Population Genetics: Classical Wright-Fisher Model and Its Extensions

Thomas McCurley
Sierra Canyon School
Chatsworth, CA

Abstract

This paper presents a mathematical analysis of the Wright-Fisher model, a fundamental framework in population genetics that describes allele frequency dynamics in finite populations. We begin with the classical neutral model, deriving exact expressions for the mean and variance of allele counts over time and establishing conditions for fixation and loss under genetic drift alone. We then extend the analysis to incorporate bidirectional mutation, proving that mutation prevents absorption and instead produces a stationary beta distribution which parameters depend on population size and mutation rates. Through computational simulations, we validate theoretical predictions for fixation probabilities, absorption times, and equilibrium distributions.

Keywords: Wright-Fisher model, population genetics, allele frequency, mutation, absorption, fixation

1 Introduction

The Wright–Fisher model is one of the central frameworks in population genetics, providing a mathematical and probabilistic description of how genetic variation evolves in finite populations. It focuses on alleles, which are different versions of the same gene that occupy the same position, or locus, on a chromosome. In a diploid population, each individual carries two alleles at each locus, one inherited from each parent. The model assumes a fixed population size with non-overlapping generations, meaning that each generation is completely

replaced by the next, and that reproduction occurs through random mating without population structure. Within this framework, the genetic composition of the next generation is formed by randomly sampling alleles from the current generation, which introduces random fluctuations known as genetic drift. As a result of drift, allele frequencies fluctuate across generations even in the absence of other evolutionary forces, and over time, the population almost surely reaches an absorbing state, in which one allele becomes fixed (present in all individuals) while the other is lost (completely absent). The likelihood of fixation depends on the allele's initial frequency, reflecting the inherently stochastic nature of inheritance in finite populations. The model can also be enriched to capture more realistic evolutionary forces. For instance, mutation allows alleles to switch from one type to another, ensuring that genetic variation can be reintroduced even after fixation. Similarly, selection accounts for differences in fitness—the relative reproductive success of different genotypes—which biases the transmission of alleles and interacts with drift in shaping genetic diversity.

Historically, the Wright–Fisher model was developed independently in the 1930s by two pioneers of population genetics: Sewall Wright in the United States and Ronald A. Fisher in the United Kingdom. Both were concerned with understanding how random sampling in finite populations influences genetic variation over time. Fisher first introduced the idea of sampling gene frequencies across discrete, non-overlapping generations in his seminal book The Genetical Theory of Natural Selection (1930), where he combined Mendelian genetics with Darwinian natural selection into a coherent mathematical framework. Around the same time, Wright (1931) published his influential paper Evolution in Mendelian Populations, which laid out a probabilistic description of how allele frequencies change under random drift, mutation, and selection in finite populations. While their approaches differed—Fisher emphasizing deterministic and statistical methods, Wright stressing stochastic processes and population subdivision—the discrete-time binomial sampling model they both described became widely known as the Wright–Fisher model. Over the decades, the model has been formalized and extended, serving as a cornerstone of modern theoretical population genetics and forming the basis for diffusion approximations, coalescent theory, and numerous applications in evolutionary biology. For accessible introductions, see Ewens (2012), Durrett (2008), and Wakeley (2009), while more advanced treatments and applications can be found in Crow and Kimura (1970) and Etheridge (2011).

2 The Wright-Fisher Model: Theoretical Framework

2.1 The Classical Wright-Fisher Model

The classical Wright-Fisher model is a discrete-time, stochastic model of allele frequency dynamics in a finite population. It provides a probabilistic description of how allele frequencies evolve across generations under the influence of genetic drift, with possible extensions to include mutation and selection. The assumptions of the model are:

- Diploid population of constant size N.
- Non-overlapping generations.
- Random mating with no population structure.
- Alleles are sampled independently to form the next generation.

Consider a biallelic locus with alleles A and a. Let the frequency of allele A in generation t be

$$p_t = \frac{X_t}{2N},$$

where X_t is the number of copies of A among the $2N$ alleles in generation t. Each new allele in generation $t+1$ is drawn independently from the parental gene pool. Thus,

$$X_t \,\big|\, X_{t-1} \sim \text{Binomial}(2N, p_{t-1})\,.$$

Equivalently, the transition probability is

$$P(X_t = k \mid X_{t-1} = i) = \binom{2N}{k} \left(\frac{i}{2N}\right)^k \left(1 - \frac{i}{2N}\right)^{2N-k}, \quad k = 0, 1, \ldots, 2N.$$

Proposition 1. The mean and variance of X_t are given by:

$$\mathbb{E}X_t = X_0, \;\text{ and }\; \mathbb{V}ar(X_t) = 2N\,X_0\Big(1 - \frac{X_0}{2N}\Big)\Big(1 - \Big(1 - \frac{1}{2N}\Big)^t\Big).$$

Proof: For given X_{t-1}, the conditional expectation and variance are:

$$\mathbb{E}\big(X_t \mid X_{t-1}\big) \;=\; 2N \cdot \frac{X_{t-1}}{2N} \;=\; X_{t-1},$$

and

$$\mathbb{V}ar\big(X_t \mid X_{t-1}\big) \;=\; 2N \cdot \frac{X_{t-1}}{2N}\Big(1 - \frac{X_{t-1}}{2N}\Big) \;=\; X_{t-1} \;-\; \frac{X_{t-1}^2}{2N}.$$

By the law of total expectation,

$$\mathbb{E}X_t \;=\; \mathbb{E}\big(\mathbb{E}[X_t \mid X_{t-1}]\big) \;=\; \mathbb{E}[X_{t-1}] \;=\; \cdots \;=\; X_0.$$

By the law of total variance,

$$\mathbb{V}ar(X_t) \;=\; \mathbb{E}\big[\mathbb{V}ar(X_t \mid X_{t-1})\big] \;+\; \mathbb{V}ar\big(\mathbb{E}[X_t \mid X_{t-1}]\big)$$

$$= \mathbb{E}\Big(X_{t-1} - \frac{X_{t-1}^2}{2N}\Big) \;+\; \mathbb{V}ar(X_{t-1}) \;=\; \mathbb{E}[X_{t-1}] \;-\; \frac{\mathbb{E}[X_{t-1}^2]}{2N} + \mathbb{V}ar(X_{t-1}).$$

$$= \mathbb{E}[X_{t-1}] \;-\; \frac{\mathbb{V}ar(X_{t-1}) \;+\; (\mathbb{E}X_{t-1})^2}{2N} + \mathbb{V}ar(X_{t-1})$$

$$= X_0 \;-\; \frac{\mathbb{V}ar(X_{t-1}) + X_0^2}{2N} \;+\; \mathbb{V}ar(X_{t-1}) \;=\; \mathbb{V}ar(X_{t-1})\Big(1 - \frac{1}{2N}\Big) \;+\; X_0\Big(1 - \frac{X_0}{2N}\Big).$$

This is an affine recurrence of the form

$$V_t = aV_{t-1} + b, \qquad a = 1 - \frac{1}{2N}, \quad b = X_0\Big(1 - \frac{X_0}{2N}\Big).$$

Solving (with $V_0 = \mathbb{V}ar(X_0) = 0$ since X_0 is fixed) yields

$$V_t = b\frac{1-a^t}{1-a} = b \cdot \frac{1 - (1 - \frac{1}{2N})^t}{\frac{1}{2N}} = 2N\, X_0\Big(1 - \frac{X_0}{2N}\Big)\Big(1 - \Big(1 - \frac{1}{2N}\Big)^t\Big).$$

Therefore,

$$\mathbb{V}ar(X_t) = 2N\, X_0\Big(1 - \frac{X_0}{2N}\Big)\Big(1 - \Big(1 - \frac{1}{2N}\Big)^t\Big). \quad \square$$

Corollary. In terms of allele frequency, $\mathbb{E}p_t \;=\; \dfrac{\mathbb{E}X_t}{2N} \;=\; \dfrac{X_0}{2N} \;=\; p_0$, and

$$\mathbb{V}ar(p_t) \;=\; \mathbb{V}ar\Big(\frac{X_t}{2N}\Big) \;=\; \frac{\mathbb{V}ar(X_t)}{(2N)^2} \;=\; \frac{1}{(2N)^2}\, 2N\, X_0\Big(1 - \frac{X_0}{2N}\Big)\Big(1 - \Big(1 - \frac{1}{2N}\Big)^t\Big)$$

$$= p_0(1-p_0)\Big(1 - \Big(1 - \frac{1}{2N}\Big)^t\Big).$$

In the Wright–Fisher model, genetic drift refers to random fluctuations in allele frequencies across generations that arise purely from chance sampling in a finite population. While genetic drift does not systematically bias allele frequencies (it is unbiased), it introduces random variation which magnitude decreases as the population size N increases.

In addition, the states $X_t = 0$ (loss of A) and $X_t = 2N$ (fixation of A) are absorbing states. From the general theory of stochastic processes (see, for example, [4]), it can be shown that the following statements are true.

- Eventually, the process reaches absorption almost surely, that is,

$$\lim_{t\to\infty} \mathbb{P}(X_t = 0 \text{ or } 2N) \;=\; \lim_{t\to\infty} \mathbb{P}(p_t = 0 \text{ or } 1) \;=\; 1.$$

- The probability of loss of allele A is $1 - p_0$, and the probability of fixation of A is p_0, that is,

$$\lim_{t\to\infty} \mathbb{P}(X_t = 0) \;=\; \lim_{t\to\infty} \mathbb{P}(p_t = 0) \;=\; 1 - p_0 \;\text{ and }\; \lim_{t\to\infty} \mathbb{P}(X_t = 2N) \;=\; \lim_{t\to\infty} \mathbb{P}(p_t = 1) \;=\; p_0.$$

- Let T_{abs} denote the time to absorption. For large N, the expected value of T_{abs} is approximately

$$\mathbb{E}T_{abs} \;\approx\; -4N\big(p_0 \ln(p_0) + (1 - p_0)\ln(1 - p_0)\big).$$

2.2 The Wright-Fisher Model with Mutation

Suppose in a basic Wright-Fisher model, a mutation $A \to a$ occurs at rate μ from one generation to another, and a mutation $a \to A$ occurs at rate ν. Then

$$X_t \sim \text{Binomial}\Big(2N,\; p_{t-1}(1 - \mu) + (1 - p_{t-1})\nu\Big).$$

First, note that by definition, since we assume that $\mu > 0$ and $\nu > 0$, absorption is not possible. If a trajectory reaches 0, allele A will be reintroduced into the population through mutation. Similarly, if allele a becomes absent, it will reappear due to mutation. Therefore, the probability of absorption in this model is zero.

Next, while exact expressions for the mean and variance of X_t (or p_t) can be derived, they are rather cumbersome. Of greater interest is the long-run behavior of this model, which we summarize in the following proposition.

Proposition 2. If the distribution of p_t is concentrated around its mean, that is, if $p_t \approx \mathbb{E}p_t$, then the limiting (stationary) distribution of p_∞ is a beta distribution with the mean

$$m_\infty \;=\; \mathbb{E}p_\infty \;=\; \lim_{t\to\infty} \mathbb{E}p_t \;=\; \frac{\nu}{\mu \,+\, \nu},$$

and an approximate variance

$$\sigma^2_\infty \;=\; \mathbb{V}ar(p_\infty) \;\approx\; \frac{m_\infty(1 - m_\infty)}{2N(\mu \,+\, \nu)}.$$

Proof: Due to the complexity of the model under consideration, we prove here only some of the stated results. Given p_{t-1}, the conditional expectation of p_t is

$$\mathbb{E}\big[p_t \,|\, p_{t-1}\big] \;=\; p_{t-1}(1-\mu) + (1-p_{t-1})\nu \;=\; (1-\mu-\nu)\, p_{t-1} \;+\; \nu.$$

Thus, the unconditional mean satisfies the following recurrence relation:

$$\mathbb{E}p_t \;=\; \mathbb{E}\big[(1-\mu-\nu)\, p_{t-1} \;+\; \nu\big] \;=\; (1-\mu-\nu)\mathbb{E}p_{t-1} \;+\; \nu.$$

Solving this with the initial condition $\mathbb{E}p_0 = p_0$ yields the following result:

$$\mathbb{E}p_t \;=\; (1-\mu-\nu)^t\, p_0 \;+\; \frac{\nu}{\mu \;+\; \nu}\Big(1-(1-\mu-\nu)^t\Big).$$

From here, the long-run mean is

$$\mathbb{E}p_\infty \;=\; \lim_{t\to\infty} \mathbb{E}p_t \;=\; \frac{\nu}{\mu \;+\; \nu}.$$

Next, by the law of total variance,

$$\begin{aligned}\sigma_t^2 \;&=\; \mathbb{V}ar(p_t) \;=\; \mathbb{E}\big[\mathbb{V}ar(p_t \,|\, p_{t-1})\big] \;+\; \mathbb{V}ar\big(\mathbb{E}[p_t \,|\, p_{t-1}]\big) \\ &=\; \frac{1}{2N}\,\mathbb{E}\left[p'_{t-1}\,(1-p'_{t-1})\right] \;+\; (1 \;-\; \mu \;-\; \nu)^2\,\sigma_{t-1}^2,\end{aligned}$$

where $p'_{t-1} \;=\; (1-\mu-\nu)p_{t-1} \;+\; \nu$. If the distribution of p_{t-1} is concentrated so that $\mathbb{E}[p'_{t-1}(1-p'_{t-1})] \;\approx\; m'_t(1-m'_t)$ where $m'_t \;=\; (1-\mu-\nu)m_t + \nu$, the variance recursion simplifies to

$$\sigma_t^2 \approx \frac{1}{2N}\, m'_t(1-m'_t) \;+\; (1-\mu-\nu)^2\sigma_{t-1}^2.$$

Passing to the limit as t grows, we obtain

$$\sigma_\infty^2 \approx \frac{1}{2N}\, m'_\infty(1-m'_\infty) \;+\; (1-\mu-\nu)^2\sigma_\infty^2$$

where $m'_\infty \;=\; (1-\mu-\nu)m_\infty \;+\; \nu$ with $m_\infty \;=\; \dfrac{\nu}{\mu \;+\; \nu}$. From here,

$$\sigma_\infty^2 \approx \frac{m_\infty(1-m_\infty)}{2N(\mu \;+\; \nu)}. \quad \square$$

Equilibrium Mean and Variance

When bidirectional mutation is introduced, with forward rate μ and backward rate ν, the mean allele frequency evolves as

$$E[p_{t+1}] = (1 - \mu - \nu)E[p_t] + \nu.$$

At equilibrium ($E[p_{t+1}] = E[p_t] = E[p_\infty]$), this yields

$$E[p_\infty] = \frac{\nu}{\mu + \nu}.$$

The stationary variance can be derived from diffusion approximations as

$$\mathrm{Var}(p_\infty) \approx \frac{E[p_\infty](1 - E[p_\infty])}{2N(\mu + \nu)}.$$

In the continuous diffusion limit, the stationary distribution of p_t is

$$p_\infty \sim \mathrm{Beta}(2N\nu,\, 2N\mu),$$

whose mean and variance correspond exactly to the expressions above.

3 Wright-Fisher Model: Applications

3.1 The Classical Wright-Fisher Model

The only parameter of this model is p_0, the initial proportion of allele A in the population. Given an observed trajectory $\{p_t,\, t = 0, \ldots, T\}$, p_0 can be estimated as the sample mean of all observed values,

$$\hat{p}_0 \;=\; \frac{1}{T} \sum_{t=1}^{T} p_t.$$

In this section, we evaluate $\hat{p}_0$ and calculate and plot the empirical mean and variance functions against the generation number, alongside the theoretical functions. We also simulate multiple trajectories to compute the proportion absorbed at 1 and the proportion absorbed at 0, as well as evaluate the average number of generations needed for absorption and compare it to the theoretical value.

3.2 The Wright-Fisher Model with Mutation

In this section, we simulate multiple trajectories, allow them to run for a sufficient number of generations, and then take their long-run values to plot a histogram. We then overlay the

fitted theoretical density curve of the stationary Beta distribution defined in Proposition 2. To estimate the mean and variance of the beta distribution, we use the following approach. Since $m_t = \mathbb{E}p_t = (1-\mu-\nu)\mathbb{E}p_{t-1} + \nu = (1-\mu-\nu)m_{t-1} + \nu$, we produce the empirical values of μ and ν by regressing $\hat{m}_t$ on $\hat{m}_{t-1}$, where $\hat{m}_t$ and $\hat{m}_{t-1}$ are sample means of trajectories for generations t and $t-1$, respectively. Then the fitted intercept and slope will satisfy

$$\widehat{\text{intercept}} = \hat{\nu}, \text{ and } \widehat{\text{slope}} = 1 - \hat{\mu} - \hat{\nu},$$

or

$$\hat{\nu} = \widehat{\text{intercept}}, \text{ and } \hat{\mu} = 1 - \widehat{\text{slope}} - \widehat{\text{intercept}}.$$

The long-run mean and variance of the beta distribution are evaluated as

$$\hat{m}_\infty = \frac{\hat{\nu}}{\hat{\mu} + \hat{\nu}},$$

and

$$\hat{\sigma}^2_\infty = \frac{\hat{m}_\infty(1-\hat{m}_\infty)}{2N(\hat{\mu} + \hat{\nu})}.$$

Parameter Estimation. The mutation parameters μ and ν can be estimated directly from temporal allele-frequency data. Because the theoretical recursion for the expectation is

$$E[p_t] = (1-\mu-\nu)E[p_{t-1}] + \nu,$$

a simple linear regression of $\hat{p}_t$ on $\hat{p}_{t-1}$ provides empirical estimates of these rates: the intercept approximates ν, and the negative of the slope deviation from one approximates $\mu+\nu$. This approach mirrors the statistical estimators applied to the simulated trajectories.

3.3 Simulation and Results

To complement the theoretical results derived above, we conducted simulations of the Wright–Fisher model under three conditions: the classical (neutral) model, the model with mutation, and the model with selection. These simulations allow direct comparison between empirical outcomes and theoretical expectations for mean allele frequency, variance, and fixation probabilities.

3.3.1 Classical Wright–Fisher Model (Neutral Drift)

We simulated 1000 independent trajectories with diploid population size $N = 300$ and initial allele frequency $p_0 = 0.3$. Each trajectory was followed until absorption at $p_t = 0$ or

$p_t = 1$. Figure 1 shows fifty representative trajectories, with the red line marking the initial frequency p_0. Random fluctuations cause the allele frequency to wander until it eventually fixes or is lost.

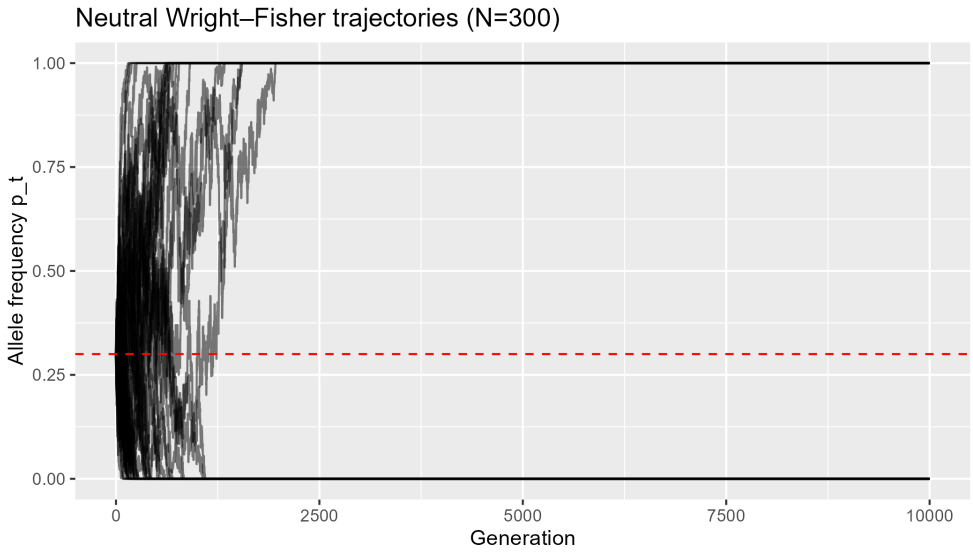

Figure 1: Fifty representative Wright–Fisher trajectories under neutrality ($N = 300$, $p_0 = 0.3$). The red line indicates the starting frequency p_0.

The empirical mean $\hat{E}[p_t]$ remained constant near p_0, as predicted for a neutral allele. The variance $\widehat{\mathrm{Var}}(p_t)$ increased with time, closely matching the theoretical expression:

$$\mathrm{Var}(p_t) = p_0(1-p_0)\left(1 - \left(1 - \frac{1}{2N}\right)^t\right).$$

Figure 2 confirms this match between theory and simulation.

Fixation statistics further validated theoretical predictions. Empirical fixation probabilities were $P(p_t \to 1) \approx 0.29$ and $P(p_t \to 0) \approx 0.71$, closely matching the theoretical probabilities p_0 and $1 - p_0$, respectively (Figure 3). The average absorption time was approximately 1450 generations, in agreement with the theoretical estimate

$$E[T_{\mathrm{abs}}] \approx -4N(p_0 \ln p_0 + (1-p_0)\ln(1-p_0)) = 1420.$$

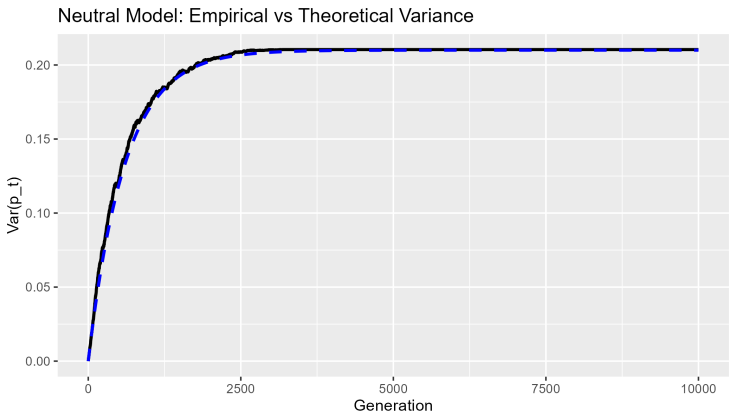

Figure 2: Empirical variance of p_t (solid black) compared with the theoretical variance (blue dashed).

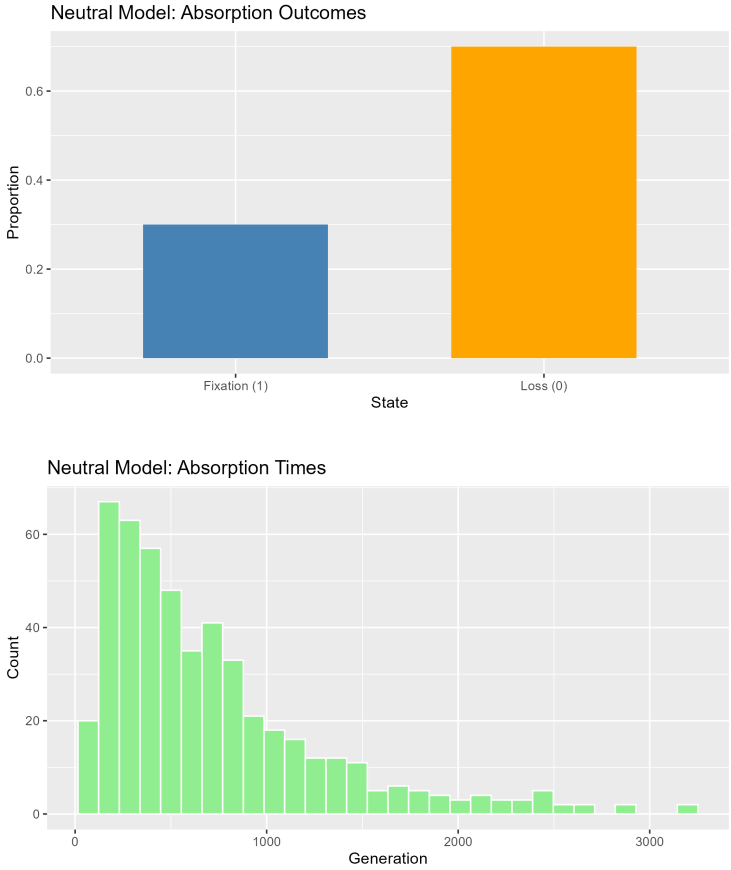

Figure 3: (Top) Proportion of trajectories absorbed at 0 or 1. (Bottom) Histogram of absorption times (generations). Empirical proportions and mean fixation times align closely with theory.

3.3.2 Wright–Fisher Model with Mutation

Next, we introduce bidirectional mutation with rates $\mu = 0.002$ and $\nu = 0.0005$ in a population of $N = 1000$ individuals, starting at $p_0 = 0.65$. Figure 4 shows that, unlike the neutral case, trajectories never reach absorption at 0 or 1: mutation continuously reintroduces alleles, maintaining polymorphism indefinitely.

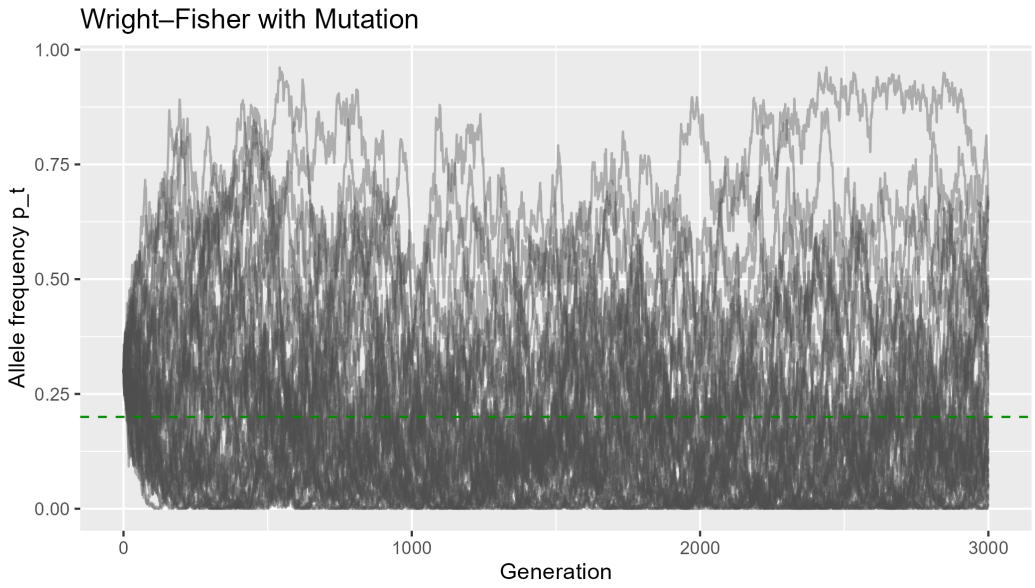

Figure 4: Wright–Fisher trajectories with mutation ($N = 1000$, $\mu = 0.002$, $\nu = 0.0005$). The process never reaches absorption due to constant mutation pressure.

The process stabilizes around the equilibrium frequency

$$E[p_\infty] = \frac{\nu}{\mu + \nu} = 0.2,$$

with stationary variance

$$\mathrm{Var}(p_\infty) \approx \frac{E[p_\infty](1 - E[p_\infty])}{2N(\mu + \nu)} = 0.032 = 3.2\%.$$

The empirical long-run histogram of allele frequencies matched the theoretical Beta($2N\nu$, $2N\mu$) stationary distribution (Figure 5).

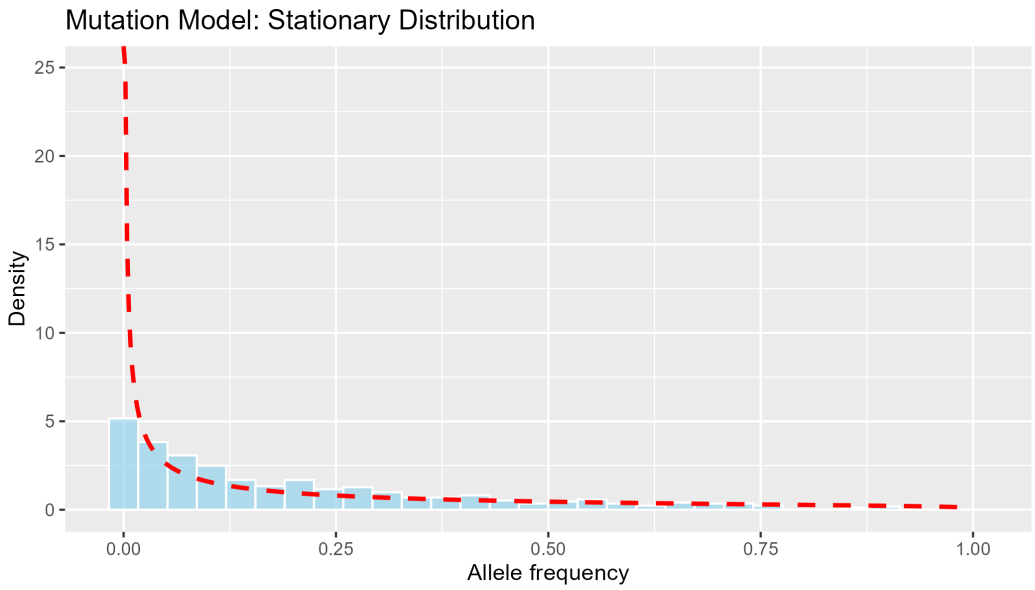

Figure 5: Histogram of stationary allele frequencies (blue) from long simulations under mutation. The red curve shows the theoretical Beta($2N\nu$, $2N\mu$) distribution, confirming the predicted steady state.

3.3.3 Summary and Interpretation

Under neutrality, allele frequencies fluctuate randomly due to genetic drift. Because no evolutionary forces act to favor either allele, these stochastic fluctuations eventually lead to absorption at fixation ($p_t = 1$) or loss ($p_t = 0$). The expected fixation probabilities are p_0 and $1 - p_0$, respectively, and the observed mean and variance dynamics in the simulations agree closely with these theoretical predictions.

When a mutation is introduced, the absorbing boundaries disappear, as both alleles can reappear through recurrent mutation. The process instead stabilizes around a stationary equilibrium

$$p^* = \frac{\nu}{\mu + \nu},$$

representing the long-term mean allele frequency maintained by the balance of mutation rates. The simulated stationary distribution matched the theoretical Beta($2N\nu$, $2N\mu$) form derived from diffusion theory, illustrating how mutation maintains polymorphism and prevents permanent fixation.

Together, these results demonstrate the two fundamental outcomes of the Wright–Fisher process: (1) pure drift leads to random fixation in finite populations, and (2) mutation introduces a steady-state balance that sustains genetic variation indefinitely. These simulations

validate the theoretical framework developed earlier and highlight the key roles of drift and mutation in shaping allele frequency dynamics.

4 Summary and Potential Future Research

The Wright–Fisher model can also be generalized to populations with multiple alleles. Rather than limiting our attention to two alleles, we can examine situations where several alleles occur at a given genetic locus.

In such cases, instead of monitoring one or two allele frequencies, we must track several simultaneously. This involves using a more complex matrix representation, but the fundamental ideas remain unchanged—we still study how allele frequencies evolve across generations under the effects of genetic drift, selection, and mutation.

While this paper focused primarily on neutral drift and mutation, future work should extend the analysis to include natural selection. Simulations with positive selection coefficients (e.g., $s = 0.05$) already demonstrate directional bias toward fixation. Further research could compare empirical fixation probabilities with theoretical diffusion results under selection, such as those derived from the Kimura equation.

Wright–Fisher Model with Selection. A useful next step is to include natural selection in the Wright–Fisher model. If one allele gives a fitness advantage of $1 + s$, then the average change in its frequency becomes

$$M(p) = s \cdot p(1 - p),$$

which means the allele is more likely to increase in the population. In the continuous-time (diffusion) version of the model, this selection factor changes the long-term (stationary) distribution of allele frequencies to

$$\phi_{\text{eq}}(p) \;\propto\; p^{2N\nu-1}(1-p)^{2N\mu-1}\, e^{2Nsp},$$

where N is population size, μ and ν are mutation rates, and s is the selective advantage. When $s > 0$, this "e^{2Nsp}" factor skews the distribution toward higher allele frequencies (closer to $p = 1$). Running simulations of this model would let us compare directly how drift, mutation, and selection each shape allele frequencies over time — and because the influence of selection depends strongly on N, this extension would give a deeper understanding beyond just neutral drift or mutation alone.

Supplemental Materials

All computational code and simulation data supporting this research are publicly accessible through a GitHub repository. Readers can find the complete collection of R scripts and datasets under the username tmccurley in the repository titled Population-Genetics-Classical-Wright-Fisher-Model-and-Its-Extensions-Repository. This resource enables readers to examine the technical implementation details, replicate the simulation experiments, and verify the statistical analyses presented throughout the paper. The repository contains code for simulating Wright-Fisher trajectories under neutral drift and mutation dynamics, computing theoretical and empirical variance measures, analyzing fixation probabilities and absorption times, and fitting stationary beta distributions to long-run allele frequency data as described in Section 3.

Acknowledgments

I would like to extend my greatest appreciation to Dr. Olga Korosteleva of California State University, Long Beach, whose guidance was instrumental both in strengthening my mathematical foundations and in helping me recognize meaningful applications of the Wright–Fisher model within modern statistical genetics. Her mentorship greatly shaped both the direction and the depth of this work.

I am also grateful for those who supported me in the earlier years of my academic life. It was through the UCLA math circle that Dr. Oleg Gleizer introduced me to this research opportunity and encouraged my exploration of advanced mathematical topics. I would also like to thank Mr. Chris Tillman, math team coach and teacher, whose love of mathematics and devotion to fostering problem-solving communities provided a perfect environment in which projects like this could succeed.

References

Crow, J. F. and M. Kimura. (1970). *An Introduction to Population Genetics Theory.* Harper & Row.

Durrett, R. (2008). *Probability Models for DNA Sequence Evolution.* Springer.

Etheridge, A. M. (2011). *Some Mathematical Models from Population Genetics.* Springer.

Ewens, W. J. (2012). *Mathematical Population Genetics 1: Theoretical Introduction*. Springer.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Gillespie, J. H. (2004). *Population Genetics: A Concise Guide*. Johns Hopkins University Press.

Kimura, M. (1964). Diffusion Models in Population Genetics. *Journal of Applied Probability*, 1(2), 177–232.

Wakeley, J. (2009). *Coalescent Theory: An Introduction*. Roberts & Company.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16(2), 97–159.