessell(2*w, nu = i)/bessell(2*w, nu = 1)
  }
}

```

```

        return(ProbE)
    }
    curve(ProbE, from=0, to=20,
          xlab="w", ylab="Prob(E_{n})", ylim=c(0,1))
    addCurve <- TRUE
}

cl <- c("red", "blue", "green", "yellow", "purple", "Orange")

for (i in 2:7) {
    curve(ProbE, from = 0, to=20, add=addCurve, xlab = "w",
          ylab = "prob", ylim=c(0,1), col=cl[i-1])
    addCurve <- TRUE

    for(n in 2:100){
        Prob <- function(w){
            Prob <- w/besselI(2*w, nu=1)
            return(Prob)
        }
        curve(Prob, from=0, to=20, add=T,
              xlab="w", ylab="Prob(E_{n})", ylim=c(0,1),
              col="black")
    }
}

legend("topright", legend = c(2:7, Inf), col=c(cl, "black"),
      lty=1)

```

3. The Contour Map of Probability Density Function of the Event that Two Randomly Chosen Samples Descended from a Single Individual (fig.5.2)

```

m <- 3000000
n <- 2
logLambda <- 0.0015
sigmaSq <- 2
epsilon <- 0.000000001
#

```

```

windows(height=11, width=11)
oldpar <- par(mfrow=c(1,1))
#
gridSize <- 200
kappa0 <- seq(0, 16, length.out= gridSize+1)[-1]
s <- seq(2, 14, length.out=gridSize+1)[-1]
k <- 2*m* logLambda/sigmaSq

lKappa <- length(kappa0)
lS <- length(s)

fZBasic <- function(z, alphaTau, k0, x0){
  #
  tauKappaFactor <- k0/(1 - exp(-alphaTau))
  #
  fZBasic <- tauKappaFactor * sqrt(x0/z) *
    exp(-tauKappaFactor*(x0 + z - 2*sqrt(x0*z))) *
    besseli(2*tauKappaFactor*(sqrt(x0*z)), 1,
    expon.scaled=TRUE)
  return(fZBasic)
}

fzbasic_v <- Vectorize(fZBasic, c("z"))

#
LofSandKappa0 <- function(s, k0){
  L <- 1 - integrate(fzbasic_v, lower = k/(k0*exp(s)),
    upper= Inf, alphaTau=s, k0=k0, x0=1)$value
  return(L)
}

phifunct <- function(w){
  phifunct <- (besseli(2*w, nu=1)*besseli(2*w, nu=n+1)-
    besseli(2*w, nu=2)*besseli(2*w, nu=n))/besseli(2*w, nu=1)^2
  return(phifunct)
}

```

```

#Define du/ds
arrayduds <- array(dim=c(lS , lKappa))

duds <- function(w,s){
  duds <- -factorial(n)*phifunct(w)/((w^(n-2))*tanh(s/2))
  return(duds)
}
#
#Define du/dk0
arraydudk0 <- array(dim=c(lS , lKappa))

dudk0 <- function(w,kappa0){
  dudk0 <- factorial(n)*phifunct(w)/(kappa0*w^(n-2))
  return(dudk0)
}
#
#Define dv/ds
arraydvds <- array(dim=c(lS , lKappa))

dvds <- function(s , kappa0){
  w <- sqrt(kappa0*k*exp(-s))/(1 - exp(-s))
  dvds <- ((1-k)/(exp(s)-1))*((exp(s)*kappa0)/k)^(1/2)
  *exp(-(k+kappa0*exp(s))/(exp(s)-1))*besselI(2*w, nu=1)
  +kappa0/(2*k*(exp(s)-1)^2)*exp(-(exp(s)*kappa0+k)/(exp(s)-1))
  *(k*exp(s)*besselI(2*w, nu=0)-(exp(s)*k/kappa0)^(1/2)
  *(2*k+exp(s)-1)*besselI(2*w, nu=1)+exp(s)*k*besselI(2*w, nu=2))
  return(dvds)
}

# Define dv/dkappa0
arraydvdk0 <- array(dim=c(lS , lKappa))

dvdk0 <- function(s , k0){

  dvdk0 <- (LofSandKappa0(s , k0+epsilon)-LofSandKappa0(s , k0))
            /epsilon
  return(dvdk0)
}

```



```

}
#

#Loop steps

for(i in 1:lS){

  for (j in 1:lKappa){
    w <- sqrt(kappa0[j]*k*exp(-s[i]))/(1 - exp(-s[i]))
    arrayduds[i,j] <- duds(w,s[i])
    arraydudk0[i,j] <- dudk0(w,kappa0[j])
    arraydvds[i,j] <- dvds(s[i],kappa0[j])
    arraydvdk0[i,j] <- dvdk0(s[i],kappa0[j])
  }
}

#

#Confidence Levels and plot its contour

confLevel <- array(dim=c(lS, lKappa))
for(iS in 1:lS){
  thisS <- s[iS]
  for(jKappa in 1:lKappa){
    thisKappa0 <- kappa0[jKappa]
    cL <- LofSandKappa0(thisS, thisKappa0)
    if(abs(cL - 1) < 1.e-10){cL <- 1 - 1.e-10}
    confLevel[iS, jKappa] <- cL
  }
}

#
contour(s[1:lS], kappa0, confLevel[1:lS,],
        col="blue", lwd=c(rep(1, 5), 2, rep(1, 5)), xlab="s",
        ylab=expression(kappa[0]), main=substitute(" Scaled
        .....population at times, kappa=4500"))
#
#Superimpose curve kappa0 = kappa*exp(-s)

```

```

kappa0Curve <- k*exp(-s)
points(s, kappa0Curve, type = "l", lty=2, col="red")
#

probFunction <- function(w){
  probFunction <- (factorial(n)/w^(n-1))*(besselI(2*w, nu=n)
    /besselI(2*w,nu=1))
  return(probFunction)
}

probArray <- array(dim=c(lS, lKappa))
#
for(iS in 1:lS){
  for(jKappa in 1: lKappa){
    w <- sqrt(kappa0[jKappa]*k*exp(-s[iS]))/
      (1 - exp(-s[iS]))
    probArray[iS, jKappa] <- probFunction(w)
  }
}
#
contour(s[1:lS], kappa0, probArray[1:lS,], method="edge",add=T,
col="black",lwd=c(rep(1, 5), 2, rep(1, 5)))
#

#Determine their intersection point

diffFromHalf <- abs(probArray - 0.5) + abs(confLevel - 0.5)
whichIndices <- which(diffFromHalf == min(diffFromHalf),
arr.ind=TRUE)
sHat<- s[whichIndices[1]]
kappa0Hat <- kappa0[whichIndices[2]]

cat("(sHat, kappa0Hat) ⊥=",c(sHat, kappa0Hat))
points(sHat, kappa0Hat, pch=19, col="yellow")
#

```

4. The Contour Map for the Joint Density Function of s and κ_0 of Two Randomly Chosen Samples (fig.5.3)

```

m <- 3000000
n <- 2
logLambda <- 0.0015
sigmaSq <- 2
epsilon <- 0.000000001
#

windows(height=11, width=11)
oldpar <- par(mfrow=c(1,1))
#
gridSize <- 200
kappa0 <- seq(0, 16, length.out= gridSize+1)[-1]
s <- seq(2, 14, length.out=gridSize+1)[-1]
k <- 2*m* logLambda/sigmaSq

lKappa <- length(kappa0)
lS <- length(s)

fZBasic <- function(z, alphaTau, k0, x0){
  #
  tauKappaFactor <- k0/(1 - exp(-alphaTau))
  #
  fZBasic <- tauKappaFactor * sqrt(x0/z)
    *exp(-tauKappaFactor*(x0 + z - 2*sqrt(x0*z)))
    *besselI(2*tauKappaFactor*(sqrt(x0*z)), 1,
      expon.scaled=TRUE)
  return(fZBasic)
}

fzbasic_v <- Vectorize(fZBasic, c("z"))

#
LofSandKappa0 <- function(s, k0){
  L <- 1 - integrate(fzbasic_v, lower = k/(k0*exp(s)),

```

```

        upper= Inf , alphaTau=s , k0=k0 , x0=1)$value
    }
    return(L)
}

phifunct <- function(w){
  phifunct <- (bessellI(2*w, nu=1)*bessellI(2*w, nu=n+1)
    -bessellI(2*w, nu=2)*bessellI(2*w, nu=n))/bessellI(2*w, nu=1)^2
  return(phifunct)
}

#Define du/ds
arrayduds <- array(dim=c(1S , lKappa))

duds <- function(w,s){
  duds <- -factorial(n)*phifunct(w)/((w^(n-2))*tanh(s/2))
  return(duds)
}
#
#Define du/dk0
arraydudk0 <- array(dim=c(1S , lKappa))

dudk0 <- function(w,kappa0){
  dudk0 <- factorial(n)*phifunct(w)/(kappa0*w^(n-2))
  return(dudk0)
}
#
#Define dv/ds
arraydvds <- array(dim=c(1S , lKappa))

dvds <- function(s , kappa0){
  w <- sqrt(kappa0*k*exp(-s))/(1 - exp(-s))
  dvds <- ((1-k)/(exp(s)-1))*((exp(s)*kappa0)/k)^(1/2)
  *exp(-(k+kappa0*exp(s))/(exp(s)-1))*bessellI(2*w, nu=1)+kappa0
  /(2*k*(exp(s)-1)^2)*exp(-(exp(s)*kappa0+k)/(exp(s)-1))
  *(k*exp(s)*bessellI(2*w, nu=0)-(exp(s)*k/kappa0)^(1/2)

```

```

      *(2*k+exp(s)-1)*besselI(2*w, nu=1)+exp(s)*k*besselI(2*w, nu=2))
    return(dvds)
  }

# Define dv/dkappa0
arraydvdk0 <- array(dim=c(lS, lKappa))

dvdk0 <- function(s, k0){

  dvdk0 <- (LofSandKappa0(s, k0+epsilon)-LofSandKappa0(s, k0))
    /epsilon
  return(dvdk0)
}
#

#Loop steps

for(i in 1:lS){

  for (j in 1:lKappa){
    w <- sqrt(kappa0[j]*k*exp(-s[i]))/(1 - exp(-s[i]))
    arrayduds[i, j] <- duds(w, s[i])
    arraydudk0[i, j] <- dudk0(w, kappa0[j])
    arraydvds[i, j] <- dvds(s[i], kappa0[j])
    arraydvdk0[i, j] <- dvdk0(s[i], kappa0[j])
  }
}

jacobian <- abs( arrayduds* arraydvdk0-arraydudk0* arraydvds)

contour(s[1:lS], kappa0, jacobian[1:lS,], method="edge",
        col="Orange", lwd=c(rep(2, 5), 3, rep(2, 5)),
        labcex=0.6, xlab="s", ylab=expression(kappa[0]),
        main=substitute("Scaled population at time s:"))
*k == kk, list(kk = k)))

confLevel <- array(dim=c(lS, lKappa))

```

```

for(iS in 1:lS){
  thisS <- s[iS]
  for(jKappa in 1:lKappa){
    thisKappa0 <- kappa0[jKappa]
    cL <- LofSandKappa0(thisS , thisKappa0)
    if(abs(cL - 1) < 1.e-10){cL <- 1 - 1.e-10}
    confLevel[iS , jKappa] <- cL
  }
}
#
contour(s[1:lS] , kappa0 , confLevel[1:lS ,] ,
        col="blue" , lwd=c(rep(1 , 5) , 2 , rep(1 , 5)) , add=T)

#

probFunction <- function(w){
  probFunction <- (factorial(n)/w^(n-1))*(besselI(2*w, nu=n)
    /besselI(2*w, nu=1))
  return(probFunction)
}

probArray <- array(dim=c(lS , lKappa))
#
for(iS in 1:lS){
  for(jKappa in 1: lKappa){
    w <- sqrt(kappa0[jKappa]*k*exp(-s[iS]))/
      (1 - exp(-s[iS]))
    probArray[iS , jKappa] <- probFunction(w)
  }
}
#
contour(s[1:lS] , kappa0 , probArray[1:lS ,] , method="edge" , add=T,
col="black" , lwd=c(rep(1 , 5) , 2 , rep(1 , 5)))
#

#Determine their intersection point

```

```

diffFromHalf <- abs(probArray - 0.5) + abs(confLevel - 0.5)
whichIndices <- which(diffFromHalf == min(diffFromHalf),
arr.ind=TRUE)
sHat<- s[whichIndices[1]]
kappa0Hat <- kappa0[whichIndices[2]]

points(sHat , kappa0Hat ,pch=19,col="yellow")

```

Bibliography

- Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. U.S. Govt. Print. Off., Washington, 1964.
- Norman T. J Bailey. *The elements of stochastic processes with applications to the natural sciences*. Wiley, New York, 1964.
- Jean-Noel Biraben. Essai sur l'évolution du nombre des hommes. *Population*, 34 (1):13–25, 1979.
- Conrad J Burden and Helmut Simon. Genetic drift in populations governed by a Galton–Watson branching process. *Theoretical Population Biology*, 109:63–74, 2016.
- N. Byers. Fermi and szilard. *arXiv preprint physics/0207094*., 2002.
- R.L. Cairn et al. Mitochondrial DNA and human evolution. *Nature*, 325(3), 1987.
- RL Cann. Genetics. Y weigh in again on modern humans. *Science*, 341(6145): 465–467, 2013.
- D. R. Cox and H. D. Miller. *The Theory of Stochastic Processes*. Chapman and Hall, London, 1978.
- Krzysztof A Cyran and Marek Kimmel. Alternatives to the Wright–Fisher model: The robustness of mitochondrial Eve dating. *Theoretical Population Biology*, 78(3):165–172, 2010.
- W. J Ewens. *Mathematical Population Genetics*, volume 27. Springer, New York, 2nd ed edition, 2004. ISBN 0387201912 (acid-free paper).
- William Feller. Diffusion processes in genetics. In *Proc. Second Berkeley Symp. Math. Statist. Prob*, volume 227, page 246, 1951.

- P. Francalacci et al. Low-pass DNA sequencing of 1200 sardinians reconstructs european y-chromosome phylogeny. *Science*, 341(6145):565–569, 2013.
- F. Galton and H. W. Watson. On the probability of the extinction of families. *Journal of the Royal Anthropological Institute*, 4:138–144, 1875.
- John H. Gillespie. *Population Genetics: A Concise Guide*. Johns Hopkins University Press, 2nd edition, 2004.
- Kees Klein Goldewijk, Arthur Beusen, and Peter Janssen. Long-term dynamic modeling of global population and built-up area in a spatially explicit way: Hyde 3.1. *The Holocene*, 2010.
- Charles J Mode, Candace K Sleeman, and Towfique Raj. On the inclusion of self regulating branching processes in the working paradigm of evolutionary and population genetics. *Frontiers in Genetics*, 4:Article 11, 2013.
- Neil O’Connell. The genealogy of branching processes and the age of our most recent common ancestor. *Advances in Applied Probability*, pages 418–442, 1995.
- M Slatkin and R R Hudson. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, 129(2): 555–62, Oct 1991.
- S Zimmerman. Population size at the time of mitochondrial eve. *Human evolution*, 16(2):117–124, 2001.