Article

Determination of Mutation Rates with Two Symmetric and Asymmetric Mutation Types

Jewgeni H. Dshalalow [1,*] , Van Minh Nguyen [1], Richard R. Sinden [2] and Ryan T. White [1]

1 Department of Mathematical Sciences, Florida Institute of Technology, College of Engineering and Science, Melbourne, FL 32901, USA
2 Department of Chemistry, Biology, and Health Sciences, South Dakota School of Mines and Technology, Rapid City, SD 57701, USA
* Correspondence: edshalalow@scienceatlantic.net

Abstract: We revisit our earlier paper, with two of the coauthors, in which we proposed an unbiased and consistent estimator $\hat{\mu}_n$ for an unknown mutation rate μ of microorganisms. Previously, we proved that the associated sequence of estimators $\{\hat{\mu}_n\}$ converges to μ almost surely pointwise on a nonextinct set Ω_0. Here, we show that this sequence converges also in the mean square with respect to conditional probability measure $\mathbb{P}_0(\cdot) = \mathbb{P}(\cdot \cap \Omega_0)/\mathbb{P}(\Omega_0)$ and that, with respect to $\mathbb{P}_0$, the estimator is asymptotically unbiased. We further assume that a microorganism can mutate or turn to a different variant of one of the two types. In particular, it can mean that bacteria under attack by a virus or chemical agent are either perishing or surviving, turning them to stronger variant. We propose estimators for their respective types and show that they are a.s. pointwise and L^2-consistent and asymptotically unbiased with respect to measure $\mathbb{P}_0$.

Keywords: Luria-Delbrück; mutation rate; bacterial mutation; stochastic estimator; mutagenesis; spontaneous mutation; bacterial genetic variations; symmetric and asymmetric mutation types

1. Introduction

In numerous biological applications related to the proliferation of microorganisms, such as bacteria, viruses, and even cancer cells, at some point of their division, they start mutating. New mutants are typically stronger and more virulent. In most common situations, bacteria divide in two forming a deterministic branching process. If a colony starts from a single organism, in the nth generation, its population size is 2^n (unless some progenies perish). Suppose at some point that one of its offspring mutates, say, with probability μ (which is typically small). Then, a new mutant divides like a nonmutant and it engenders a deterministic branching process. As the rate of reverse mutation on the same phenotype, called *revertant*, is extremely small [1,2], we do not consider this option in our paper. The mutation probability μ is called the *mutation rate*.

1.1. Our Approach Compared to Other Methods

The mutation rate is a common target in biomedical research and its applications. Ever since the seminal paper by Luria and Delbrück [3] in 1943, numerous works on determination of mutation rate have emerged (cf., [4–14]). They all have dealt with various estimates for μ.

The deterministic branching process (a *rooted tree*) contained a branching subprocess (random subtree) of non mutants that turned out to be a Galton–Watson process, although the target process was that of new mutants. The new mutants and pre-existing mutants form non-Markov processes.

Why do we deem the branching processes approach more rigorous? In Luria and Delbrück [3], the authors assumed that the number of mutations $M(\tau)$ that occur in interval $(t, t + \tau)$ equals $\mu \tau N_t$, where N_t is the total number of bacteria at time t. This assumption

is inaccurate, because to determine the mutation rate μ, they had to exclude from N_t the number of pre-existing mutants, even if this number is small. We actually prove that even if the mutation rate μ is small, the number of pre-existing mutants can converge to infinity and thus cannot be neglected. Moreover, $M(\tau)$ would be better off to equal some $\mu\tau 2X_t$, where X_t is the number of nonmutants at time t, which doubles in interval $(t, t + \tau)$ unless τ is too small. Stewart [15] also states that the Luria and Delbrück method is inaccurate.

Furthermore, if that would not be the case and N_t would be approximately equal to the number of nonmutants at time t, then the Luria and Delbrück method is still problematic. Here is why: suppose X_n is the number of nonmutants in generation n and suppose that each nonmutant can turn to a mutant with equal probability μ. Thus, each nonmutant is a Bernoulli r.v. with parameter μ. The number of newly born mutants NM_n in generation n is thus binomial with parameters (X_n, μ), that is, $NM_n \in [\mathbb{B}(X_n, \mu)]$. Because X_n is large and μ is small, the number NM_n is approximately Poisson, with parameter $\lambda_n = NM_n\mu$, which carries unknown NM_n and μ. The probability that there is no mutant at generation n would have been $p_0 = e^{-\lambda_n}$, which the authors claimed could be determined experimentally and is fairly simple. However, there is still a problem of finding NM_n, which is not equal to the total number cells N_n, nor does it equal the number of all mutants M_n at generation n. This is the reason why we think this method is unpredictably inaccurate (although Zheng in [16,17] claims that the Luria/Delbrück method is the best). Precisely, $\mu \approx NM_n/2X_{n-1}$ ($\neq M_n/N_n = M_n/2^n$ as the authors allude), provided that the latter is consistent in some sense and that n is sufficiently large. None of these requirements has ever been met in their work, nor in the work of their followers. This problem was also pointed out by Stewart [15].

Another problem in the statistical evaluation of the popular p_0-method is that in order to maintain a "nonmutant" plate (which is required to reasonably estimate p_0), it applies to counts rendered at early times or for cultures with extremely low mutation rates to warrant at least one plate with no mutants. However, this carries an adverse effect on statistical significance.

In various other alternatives, many formulas contain the ratio M_n/N_n as a major component in their estimators (rather $M_{n+k}/N_{n+k} - M_n/N_n$) which has a similar shortcoming and it cannot be salvaged by logarithmic or other multipliers often occurring in the literature. We noticed that no work offers a general formula for the determination of μ (whether practical or not). All heuristic methods may work in some special cases, but none are claimed to be general. In the literature on mutation rates, we see that some methods claimed to be efficient when following certain heuristic instructions imposed on μ (which was to be determined) and the mean number of mutants. It is not clear how the results obtained for μ were validated. We compared several tables of the results produced from different methods and they all vary.

In Niccum et al. [18], we offered the first rigorous method of determining μ by a pointwise rapidly convergent sequence $\{\hat{\mu}_n\}$ of estimators. At first, the original formula for $\hat{\mu}_n$ we operated on looked impractical because it contained newly formed mutants that are difficult to tell from pre-existing mutants, but it was modified by replacing it with an identical formula requiring just the knowledge of the mutants in any two consecutive generations that turned out to be practical in real labs. That "it was still difficult to observe" as stated in Zheng [17] was not the situation with a series of our lab experiments [18] performed by us and our students (who coauthored the paper) and beyond. It also needs to be noted that the proposed estimator had a rapid speed of convergence that has been validated in numerous subsequent simulations rendered by our students in senior undergraduate and graduate projects with no single exception.

1.2. Synchronized Cultures

In our original mathematical model [18], we assumed that bacteria replicate synchronously. We retain this mathematical assumption in the present work as well. As explained in [18], the real count in the lab is a little different. This method has been rendered for decades and the count of mutants (that is compatible with our theoretical assump-

tion) agrees with quite a few experiments made by different researchers, which obviously renders coincidences (and thus inaccuracies) probabilistically impossible.

Yet for the skeptics of our mathematical methodology, we would like to mention the common practice of *synchronization* of bacterial cultures known in the biological literature. First, a *synchronized culture* is a culture that contains cells in the same growth stage when they replicate, as opposed to asynchronous growth when the replication cycles are random. Most known cell divisions in a population are asynchronous. It is an observed, although unproven, assumption that random deviations of the cycle lengths are *symmetric* around the mean time of division. However, that needs to be rigorously validated. If this holds true, it is easy to deduce that a mere observations of a bacterial culture would give a pretty accurate identification of an underlying generation.

Alternatively, there are widely used synchronization approaches from 1957 (cf. Campbell [19]) and earlier and these continue to be described to the present time (cf. Chang et al. [20] of 2019). A popular technique, known as Helmstetter–Cummings [21,22] of 1968 and 2015 deals with an unsynchronized culture filtered through cellulose nitrate membrane filter. There are numerous other synchronization techniques by Anderson and Pettijohn [23], Kepes and Kepes [24,25], Kubitcheck [26], Ling and Chang [27], Noack, Klöden, and Bley [28], Shehata and Marr [29], and Wallden et al. [30], to name a few. A rigorous description of all those methods lies outside the scope of this paper.

Note that experimentally cells may not remain synchronized during growth complicating analysis. As mentioned above, it would be interesting to conduct lab experiments and learn about the statistical impact of unsynchronized replication. It seems that even if we assume the replication to be synchronous while in facts it is not, the deviations will be symmetric and not large, and as the result, the incorrect assumption will not corrupt the counts, because so far, the observed numbers were not conflicting with our experiments, and they agreed with experiments conducted in other labs. This conjecture is by no means intended to replace a rigorous investigation, which we plan to render.

1.3. Mutation or No Mutation?

There are many anecdotal stories about antibiotics causing antibiotic resistance and breeding new forms of virulent bacteria, in quite a few cases offering little to no remedy. One such notorious form is called *MRSA*, which stands for *methicillin-resistant Staphylococcus aureus* (*S. aureus*), a type of bacteria that has become resistant to potent antibiotics used to treat ordinary staph infections. It thus means that it is tougher to treat compared to other strains of *Staphylococcus*.

MRSA most commonly causes skin infections and can sometimes trigger pneumonia and other conditions. The symptoms largely depend on which part of the body is affected. Untreated, MRSA can become severe, leading to sepsis. MRSA infections are usually seen in individuals who were previously or currently in hospitals or other health care facilities such as dialysis centers or nursing homes. Many professionals are alarmed by the spread of MRSA as due to its resistance to many types of drugs; it is becoming known as a "super bug".

On the other hand, there is a general consensus that while bacteria mutate, this process is neither impacted nor exacerbated by the use of antibiotics or other chemical agents. They claim that mutations occur periodically without any external force. However, that natural selection allows these to thrive over other bugs that do not have resistance. (It must be noted that antibiotics are natural products of other organisms that also try to keep bacteria away.)

Before we elaborate it, we wonder what a *mutation* is referred to in cellular/molecular biology. A *mutation* is formally defined as an alteration to the gene sequence of an organism. A "natural mutation" occurs spontaneously and can happen during bacterial DNA replication where enzymes choose a wrong base to pair with the original DNA, or due to the environmental UV light, modifying the structure of the original DNA strand that allows it to pair with bases that are not normally compatible.

While the entire bacterial genome is susceptible to mutation, certain mutations are more easily detected than others because of selection. An example is the *trpE* gene in the *Escherichia coli* (*E. coli*) WP2 strain [31]. This high mutation region allows researchers to detect mutation on the strain using the so-called *reverse mutation assay*, in which the bacteria are grown in an amino-acid-deficient culture to test if the mutation occurred (trp^+) or not (trp^-) [32].

The natural mutation rate for *E. coli* in a rich medium mutation rate is 10^{-3} per genome per generation [33]. The mutation rate of *methicillin-resistant Staphylococcus aureus* (*S. aureus*)—MRSA strain is similar to the reported mutation rate of wild-type *E. coli* [34]. The elevated mutation rate is suspected to cause the increased rate of bacteria vancomycin-resistant genes, especially for the vancomycin-intermediate-resistant *S. aureus* strains—VRSA [35].

1.4. Other Means of Genetic Change

Other means of genetic change exist, in addition to classical spontaneous or genotoxicant-induced mutation, and, in some cases, these resulting genetic changes may also be covered by our model in as much as they may represent stochastic events. While the mechanisms responsible for these changes are well understood and these processes are not considered as classical mutations, they do result in genetic change in the affected organism. In some cases, the probabilities of these events are rare and, as such, the resulting genetic change may be described by probabilistic models. For the purposes of a more comprehensive discussion of other modes genetic changes, these alternative modes of change will be briefly described.

Transformation. Transformation is a process in which a bacteria takes up exogenous (free) DNA from its surrounding environment. This DNA can come from dead (lysed) bacteria. Transformation can happen artificially (such as a heat shock/electroporation), or naturally. A natural transformation occurs when there is a cooperative expression of multiple genes under specific conditions called competence [36]. This state is induced by a limited nutrition (usually amino-acid-deficient environment), associated with the stationary phase. For some other bacteria, transformation occurs most efficiently at the end of exponential growth approaching the stationary phase.

Transformation is one process that can lead to *Horizontal Gene Transfer* (HGT). There is a hypothesis about the HGT facilitated by the transformation of antibiotic-induced cell-wall-deficient bacteria [37]. This hypothesis states that before the bacteria die, they generate a cell-wall-deficient form and increase the uptake rate of exogenous DNA, because the DNA from dying bacteria release DNA. They are not doing the uptake as they are dead and could not then restore life to propagate this. Combining with the fact that patients who were treated with diverse antibiotics are more prone to get infected with MRSA makes this hypothesis more credible.

In this process, two bacterial cells attach, and plasmids and free intracellular DNA fragments travel from the donor cell to the target cell. In the target cell, plasmids can begin to replicate, whereas DNA fragments may become incorporated in the target cell's genome. A study on MRSA conjugation stated that this mechanism could be one of the pathways by which methicillin resistance is transmitted among *S. aureus* strains [38].

Transduction. Another mechanism for bacteria genome variation is called *transduction*. When bacteria are infected with a virus (usually a bacteriophage), the virus uses the cell's materials to produce new viruses and kill the bacteria when it reaches some threshold. During the production of new viruses, some fragments of the original bacteria's DNA are taken up into the virus' DNA. When this virus infects other cells, it also introduces the DNA to the new cell. If the infection goes to the latent phase and the fragments contain antibiotic genes, the bacteria become antibiotic-resistant [39]. Virus latency (or viral latency) is the ability of a pathogenic virus to lie dormant (latent) within a cell, denoted as the lysogenic part of the viral life cycle. If virus becomes latent, it neither reproduces nor kills the host cell.

So, as we see it, there is a variety of genetic alterations to the bacteria that are technically not mutations, but they often produce virulent species. While it is a common conjecture that bacteria do not mutate under chemical agents such as antibiotics, the latter does cause formidable DNA changes turning some strains to dangerous forms that can be even life-threatening. Most of these transformations are detectable and they can be observed and analyzed mathematically to determine the odds of their changes whether or not we call them *mutations*. Therefore, it stands to reason to consider those phenomena and place them in the same category as mutations and apply the same mathematical tools to establish rigorous estimators of their genetic transformations. It is also possible that an epigenetic change can occur that may be detectable by selection.

1.5. The Results

In the present paper, we show that the estimator proposed in [18] is not only almost surely pointwise consistent but also consistent in the L^2-norm giving us yet another mathematical confirmation of the goodness of the proposed estimator. We also prove the convergence with respect to other probability measures. We further assume that a microorganism can mutate in one of the two types, which in particular, can model the spread of virulent antibiotics-resistant bacteria (referred to in the media as "superbugs") that evolve under attacks by chemical agents so that they either die or mutate to survive the attacks. Consequently, mutated or altered bacteria's variants that survive are referred to as *type 1 mutants* and other mutants are called *type 2 mutants*. We also propose estimators for either type of mutants. These estimators turn out to be unbiased and consistent in every common sense, that is, almost surely pointwise and in the L^2-norm. (Some other forms of consistency are discussed throughout the forthcoming sections.)

We note that while death of bacteria is present in most bacterial cultures, it is rarely used in mathematical modeling. We address this matter in the present article and even generalize our model allowing replicated offspring to turn to one of the two mutant types with probabilities p and $1 - p$, respectively. It makes sense to call such division *symmetric* when $p = \frac{1}{2}$. In the event that this represents bacterial death, that, as mentioned, takes place under the use of a chemical agent (such as antibiotic) or bacteriophage attack.

2. Types of Stochastic Convergence

To make our paper self-contained, we present a short background on types of convergence of random variables (r.v.'s) related to the consistency of our estimators. (Cf. Dshalalow [40].)

Definition 1 (Types of convergence for a sequence of r.v.'s)**.** *Let* $Z, Z_1, Z_2, \ldots$ *be a sequence of r.v. on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *We say that the sequence* $\{Z_n\}$ *converges to r.v.* Z

(i)　　　*in probability* $\left(\text{in notation } Z_n \xrightarrow{\mathbb{P}} Z\right)$ *if*

$$\lim_{n \to \infty} \mathbb{P}\{|Z_n - Z| > \varepsilon\} = 0, \text{ for each } \varepsilon > 0,$$

(ii)　　　*in the p-norm (or in the L^p-norm or in the p-th mean)* $\left(\text{in notation } Z_n \xrightarrow{L^p} Z\right)$ *if*

$$\lim_{n \to \infty} \mathbb{E}[(Z_n - Z)^p] = 0, p \geq 1,$$

(iii)　　　*in the mean (or in the L^1-norm)* $\left(\text{in notation } Z_n \xrightarrow{L^1} Z\right)$ *if*

$$\lim_{n \to \infty} \mathbb{E}[Z_n - Z] = 0,$$

(iv)   *almost surely (a.s.) (or $\mathbb{P}$-almost surely) $\left(\text{in notation } Z_n \xrightarrow{a.s.} Z\right)$ if*

$$\mathbb{P}\left\{\omega \in \Omega : \lim_{n\to\infty} |Z_n(\omega) - Z(\omega)| = 0\right\} = 1.$$

The latter means that Z_n converges to Z pointwise for almost all $\omega \in \Omega$.

Theorem 1. *The following relations hold true:*

$$Z_n \xrightarrow{a.s.} Z \Rightarrow Z_n \xrightarrow{\mathbb{P}} Z \quad and \quad Z_n \xrightarrow{L^2} Z \Rightarrow Z_n \xrightarrow{L^1} Z \Rightarrow Z_n \xrightarrow{\mathbb{P}} Z$$

Definition 2. *If $\hat{\theta}_n$ is an estimator of some parameter θ, then $\hat{\theta}_n$ is a statistic (Borel measurable function) of a random sample $X_1, \ldots, X_n \in [X]$. In general, suppose $X_1, X_2, \ldots$ is a sequence of r.v.'s on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$ adapted to $(\mathcal{F}_n)$. Notice that the r.v.'s $X_1, X_2, \ldots$ need not be independent nor identically distributed. Let*

$$\hat{\theta}_n := \delta(X_1, \ldots, X_n)$$

be a statistic of the sample of the first n r.v.'s. Then, $\{\hat{\theta}_n\}$ is a sequence of estimators associated with $\{X_n\}$ and induced by a Borel measurable function δ. Let θ be a real number referred to as a parameter of $\{X_n\}$.

(i)   *The sequence $\{\hat{\theta}_n\}$ is unbiased if $\mathbb{E}\hat{\theta}_n = \theta$ holds for all n.*

(ii)   *The sequence $\{\hat{\theta}_n\}$ is asymptotically unbiased if $\lim_{n\to\infty} \mathbb{E}\hat{\theta}_n = \theta$*

(iii)   *The sequence $\{\hat{\theta}_n\}$ is consistent*

(a)   *in probability if $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$*

(b)   *in the pth mean (or in the L^p-norm) if $\hat{\theta}_n \xrightarrow{L^p} \theta$ for some $p \geq 1$*

(c)   *almost surely if $\hat{\theta}_n \xrightarrow{a.s.} \theta$*

We say that $\hat{\theta}_n$ is a consistent estimator in probability, pth mean, or almost surely, respectively, if so is the sequence $\{\hat{\theta}_n\}$.

3. A Background on Stochastic Estimators of Mutation Rate

Consider a population of microorganisms, such as bacteria, stemming from a single parent. Assume that the bacteria replicate by the division in exactly two progeny. At generation n, there are exactly 2^n bacteria in the population forming a deterministic branching process, provided that the bacteria do not die. Suppose that a generation began with a single parent bacterium (at generation zero) and beginning from generation 1, each bacteria, independently of the others, mutates with probability $\mu \in \left(0, \frac{1}{2}\right)$. The parent bacterium of the generation is thus assumed to be a nonmutant organism. Let X_n be the total number of nonmutants at generation n, NM_n be the total number of new mutants at generation n, and M_n be the total number of all (preexisting and new) mutants at generation n. We assume that μ (referred to as the *mutation rate*) is constant at every generation. To determine μ experimentally, we proposed in [18] the estimator

$$\hat{\mu}_{n+1} = NM_{n+1}/2X_n = (2X_n - X_{n+1})/2X_n = 1 - \frac{X_{n+1}}{2X_n} \tag{1}$$

which turned out to be unbiased and consistent a.s. on the non extinction set Ω_0 (defined below) such that $\mathbb{P}(\Omega_0) = 1 - \mu^2/\nu^2$, where $\nu = 1 - \mu$. (For various bacteria, such as *E. coli*, this probability is extremely close to 1.)

Denote

$$C_n = \{\omega : X_n(\omega) > 0\} \text{ and } \Omega_0 = \bigcap_{n=1}^{\infty} C_n. \tag{2}$$

Since the sequence $\{C_n\}$ is monotone decreasing, the notation Ω_0 in (2) makes sense and Ω_0 is referred to as the *nonextinction set*. As mentioned above, $\mathbb{P}(\Omega_0) = 1 - \frac{\mu^2}{\nu^2} = 1 - \phi$.

In light of our next applications, it is more convenient to turn to the confined space $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$, where

$$\mathbb{P}_0(\cdot) = \mathbb{P}(\cdot \cap \Omega_0)/\mathbb{P}(\Omega_0) = \mathbb{P}(\cdot|\Omega_0). \tag{3}$$

If $X : \Omega_0 \to \overline{\mathbb{R}}$ is a r.v. on the confined space $(\Omega_0, \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$, then (cf., Dshalalow [41]), the expected value $\mathbb{E}_0$ (with respect to measure $\mathbb{P}_0$) is

$$\mathbb{E}_0 X = \int X d\mathbb{P}_0 = \frac{1}{\mathbb{P}(\Omega_0)} \int_{\Omega_0} X d\mathbb{P} = \frac{1}{\mathbb{P}(\Omega_0)} \mathbb{E}[\mathbf{1}_{\Omega_0} X]. \tag{4}$$

4. Preliminaries

Let $W_n = \frac{X_n}{(2\nu)^n}, n = 0, 1, \ldots$, be the normalized sequence of the respective nonmutants populations. This sequence is a martingale that almost surely on Ω_0 converges to a r.v. W_∞ (cf. [42]). Furthermore, Theorem 2 [Section 6, Chapter I] of [42] reveals the nature of this r.v., also stating that the convergence of W_n to W_∞ is also in the mean square.

Note that Theorem 2 below is borrowed from [42] originally formulated and proved as Theorem 2 . It could be more rigorously crafted under the specification of the confined probability space $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$ introduced in (3) and (4). We proceed with those key assumptions throughout the rest of the paper.

Theorem 2 ([42]). *If $\nu > \frac{1}{2}$, then*

(i)　　　$\|W_n - W_\infty\|_{L^2} \to 0$.

(ii)　　$\mathbb{E}W_\infty = 1$, $VarW_\infty = \frac{1}{2\nu} \cdot \frac{\sigma^2}{2\nu-1} = \frac{\mu}{\nu-\mu}$.

(iii)　　$\mathbb{P}\{W_\infty = 0\} = \phi = 1 - \mathbb{P}(\Omega_0)$.

From Theorem 2 *(iii)*, we have that $\mathbb{P}_0\{W_\infty > 0\} = 1$.

Assume that $\nu > \frac{1}{2}$. Since $W_n \to W_\infty : \Omega \to \{0\} \cup (0, \infty]$ with the distribution ϕ and $1 - \phi$, respectively, the sequence $\{X_n\}$ converges to

$$X_\infty = \lim_{n \to \infty} W_n(2\nu)^n$$

where X_∞ is defined on Ω and valued in $\{0\} \cup \{\infty\}$, with the distribution ϕ and $1 - \phi = \mathbb{P}(\Omega_0)$, respectively. Therefore, on the confined probability space, $(\Omega_0, \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$, X_∞ equals ∞ $\mathbb{P}_0$-a.s. Hence, the r.v. $\frac{1}{X_\infty}$ on $(\Omega_0, \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$ equals 0, $\mathbb{P}_0$-a.s., which implies that

$$\mathbb{E}_0\left[\frac{1}{X_\infty}\right] = 0 \tag{5}$$

Now, because $X_n \geq 1$ on $C_n = \{X_n > 0\}$, $\frac{1}{X_n} \leq 1$ on C_n. Therefore,

$$\frac{1}{X_n}\mathbf{1}_{C_n} \leq \mathbf{1}_{C_n}.$$

Because $\mathbf{1}_{C_n}$ is $L^1(\mathbb{P})$-integrable, using the Lebesgue dominated convergence theorem (LDCT),

$$\lim_{n \to \infty} \mathbb{E}\left[\frac{1}{X_n}\mathbf{1}_{C_n}\right] = \mathbb{E}\left[\lim_{n\to\infty}\frac{1}{X_n}\mathbf{1}_{C_n}\right] \qquad = \mathbb{E}\left[\lim_{n\to\infty}\frac{1}{X_n} \cdot \lim_{n\to\infty}\mathbf{1}_{C_n}\right]$$

$$= \mathbb{E}\left[\frac{1}{X_\infty}\mathbf{1}_{\Omega_0}\right] = 0 \tag{6}$$

Thus, we proved the following lemma,

Lemma 1. $\mathbb{E}_0\left[\frac{1}{X_\infty}\right] = \mathbb{E}\left[\frac{1}{X_\infty}\mathbf{1}_{\Omega_0}\right] = 0$, *where* $\mathbb{E}_0$ *and* $\mathbb{E}$ *are the integrals on respective spaces* $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$ *and* $(\Omega, \mathcal{F}, \mathbb{P})$.

Next, recall [18] that NM_{n+1} is the number of new mutants at generation $n+1$ equal to

$$NM_{n+1} = 2X_n - X_{n+1}.$$

Thus,

$$\frac{NM_{n+1}}{(2\nu)^{n+1}} = \frac{1}{\nu}W_n - W_{n+1} \to \frac{\mu}{\nu}W_\infty \text{ a.s. pointwise.}$$

Similarly, because $NM_{n+1} \geq 1$ a.s. on $A_{n+1} = \{NM_{n+1} > 0\}$, it follows that $\frac{1}{NM_{n+1}} \leq 1$ on A_{n+1} that implies

$$\frac{1}{NM_{n+1}}\mathbf{1}_{A_{n+1}} \leq \mathbf{1}_{A_{n+1}}.$$

By the LDCT and because $A_{n+1} \subseteq C_n$,

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{NM_{n+1}}\mathbf{1}_{A_{n+1}}\right] = \mathbb{E}\left[\lim_{n\to\infty}\frac{1}{NM_{n+1}}\mathbf{1}_{A_{n+1}}\right] \leq \mathbb{E}\left[\lim_{n\to\infty}\frac{1}{NM_{n+1}}\mathbf{1}_{C_n}\right]$$

$$= \mathbb{E}\left[\frac{1}{NM_\infty}\mathbf{1}_{\Omega_0}\right] = 0. \qquad (7)$$

Using the same arguments as in the proof of Lemma 1, we easily conclude that

$$\mathbb{E}_0\left[\frac{1}{NM_\infty}\right] = 0.$$

In a nutshell,

Lemma 2. $\mathbb{E}_0\left[\frac{1}{NM_\infty}\right] = \mathbb{E}\left[\frac{1}{NM_\infty}\mathbf{1}_{\Omega_0}\right] = 0$, *where* $\mathbb{E}_0$ *and* $\mathbb{E}$ *are the integrals on respective spaces* $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$ *and* $(\Omega, \mathcal{F}, \mathbb{P})$.

5. The $L^2(\mathbb{P}_0)$-Convergence of $\hat{\mu}_{n+1}$

Theorem 3. *Let* $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$ *be the confined space introduced in Section 3 and* $(\Omega, \mathcal{F}, \mathbb{P})$ *be the original space. The estimator* $\hat{\mu}_{n+1}$ *is* $\mathbb{P}_0$-*asymptotically unbiased and it converges to* μ *in the* $L^2(\mathbb{P}_0)$-*norm.*

Proof. We use the formula

$$\text{Var}_0\hat{\mu}_{n+1} = \mathbb{E}_0\hat{\mu}_{n+1}^2 - [\mathbb{E}_0\hat{\mu}_{n+1}]^2 = \frac{1}{\mathbb{P}(\Omega_0)}\mathbb{E}\left[\hat{\mu}_{n+1}^2\mathbf{1}_{\Omega_0}\right] - \left\{\frac{1}{\mathbb{P}(\Omega_0)}\mathbb{E}\left[\hat{\mu}_{n+1}\mathbf{1}_{\Omega_0}\right]\right\}^2. \quad (8)$$

(i)　　$\hat{\mu}_{n+1}\mathbf{1}_{\Omega_0} \leq \hat{\mu}_{n+1}$ (dominating sequence that is $\mathbb{P}$-integrable with $\int \hat{\mu}_{n+1}d\mathbb{P} = \mathbb{E}\hat{\mu}_{n+1} = \mu$). Thus, by the LDCT,

$$\lim_{n\to\infty} \mathbb{E}\left[\hat{\mu}_{n+1}\mathbf{1}_{\Omega_0}\right] = \mathbb{E}\left[\lim_{n\to\infty}\hat{\mu}_{n+1}\mathbf{1}_{\Omega_0}\right] = \mu\mathbb{P}(\Omega_0).$$

implying that

$$\mathbb{E}_0\hat{\mu}_{n+1} \to \mu \qquad (9)$$

which means that the estimator $\hat{\mu}_{n+1}$ is $\mathbb{P}_0$-asymptotically unbiased.

(ii) Next, we show that $\hat{\mu}_{n+1}$ is $L^2(\mathbb{P}_0)$-integrable. First, from (1),

$$\mathbb{E}\left[\hat{\mu}_{n+1}^2 \Big| X_n\right] = \mathbb{E}\left[\left(\frac{NM_{n+1}}{2X_n}\right)^2 \Big| X_n\right] = \frac{1}{4X_n^2}\mathbb{E}\left[NM_{n+1}^2 \Big| X_n\right].$$

Given the σ-algebra $\sigma(X_n)$, the r.v. $NM_{n+1} \in [\mathbb{B}(2X_n, \mu)]$, i.e., NM_{n+1} is conditionally binomial with parameters $2X_n$ and μ. Thus,

$$\mathbb{E}\left[\hat{\mu}_{n+1}^2 \Big| X_n\right] = \frac{1}{4X_n^2}\left[2X_n\mu\nu + 4X_n^2\mu^2\right] = \frac{1}{X_n}\frac{1}{2}\mu\nu + \mu^2$$

implying that

$$\mathbb{E}\left[\hat{\mu}_{n+1}^2 \mathbf{1}_{C_n} | X_n\right] = \frac{1}{X_n}\frac{1}{2}\mu\nu\mathbf{1}_{C_n} + \mu^2\mathbf{1}_{C_n}. \qquad (10)$$

and that

$$\mathbb{E}\left[\hat{\mu}_{n+1}^2 \mathbf{1}_{C_n}\right] = \frac{1}{2}\mu\nu\mathbb{E}\left[\frac{1}{X_n}\mathbf{1}_{C_n}\right] + \mu^2\mathbb{P}(C_n).$$

From (6) and (10), we have

$$\mathbb{E}\left[\hat{\mu}_{n+1}^2 \mathbf{1}_{C_n}\right] \to \mu^2\mathbb{P}(\Omega_0).$$

Since $\hat{\mu}_{n+1}^2\mathbf{1}_{\Omega_0} \le \hat{\mu}_{n+1}^2\mathbf{1}_{C_n}$ and since $\hat{\mu}_{n+1}^2\mathbf{1}_{C_n}$ is $\mathbb{P}$-integrable, so is $\hat{\mu}_{n+1}^2\mathbf{1}_{\Omega_0}$. Furthermore, because $\hat{\mu}_{n+1}^2 \le 1$ and because $\hat{\mu}_{n+1}^2\mathbf{1}_{\Omega_0} \to \mu^2\mathbf{1}_{\Omega_0}$ a.s., by the LDCT,

$$\mathbb{E}\left[\hat{\mu}_{n+1}^2\mathbf{1}_{\Omega_0}\right] \to \mu^2\mathbb{P}(\Omega_0)$$

implying that

$$\mathbb{E}_0\,\hat{\mu}_{n+1}^2 = \frac{1}{\mathbb{P}(\Omega_0)}\mathbb{E}\left[\hat{\mu}_{n+1}^2\mathbf{1}_{\Omega_0}\right] \to \mu^2. \qquad (11)$$

(iii) Using (9) and (11),

$$\lim_{n\to\infty}\|\hat{\mu}_{n+1} - \mathbb{E}_0\hat{\mu}_{n+1}\|_{L^2(\mathbb{P}_0)}^2 = \lim_{n\to\infty}\text{Var}_0\hat{\mu}_{n+1}$$
$$= \lim_{n\to\infty}\left\{\mathbb{E}_0\left[\hat{\mu}_{n+1}^2\right] - (\mathbb{E}_0\hat{\mu}_{n+1})^2\right\} = \mu^2 - \mu^2 = 0 \qquad (12)$$

or

$$= \frac{1}{\mathbb{P}(\Omega_0)}\mu^2\mathbb{P}(\Omega_0) - \frac{1}{\mathbb{P}^2(\Omega_0)}\mu^2\mathbb{P}^2(\Omega_0) = 0.$$

(iv) Notice that while $\hat{\mu}_{n+1}$ is an unbiased estimator of μ with respect to measure $\mathbb{P}$, it need not be unbiased with respect to measure $\mathbb{P}_0$. As mentioned, however, $\hat{\mu}_{n+1}$ is asymptotically unbiased with respect measure $\mathbb{P}_0$. To establish the $L^2(\mathbb{P}_0)$-convergence of $\hat{\mu}_{n+1}$ to μ we use (9) and (12):

$$\|\hat{\mu}_{n+1} - \mu\|_{L^2(\mathbb{P}_0)} \le \|\hat{\mu}_{n+1} - \mathbb{E}_0\hat{\mu}_{n+1}\|_{L^2(\mathbb{P}_0)} + \|\mathbb{E}_0\hat{\mu}_{n+1} - \mu\|_{L^2(\mathbb{P}_0)}$$
$$= \sqrt{\text{Var}_0\hat{\mu}_{n+1}} + |\mathbb{E}_0\hat{\mu}_{n+1} - \mu| \to 0 \text{ as } n \to \infty.$$

Therefore, we proved that $\hat{\mu}_{n+1} \xrightarrow{L_2(\mathbb{P}_0)} \mu$. $\quad\square$

6. The Speed of Convergence

In [18], we validated the pointwise convergence of estimator $\hat{\mu}_{n+1}$ to μ by simulation and most importantly noticed that the speed of convergence was fast. Considering the

replication time of various bacteria (20–40 min), it generally takes not more than 12 to 15 h of lab experiments (or 25 to 30 generations) to attain to a high accuracy approximation of μ. We wondered if similar qualities of $\hat{\mu}_{n+1}$ apply to the L^2-convergence (mean square) in some form. The considerations below are carried out in the confined space $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$.

Let

$$\xi_i^{[n]} = \begin{cases} 1, & i\text{th non mutant in generation } n \text{ does not mutate} \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Then, $\eta_i^{[n]} = 1 - \xi_i^{[n]}$ is 1 if the ith nonmutant mutates and 0 otherwise.

We drop the superscript n in $\xi's$, and $\eta's$ assuming that they are iid Bernoulli r.v.'s with parameters ν and μ, respectively. Then,

$$X_{n+1} = \sum_{i=1}^{2X_n} \xi_i, n = 0, 1, \ldots, X_0 = 1$$

and

$$NM_{n+1} = 2X_n - X_{n+1} = 2X_n - \sum_{i=1}^{2X_n} \xi_i = \sum_{i=1}^{2X_n} \eta_i$$

implying that

$$\hat{\mu}_{n+1} = \frac{\sum\limits_{i=1}^{2X_n} \eta_i.}{2X_n}, \text{ where } \eta_i \in \mathbb{B}[1, \mu].$$

Setting $X_n = \mathfrak{X}$ (constant), $\hat{\mu}_{n+1}$ is the sample mean of $2\mathfrak{X}$ Bernoulli r.v.'s and thus with $\mathfrak{X}$ large, by the CLT (central limit theorem), $\hat{\mu}_{n+1}$ has a Gaussian distribution with parameters μ (mean) and $\mu\nu/2\mathfrak{X}$ (variance). Now, with n being just 10 or larger and μ small, $X_n = \mathfrak{X}$ must be a large number. Thus, with simulation, because X_n is obtained empirically it is a constant and as noticed $\hat{\mu}_{n+1} \in N[\mu, \mu\nu/2\mathfrak{X}]$, good for all $n \geq 10$. The variance $\mu\nu/\mathfrak{X}$ of $\hat{\mu}_{n+1}$ is small and must approach 0.

Suppose we "observe" bacterial replication on M plates where each colony starts with a single parent. Denote $\hat{\mu}_{n+1}^{[k]}$ the estimator associated with kth plate, $k = 1, \ldots, M$. For notational brevity, we will use the same symbol $\hat{\mu}_{n+1}^{[k]}$ for its estimate (i.e., observed value). Denote by $\bar{\mu}_{n+1}$ the sample mean of M estimators (estimates) $\hat{\mu}_{n+1}^{[k]}$, $k = 1, \ldots, M$. Thus,

$$\hat{\sigma}_{n+1}^2 = \frac{1}{M-1} \sum_{k=1}^{M} \left(\hat{\mu}_{n+1}^{[k]} - \bar{\mu}_{n+1} \right)^2$$

is the unbiased sample variance that is known to be consistent and weakly convergent to $\mu\nu/2\mathfrak{X}$ as $M \to \infty$. However, we expectthat $\hat{\sigma}_{n+1}$ rapidly converges to 0 with just reasonable n and M. This would be a heuristic feel of the true convergence speed of the variance of $\hat{\mu}_{n+1}$ (cf. Equation (12)).

Example 1. *To validate the results above, we will simulate the process for several fixed values of μ and compute the sample estimates of the mean and variance of μ, i.e., $\hat{\mu}_n$ and $\hat{\sigma}_n^2$, for generations $n = 1, 2, \ldots, 25$ for different numbers of plates ranging from $M = 1$ to $M = 20$. The goals are twofold: (1) to validate the accuracy of the estimators empirically and (2) to experimentally examine the convergence of the estimators for various choices of generations n and plates M. The estimators are more promising for practical use if convergence occurs for relatively small n and M.*

To compute estimates, we must simulate M paths of the process for each experiment. For each sample path, we start with $X_0 = 1$ nonmutant at the 0th generation and, for each generation $n \geq 1$, we assume the number of nonmutants is a binomial random variable $X_n \in \mathbb{B}(2X_{n-1}, 1 - \mu)$ for a preselected mutation rate μ. We generated M paths from $n = 1$

to $n = 30$. The estimated mutation rate is computed at each generation $n = 0$ to $n = 29$ as $\hat{\mu}_{n+1} = NM_{n+1}/2X_n$ (rewritten as $1 - X_{n+1}/2X_n$ for convenience) and the mean $\overline{\mu}_n$ is taken for each experiment with M plates. The sample variance will also computed. Ideally, the estimate $\hat{\mu}_n$ will be near μ and the estimate $\hat{\sigma}_n^2$ will approach 0 quickly for relatively small M.

We follow this approach in several examples. In Figure 1, we let $\mu = 0.01$ and simulated the experiment for 25 generations for each selection of plates $M \in \{2, 5, 10, 20\}$ (We did not use $M = 1$ as we evaluate sample variance with degree of freedom of 1). There are several intuitive occurrences here.
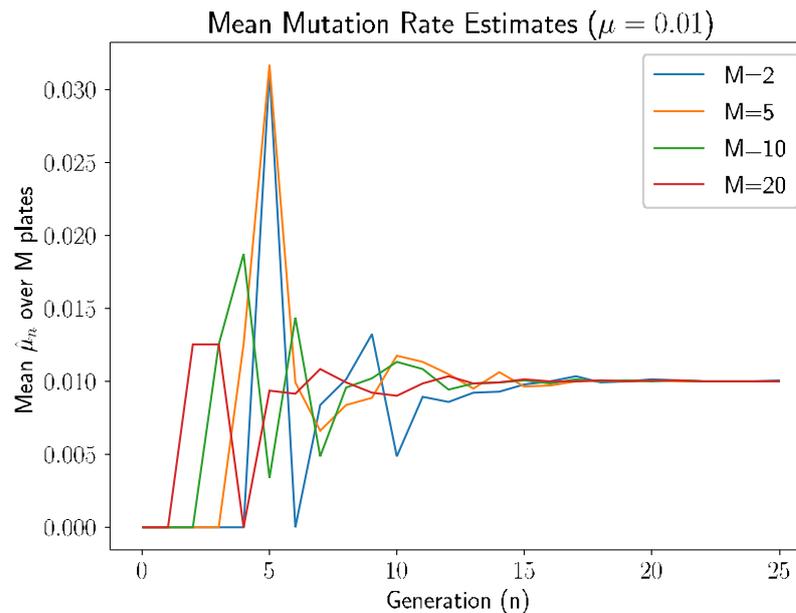
Figure 1. The mean mutation rate estimates of $\mu = 0.01$ for $M = 2, 5, 10, 20$ plates for generations $n = 1$ to 25.

Since $\hat{\mu}_{n+1} = NM_{n+1}/2X_n$, the estimate is 0 if there are no new mutants in generation $n + 1$. As we see in Figure 1, the estimates are indeed 0 before the first mutant appears in each experiment. An interesting pattern emerges here: for $M = 20$ and $M = 10$ plates, the first mutant appears in generation 1 and 2, respectively; for $M = 5$ plates, the first mutant appears at generation 3; and for $M = 2$, the first mutant appears at generation 4. The time of the first mutation is inversely related to the number of plates, which makes sense because there are more opportunities for mutations with more plates and, therefore, the estimate is nonzero sooner with more plates.

Second, the estimates behave erratically between the initial escape from 0 until roughly generation 7–8. During the early generations, there are smaller nonmutant populations in the plates, meaning intuitively that the random mutations have larger impacts on the estimates.

As such, variance is higher when the prior generation's nonmutant population is smaller, and this nonmutant population grows over time in expectation since the mutation rate $\mu \ll 0.05$, so we should expect to high variances in early generations. After 7 or 8 generations, the variance seems to dissipate enough to allow the estimates start to approach the true mutation rate $\mu = 0.01$.

Continuing beyond generation 8, estimates in all experiments approach $\mu = 0.01$, with the experiments with fewer plates, especially $M = 2$, taking a longer time to converge—again, this is expected as the conditional variance of $\overline{\mu}_{n+1}$ is inversely proportional to M^2, and a shrinking variance should result in convergence of the unbiased estimator.

Clearly, and unsurprisingly, the estimates have higher quality for larger generation n and larger number of plates M, but more interestingly, what does this mean for experimental use of the estimator? First, we define a modified relative error for the estimate

$$e_{0.01} = \max\left(\left|\frac{\overline{\mu}_n - \mu}{\mu}\right|, 0.01\right),$$

which is the relative error or 0.01, whichever is larger. Figure 2 is a heatmap of this modified relative error for each pair n and M from our simulations.
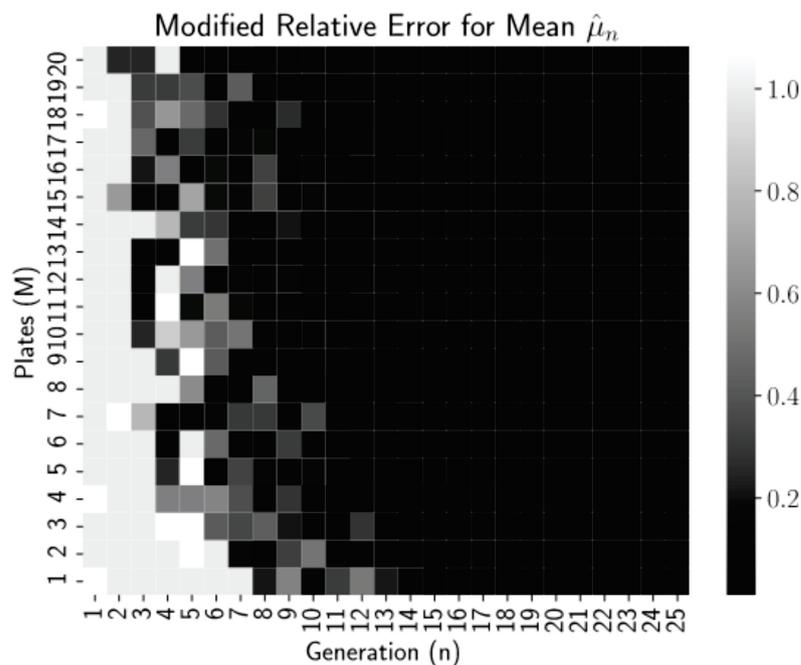
Figure 2. Heatmap of modified relative errors $e_{0.01}$ of mean mutation rate estimates.

Thresholding the relative error as $e_{0.01}$ allows us to see clearly in Figure 2 that relative errors of less than 0.01 occurred for all generations $n \geq 20$ regardless of the number of plates, but for more plates, near $M = 20$, these small errors occur reliably earlier, at generations $n < 15$. Thus, even low numbers of plates can be used if the experiment can continue for more generations. This is a promising finding for biology lab protocol, as growing more than 20–25 dishes per experiment is highly infeasible. After 15–20 generations (starting with 1 starting cells), we have the mean converged.

Beyond the behavior of the mean mutation rate themselves, we can consider the estimated variance of the mutation rate, $\hat{\sigma}_{n+1}^2$. Figure 3 shows the estimated sample variance for several experiments with $M = 2, 5, 10,$ and 20 plates.
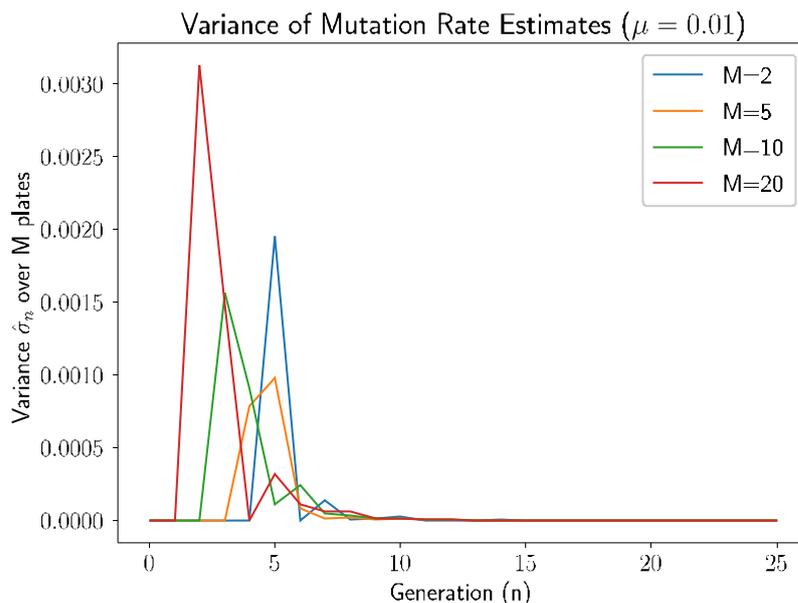
Figure 3. The estimated variance of mutation rate for $\mu = 0.01$, $M = 2, 5, 10, 20$ plates for generations $n = 1$ to 25.

As we see in Figure 3, the variance converges to 0 within 10–12 generations regardless of the number of plates, as prior analysis suggested should happen.

7. Estimator $\hat{p}_n$

We now assume that once a cell or microbe mutates, it turns into one of the two mutant types, with probabilities p and $q = 1 - p$, respectively. The type i mutant then divides in accordance with a deterministic branching process throughout the generations and it does not mutate any further, nor does it alter its type. In particular, we can also interpret mutation types as a continual existence upon mutation (type 1) with probability p or the death (type 2) with probability $q = 1 - p$. Note that our analysis in this section is interchangeably carried out in the original and confined probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0)$, respectively. In all cases, the underlying functionals can be distinguished by the absence or presence of subscript 0.

Denote by

$$\hat{p}_n = \frac{NM_n^{[1]}}{NM_n} \text{ and } \hat{q}_n = \frac{NM_n^{[2]}}{NM_n} \tag{14}$$

the proportions of type 1 and 2 new mutants at the nth generation from among the NM_n new mutants, respectively. We will show that the estimator $\hat{p}_n$ is unbiased and consistent (and so is $\hat{q}_n$).

Recall that $A_n = \{NM_n > 0\}$. Denote $g_n(z) = \mathbb{E}z^{X_n}$. Below, we observe an interesting and unexpected result concerning the sequence $\{A_n\}$. Whereas it obviously is not monotone, as the following two assertions show, the sequence $\{\mathbb{P}(A_n)\}$ of their measures under a weak and natural assumption, is strictly monotone increasing (we recollect that the sequence $\{\mathbb{P}(C_n)\}$ with respect to nonmutants is monotone nonincreasing) and furthermore, like $\{\mathbb{P}(C_n)\}$, $\{\mathbb{P}(A_n)\}$ converges to $\mathbb{P}(\Omega_0) = 1 - \phi$.

Theorem 4. *The following two assertions are valid.*

(i) $\mathbb{P}(A_n^c) = g_{n-1}(v^2)$.

(ii)   *If $\mu < v^2 < 1$, the sequence $\{\mathbb{P}(A_n)\}$ is strictly monotone increasing.*

Proof.

(i)　　　Let $h_n(z) = \mathbb{E}z^{NM_n}$. Then, because $NM_{n+1} \in [\mathbb{B}(2X_n, \mu)]$ given $\sigma(X_n)$,

$$h_{n+1}(z) = \mathbb{E}\left[\mathbb{E}\left[z^{NM_{n+1}}X_n\right]\right] = \mathbb{E}\left[(\mu z + v)^{2X_n}\right] = g_n\left[(\mu z + v)^2\right].$$
$$\Rightarrow \mathbb{P}\left(A_{n+1}^c\right) = h_{n+1}(0) = g_n\left(v^2\right).$$

(ii)　　　We show that $\left\{g_n\left(v^2\right)\right\}$ is strictly monotone decreasing.

First off, $p(z) = (vz + \mu)^2$ is strictly monotone increasing for $z \in [0, 1]$. Now since $v^2 < 1$ and $p(1) = \max\{p(z) : z \in [0, 1]\} = 1$,

$$p\left(v^2\right) = g_1\left(v^2\right) = \left(v^3 + \mu\right)^2 < 1.$$

Now, we show that

$$g_{n+1}\left(v^2\right) < g_n\left(v^2\right)$$

or that

$$g_{n+1}(v^2) = p\left[g_n\left(v^2\right)\right] < g_n\left(v^2\right).$$

Let $g_n\left(v^2\right) = z$. Consider $p(z) \diamond z$, where $\diamond$ stands for one of the three relations, $<, >,$ or $=$. We have

$$(vz + \mu)^2 - z \diamond 0$$

or

$$v^2 z^2 - (1 - 2v\mu)z + \mu^2 \diamond 0$$

The discriminant of the equation $v^2 z^2 - (1 - 2v\mu)z + \mu^2 = 0$ is

$$D = (1 - 2v\mu)^2 - 4v^2\mu^2 = 1 - 4\mu v = 1 + v^2 - 4v = (2v - 1)^2$$
$$= \begin{cases} = 0, & \mu = v = \frac{1}{2} \\ > 0, & v \neq \frac{1}{2} \end{cases}.$$

Furthermore, after some algebra,

$$v^2 z^2 - (1 - 2v\mu)z + \mu^2 = v^2(z - z_1)(z - z_2),$$

where $z_1 = \frac{\mu^2}{v^2}$ (is the extinction probability ϕ) and $z_2 = 1$. Therefore, $v^2 z^2 - (1 - 2v\mu)z + \mu^2 < 0$ (and $\diamond$ is $<$) for all $z \in \left(\frac{\mu^2}{v^2}, 1\right)$. In conclusion,

$$\text{if } z = v^2, \text{ then } v^2 > \frac{\mu^2}{v^2} \text{ whenever } v^2 > \mu.$$

Therefore, if $v^2 > \mu$, $p\left(v^2\right) = g_1\left(v^2\right) = \left(v^3 + \mu\right)^2 < g_0\left(v^2\right) = v^2$. Next, since $p(z)$ is strictly monotone increasing in $z \in [0, 1]$ and because $p\left(v^2\right) < v^2$, it follows that

$$g_2\left(v^2\right) = p\left[p\left(v^2\right)\right] < g_1\left(v^2\right) = p\left(v^2\right).$$

Finally, since $p(z)$ is strictly monotone increasing, by the principle of mathematical induction,

$$g_{n+1}\left(v^2\right) = \underbrace{p\left(p\left(p\dots\left(v^2\right)\right)\right)}_{n+1 \ p's} < g_n\left(v^2\right) = \underbrace{p\left(p\dots\left(v^2\right)\right)}_{n \ p's}$$

that proves the assertion that $\{g_n(\nu^2)\}$ is a strictly monotone decreasing sequence and so is the sequence $\{\mathbb{P}\{NM_n = 0\} = \mathbb{P}(A_n^c)\}$. $\square$

Corollary 1. $\lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \mathbb{P}\{NM_n > 0\} = 1 - \phi = \mathbb{P}(\Omega_0) = 1 - \frac{\mu^2}{\nu^2}$.

Proof. Let $\alpha_n = g_n(\nu^2)$. Then, from the equation

$$g_{n+1}(\nu^2) = p\left[g_n(\nu^2)\right] \text{ holding for } n = 0, 1, \dots,$$

we have

$$\alpha_{n+1} = p(\alpha_n), n = 0, 1, \dots \tag{15}$$

Furthermore, since $\{\alpha_n\}$ is monotone decreasing, $\alpha := \lim_{n\to\infty} \alpha_n$ exists and it is unique. Because $p(z)$ is continuous, this limit from (15) can be found from the equation

$$\alpha = p(\alpha).$$

From the theory of branching processes, we know (cf. Kimmel and Axelrod [43]) that the extinction probability ϕ is the smallest positive root of the equation $z = p(z)$. In our case, $p(z) = (\nu z + \mu)^2$ and the extinction probability $\phi < 1$ if the mean $2\nu > 1$ which is routinely met when $\mu < \frac{1}{2}$. The equation $z = (\nu z + \mu)^2$ has two roots: 1 and the other one $\phi = \frac{\mu^2}{\nu^2}$ which is the smallest of two. Consequently,

$$\alpha = \lim_{n\to\infty} \mathbb{P}\{NM_n = 0\} = \phi = \frac{\mu^2}{\nu^2}$$

ends up being equal to the extinction probability. Therefore,

$$\lim_{n\to\infty} \mathbb{P}\{NM_n > 0\} = \mathbb{P}(\Omega_0) = 1 - \frac{\mu^2}{\nu^2}.$$

$\square$

Since $\hat{p}_n = \frac{NM_n^{[1]}}{NM_n}$, given NM_n, $NM_n^{[1]} \in [\mathbb{B}(NM_n, p)]$ and therefore,

$$\mathbb{E}[\hat{p}_n | NM_n] = \frac{1}{NM_n}\mathbb{E}\left[NM_n^{[1]} \Big| NM_n\right] = \frac{1}{NM_n}NM_n p = p \;\Rightarrow\; \mathbb{E}\hat{p}_n = p.$$

Furthermore,

$$\mathbb{E}[\hat{p}_n \mathbf{1}_{A_n} NM_n] = p\mathbf{1}_{A_n} \;\Rightarrow\; \mathbb{E}\hat{p}_n \mathbf{1}_{A_n} = p\mathbb{P}(A_n)$$
$$\Rightarrow\; \mathbb{E}_n \hat{p}_n = p.$$

Consequently, $\hat{p}_n$ is $\mathbb{P}$- and $\mathbb{P}_n$-unbiased, where $\mathbb{P}_n(\cdot) = \mathbb{P}(\cdot \cap A_n)/\mathbb{P}(A_n)$.

Theorem 5. *If $\nu > \frac{1}{2}$, $\hat{p}_n$ converges to p in the $L^2(\mathbb{P}_n)$-norm, $L^2(\mathbb{P}_0)$-norm, and a.s. pointwise. Furthermore, $\hat{p}_n$ is asymptotically $\mathbb{P}_0$-unbiased.*

Proof.

(i) We start off with

$$\mathbb{E}\left[\hat{p}_n^2 \mathbf{1}_{A_n} \Big| NM_n\right] = \frac{1}{NM_n^2}\mathbf{1}_{A_n}\left[NM_n pq + NM_n^2 p^2\right] = \left[p^2 + \frac{1}{NM_n}pq\right]\mathbf{1}_{A_n} \tag{16}$$

implying

$$\mathbb{E}\left[\hat{p}_n^2 \mathbf{1}_{A_n}\right] = pq\mathbb{E}\left[\frac{1}{NM_n}\mathbf{1}_{A_n}\right] + p^2\mathbb{P}(A_n). \tag{17}$$

Because $\frac{1}{NM_n}\mathbf{1}_{A_n} \leq 1$, $\hat{p}_n^2\mathbf{1}_{A_n}$ is $\mathbb{P}$-integrable. By Lemma 2 (cf. Equation (7)),

$$\mathrm{Var}_n\hat{p}_n = pq\frac{1}{\mathbb{P}(A_n)}\mathbb{E}\left[\frac{1}{NM_n}\mathbf{1}_{A_n}\right] + p^2 - p^2 = pq\frac{1}{\mathbb{P}(A_n)}\mathbb{E}\left[\frac{1}{NM_n}\mathbf{1}_{A_n}\right] \to 0$$

Thus, because $\hat{p}_n$ is $\mathbb{P}_n$-unbiased,

$$\|\hat{p}_n - p\|_{L^2(\mathbb{P}_n)} = \|\hat{p}_n - \mathbb{E}_n\hat{p}_n\|_{L^2(\mathbb{P}_n)} = \sqrt{\mathrm{Var}_n\hat{p}_n} \to 0.$$

(ii) Now we show that
$$\hat{\mu}_{n+1}^{[1]} = NM_{n+1}^{[1]}/2X_n \to \mu p \text{ a.s.}$$

The latter is true due to the following: As per our assumption, every new mutation takes place with a constant probability μ. The new mutants form process $\{NM_n\}$. The estimator $\hat{\mu}_n$ of μ is consistent in the sense that $\hat{\mu}_n \to \mu$ a.s. pointwise. If a new mutant is formed from a nonmutant, in our present model, it becomes a type 1 new mutant with a constant probability p, i.e., a type 1 new mutant emerges from a nonmutant with probability $p\mu$. We then focus on X_n and $NM_{n+1}^{[1]}$ of nonmutants in generation n and type 1 new mutants in generation $n+1$, respectively. We can then reduce this model to the original one with $p\mu$ being a new mutation rate. Thus, we define the estimator $\hat{\mu}_{n+1}^{[1]}$ of $p\mu$ as $NM_{n+1}^{[1]}/2X_n$, and the latter must converge to $p\mu$ a.s. pointwise [18].
Consequently, from Theorem 3, $\hat{\mu}_{n+1}^{[1]}$ converges to $p\mu$ in the $L^2(\mathbb{P}_0)$-norm.

(iii) Now, we turn to $\hat{p}_{n+1}$ defined as $NM_{n+1}^{[1]}/NM_{n+1}$. Dividing the numerator and denominator by $2X_n$, we get

$$\hat{p}_{n+1} = \frac{NM_{n+1}^{[1]}/2X_n}{NM_{n+1}/2X_n} = \frac{\hat{\mu}_{n+1}^{[1]}}{\hat{\mu}_{n+1}} \to \frac{p\mu}{\mu} = p, \text{ a.s. on } \Omega_0.$$

More specifically, if the numerator converges to $p\mu$ pointwise on Ω_1 and the denominator converges pointwise to μ on Ω_2 such that $\mathbb{P}(\Omega_1) = \mathbb{P}(\Omega_2) = \mathbb{P}(\Omega_0)$,then $\hat{p}_{n+1} \to p$ pointwise on set $\Omega_1 \cap \Omega_2$ and clearly $\mathbb{P}(\Omega_1 \cap \Omega_2) = \mathbb{P}(\Omega_0)$. Thus, the convergence of $\hat{p}_n$ to p is a.s. pointwise on Ω_0.
Consequently,
$$\hat{p}_n^2 \to p^2 \text{ a.s. pointwise on } \Omega_0 \tag{18}$$

and

$$\hat{p}_n^2\mathbf{1}_{\Omega_0} \to p^2\mathbf{1}_{\Omega_0} \text{ a.s.}$$

Because $\hat{p}_n^2 \leq 1$ for all ω's, by the LDCT and Equation (18) ,

$$\lim_{n\to\infty}\mathbb{E}\hat{p}_n^2\mathbf{1}_{\Omega_0} = \mathbb{E}\left[\lim_{n\to\infty}\hat{p}_n^2\mathbf{1}_{\Omega_0}\right] = p^2\mathbb{P}(\Omega_0).$$

(iv) $\mathrm{Var}_0\hat{p}_n = \mathbb{E}_0\hat{p}_n^2 - [\mathbb{E}_0\hat{p}_n]^2 = \frac{1}{\mathbb{P}(\Omega_0)}\mathbb{E}\left[\hat{p}_n^2\mathbf{1}_{\Omega_0}\right] - \frac{1}{\mathbb{P}^2(\Omega_0)}\mathbb{E}^2\left[\hat{p}_n\mathbf{1}_{\Omega_0}\right]$
$\hat{p}_n\mathbf{1}_{\Omega_0} \leq \hat{p}_n$ and $\hat{p}_n$ is $\mathbb{P}$-integrable ($\int \hat{p}_n d\mathbb{P} = p$)
Thus by the LDCT,
$$\lim_{n\to\infty}\mathbb{E}\hat{p}_n\mathbf{1}_{\Omega_0} = p\mathbb{P}(\Omega_0)$$

and

$$\mathbb{E}_0\hat{p}_n \to p,$$

so that $\hat{p}_n$ is asymptotically $\mathbb{P}_0$-unbiased. Consequently,

$$\mathrm{Var}_0\hat{p}_n \to p^2 - p^2 = 0.$$

(v) $\|\hat{p}_n - p\|_{L^2(\mathbb{P}_0)} \leq \|\hat{p}_n - \mathbb{E}_0\hat{p}_n\|_{L^2(\mathbb{P}_0)} + \|\mathbb{E}_0\hat{p}_n - p\|_{L^2(\mathbb{P}_0)}$
$= \sqrt{\mathrm{Var}\hat{p}_n} + |\mathbb{E}_0\hat{p}_n - p| \to 0.$

$\square$

Remark 1.

(i)   *If the second mutation type of bacteria are dead, the formula $\hat{p}_{n+1} = \frac{NM_{n+1}^{[1]}}{NM_{n+1}}$ is impractical for a lab utility, since the dead cells are not counted. Here is a way out.*

First, notice that in this case, $NM_{n+1}^{[1]} = M_{n+1} - 2M_n$ (i.e., the same formula as for NM_{n+1} with only one mutation type). This is exactly because the dead bacteria cannot be counted. Next, we have

$$2X_n = X_{n+1} + NM_{n+1}^{[1]} + NM_{n+1}^{[2]}, \tag{19}$$

so that the "deficit" of the count is

$$NM_{n+1}^{[2]} = 2X_n - X_{n+1} - NM_{n+1}^{[1]} = 2X_n - X_{n+1} - M_{n+1} + 2M_n. \tag{20}$$

Therefore,

$$\hat{p}_{n+1} = \frac{NM_{n+1}^{[1]}}{NM_{n+1}} = \frac{NM_{n+1}^{[1]}}{NM_{n+1}^{[1]} + NM_{n+1}^{[2]}} = \frac{M_{n+1} - 2M_n}{2X_n - X_{n+1}}. \tag{21}$$

(ii)   *If the second mutation type are not dead bacteria, then $NM_{n+1}^{[1]} \neq M_{n+1} - 2M_n$, but*

$$NM_{n+1}^{[1]} = M_{n+1}^{[1]} - 2M_n^{[1]}. \tag{22}$$

However, Equation (19) still applies, whereas Equation (20) can be modified as

$$NM_{n+1}^{[2]} = 2X_n - X_{n+1} - NM_{n+1}^{[1]} = 2X_n - X_{n+1} - M_{n+1}^{[1]} + 2M_n^{[1]}. \tag{23}$$

Obviously, Equation (21) can be modified as follows

$$\hat{p}_{n+1} = \frac{M_{n+1}^{[1]} - 2M_n^{[1]}}{2X_n - X_{n+1}}, \tag{24}$$

where superscripts [1] and [2] stand for the first and second type mutants, respectively. In a nutshell, if type one bacteria die upon mutation, the formula for the estimator $\hat{p}_n$ of p is simple and it does not require one to distinguish the bacteria types. This is not the case with two surviving types, as per Equation (24).

8. Simulation of Estimators for Types 1 and 2

To validate the convergence and its speed, we render simulation analogous to that in Section 6. After generation of nonmutant paths similar to Section 6, we evaluate new mutant (all types) at each generation using $NM_{n+1} = 2X_n - X_{n+1}$. For each new mutant generation, we run a binomial random number generator with distribution $\mathbb{B}(NM_{n+1}, p)$ for mutation type 1, assuming type 2 is the rest of the mutation population with distribution $\mathbb{B}(NM_{n+1}, q = 1 - p)$, because $NM_{n+1} = NM_{n+1}^{[1]} + NM_{n+1}^{[2]}$. This new mutant type i generation is added to the current population of type i mutant (previous generation mutant doubled): $M_{n+1}^{[i]} = NM_{n+1}^{[i]} + 2 \cdot M_n^{[i]}$, with the assumption that $M_0^{[i]} = 0$ (initially, we have 0 mutants). The algorithm is looped for n generations and we store the result in a new $(n + 1) \times M$ matrix with first row being zero row. Type 1 mutation rate is estimated by Equation (21): $\hat{p}_{n+1} = \frac{M_{n+1}^{[1]} - 2M_n^{[1]}}{2X_n - X_{n+1}}$ and the output was an $n \times M$ matrix. Mean mutation

and variance of the mutation rate were evaluated using the sample variance equation in Section 6, knowing the convergence is established with Theorem 5.

For this simulation, we let the probability that a newly mutated bacterium is of type 1 be $p = 0.25$. Figure 4 is the simulation results—specifically, the mean of the type 1 mutation rate estimate.
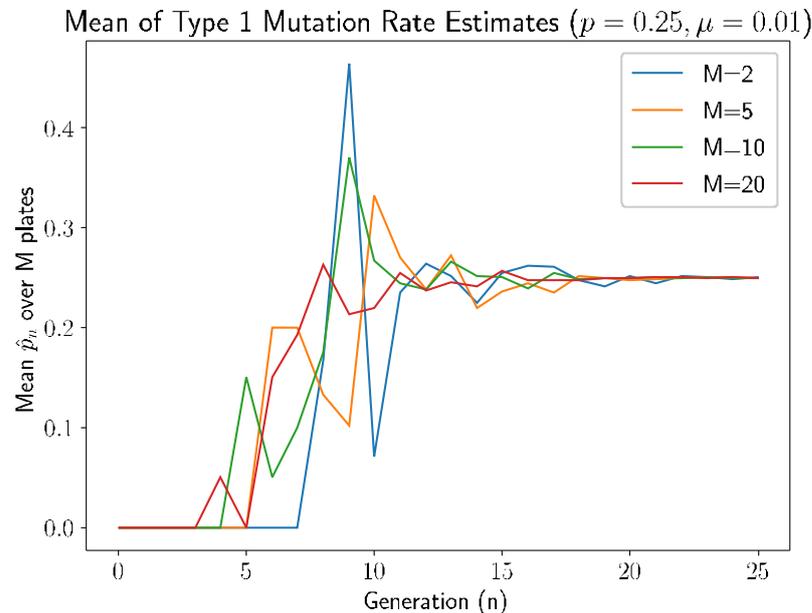
Figure 4. The mean of type 1 mutation rate estimates of $p = 0.25$ for $M = 2, 5, 10, 20$ plates for generations $n = 1$ to 25.

Similar to simulation in Section 6, we see a pattern here: the estimate is 0 if there are no new mutants. In this simulation, first mutation does not occur until generation 3 for $M = 20$, generation 4 for both $M = 10$ and $M = 5$, and generation 7 for $M = 2$. The time of the first type 1 mutant is also inversely related to the number of plates, due to the appearance of type 1 mutation depends on the time of first mutation, which is because of, as discussed previously, more opportunities for mutations to occur with more plates.

The estimates stabilize after 9–10 generations, which is a couple of generations lag after first mutation occurs, as established above. This can be explained due to the probability of getting type 2 mutation is larger than type 1 and, due to this type 1 mutation, take more generation to appear. The following modified relative error heatmap (Figure 5) also describes this phenomenon.

Figure 5 shows the similar pattern as Figure 2: for all generations $n \geq 20$, regardless of the number of plates, the relative error is less than 0.01; the more plates there are, the less occurrence of large relative errors in earlier generations. For all number of plates, the threshold for small relative error is $n \geq 16$; which is 1 to 2 generation behind the observed threshold from Figure 2 ($n \geq 14$). Similar to Section 6, this result is a promising finding for biology researchers as the number of dishes is feasibly small to be made in laboratory settings.
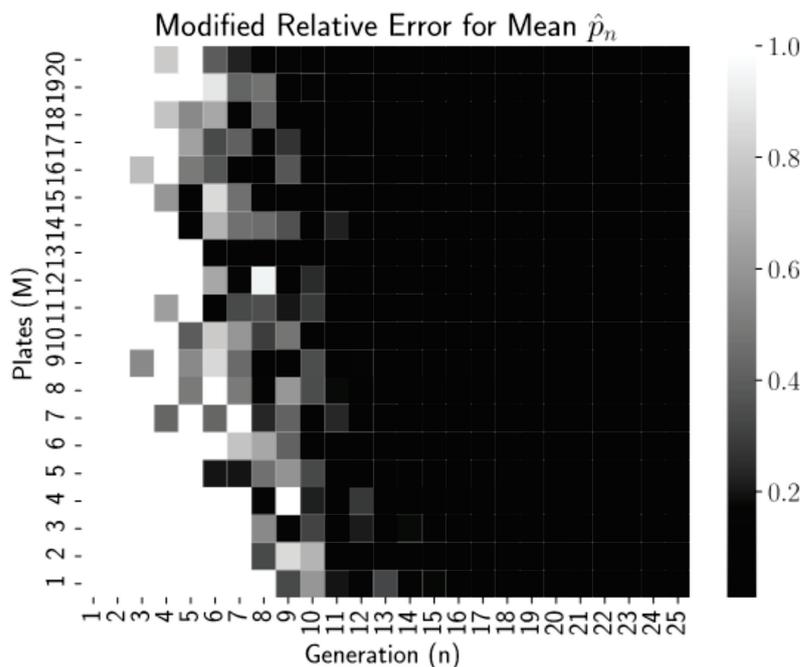
Figure 5. Heatmap of modified relative errors $e_{0.01}$ of mean type 1 mutation rate estimates.

As described in Figure 6, similar to simulation in Section 6, the estimated variance of $\hat{p}_n$ quickly converges to 0 after 13–17 generations (which is also a couple of generations behind the convergence of variance discussed in Section 6), regardless of the number of plates.
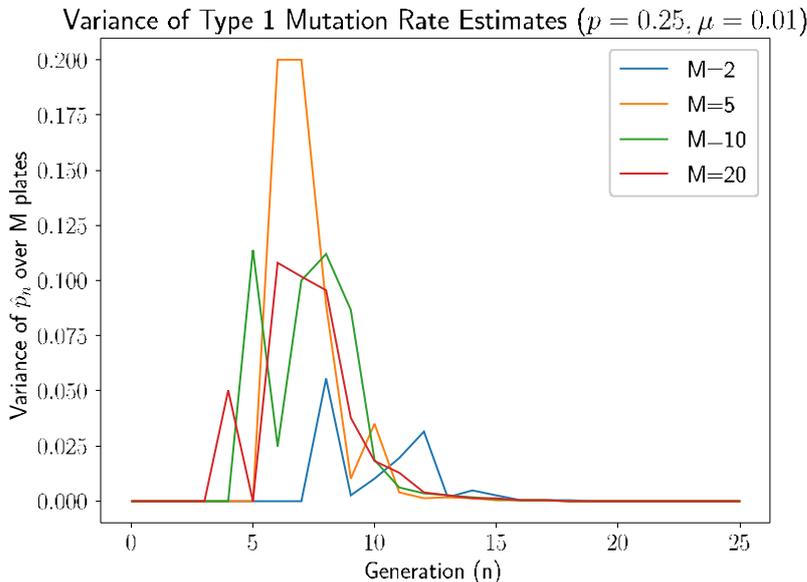
Figure 6. The estimated variance of type 1 mutation rate for $p = 0.25$, $M = 2, 5, 10, 20$ plates for generations $n = 1$ to 25.

9. Summary

In 2012, Niccum et al. [18] offered a simple stochastic estimator $\hat{\mu}_n$ of μ in a most rigorous fashion proving that it was unbiased and consistent. The latter meant that the sequence $\{\hat{\mu}_n\}$ was almost surely pointwise convergent to μ on a non extinct set. Best of all, the pointwise convergence proved to be fast, as it was validated by countless simulation runs for various values of μ.

In this paper, we studied the stochastic process describing bacterial mutation that begins at some point of their division. In our earlier paper [18] (by two of the present

coauthors), we proposed a stochastic estimator $\hat{\mu}_n$ ($n = 1, 2, \ldots$) of the mutation rate μ (or more precisely, mutation probability) and showed that the sequence $\{\hat{\mu}_n\}$ converges to μ a.s. pointwise on the non extinct set $\Omega_0 \subseteq \Omega$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space on which all processes were considered. $\mathbb{P}(\Omega_0) = 1 - \frac{\mu^2}{\nu^2}$, where $\nu = 1 - \mu$, and thus, $\mathbb{P}(\Omega_0) \approx 1$ if μ is small which is the case in many practical situations. The stochastic estimator $\hat{\mu}_n$ obeys Equation (1),

$$\hat{\mu}_{n+1} = \frac{NM_{n+1}}{2X_n},$$

where X_n is the number of nonmutants in generation n and NM_{n+1} is the number of newly formed mutants in generation $n + 1$.

One of the subjects of interest was to establish a different type of convergence, namely L^2. We formed the confined space $(\Omega_0, \mathcal{F}_0 = \mathcal{F} \cap \Omega_0, \mathbb{P}_0 = \mathbb{P}/\mathbb{P}(\Omega_0))$ and established the following results.

1. The sequence of estimators $\{\hat{\mu}_n\}$ is $\mathbb{P}_0$-asymptotically unbiased (that is, $\mathbb{E}_0 \hat{\mu}_{n+1} \to \mu$) and it converges to μ in the $L^2(\mathbb{P}_0)$-norm.
2. In our earlier paper [18], even though we did not rigorously established the speed of pointwise convergence of $\{\hat{\mu}_n\}$ to μ, we conducted simulation as well as numerous times thereafter, showing that the convergence was rapid. With the assumed $\mu = 7.6 \times 10^{-6}$ (the number produced in our lab results), a good accuracy was observed in the 30th generation. We also wanted to test the L^2-speed of convergence in the context of our paper. We used $\hat{\mu}_{n+1}$ in the form

$$\hat{\mu}_{n+1} = \frac{\sum\limits_{i=1}^{2X_n} \eta_{i\cdot}}{2X_n}, \text{ where } \eta_i \in \mathbb{B}[1, \mu]$$

setting X_n as a constant, due to the step-by-step simulation procedures. With generation n relatively small, like 10 or greater, because X_n was large, $\hat{\mu}_{n+1} \in N[\mu, \mu(1-\mu)/2X_n]$ (Gaussian) using the CLT argument. The variance of $\hat{\mu}_{n+1}$ is thus $\mu(1-\mu)/2X_n$ being virtually zero. Yet, we formed the unbiased sample variance

$$\hat{\sigma}_{n+1}^2 = \frac{1}{M-1} \sum_{k=1}^{M} \left(\hat{\mu}_{n+1}^{[k]} - \overline{\mu}_{n+1} \right)^2$$

where M is the number of plates, with bacterial colonies on each and with $\overline{\mu}_{n+1}$ being the sample mean of M estimators (estimates) $\hat{\mu}_{n+1}^{[k]}$, $k = 1, \ldots, M$. The conclusion thus was that $\hat{\sigma}_{n+1}^2$ will converge to zero with M not very large. Indeed, in Example 1, we fully validated it with $\mu = 0.01$, $M = 2, 5, 10, 20$ and with $n \geq 10$. Granted, the convergence would be slower for a smaller μ, but the pointwise convergence of $\hat{\mu}_n$ to $\mu \approx 7.6 \times 10^{-6}$ was around $n = 30$ and with $M = 10$ or 20.
3. We further generalized our model by allowing bacteria mutate into one of the two mutant types, with constant probabilities p and q. One of the interpretation of this model is when the second type could be a dead bacterium (that could be attacked by viruses or chemical agent). As mentioned in the introduction, it can depose its DNA to be picked up by other bacteria. We proposed two respective estimators

$$\hat{p}_n = \frac{NM_n^{[1]}}{NM_n} \text{ and } \hat{q}_n = \frac{NM_n^{[2]}}{NM_n}$$

and proved that they were unbiased and consistent. Here $NM^{[i]}$ is the number of newly formed type 1 mutants. Namely, we proved in Theorem 5 that $\hat{p}_n$ converged to p in the $L^2(\mathbb{P}_n)$-norm, $L^2(\mathbb{P}_0)$-norm, and a.s. pointwise, and $\hat{p}_n$ was asymptotically $\mathbb{P}_0$-unbiased. To prove this result, we had to show that the sequence $\{\mathbb{P}(A_n)\}$ was strictly

monotone increasing and converged to $\mathbb{P}(\Omega_0) = 1 - \frac{\mu^2}{(1-\mu)^2}$. As in the above case, to test the speed of the L^2 convergence, we ran simulation with $\mu = 0.01$, $p = 0.25$, and the number of plates $M = 2, 5, 10, 20$ showing a good speed for the mean and variance for $n \approx 20$ and 15, respectively.

Author Contributions: Conceptualization, J.H.D. and R.R.S.; methodology, J.H.D. and R.R.S.; software, V.M.N. and R.T.W.; validation, J.H.D., V.M.N. and R.T.W.; formal analysis, J.H.D. and V.M.N.; data curation, V.M.N. and R.T.W.; writing—original draft preparation, J.H.D. and V.M.N.; writing—review and editing, J.H.D., V.M.N. and R.T.W.; visualization, V.M.N. and R.T.W.; supervision, J.H.D. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Simulation source code and data will be made available upon request to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

a.s.	almost surely
DNA	Deoxyribonucleic Acid
HGT	Horizontal Gene Transfer
LDCT	Lebesgue Dominated Convergence Theorem
MRSA	Methicillin-Resistant *Staphylococcus aureus*
VRSA	Vancomycin-Resistant *Staphylococcus aureus*

References

1. Foster, P.L. Mechanisms of Stationary Phase Mutation: A Decade of Adaptive Mutation. *Annu. Rev. Genet.* **1999**, *33*, 57–88. [CrossRef] [PubMed]
2. Goodson-Gregg, N.; De Stasio, E.A. Reinventing the Ames Test as a Quantitative Lab That Connects Classical and Molecular Genetics. *Genetics* **2009**, *181*, 23–31. [CrossRef] [PubMed]
3. Luria, S.E.; Delbrück, M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics* **1943**, *28*, 491–511. [CrossRef] [PubMed]
4. Drake, J.W. A Constant Rate of Spontaneous Mutation in DNA-based Microbes. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7160–7164. [CrossRef]
5. Foster, P.L. Methods for determining spontaneous mutation rates. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 2006; Volume 409, pp. 195–213. [CrossRef]
6. Jones, M.E.; Thomas, S.M.; Rogers, A. Luria-Delbruck Fluctuation Experiments: Design and Analysis. *Genetics* **1994**, *136*, 1209–1216. [CrossRef]
7. Ma, W.T.; Sandri, G.V.; Sarkar, S. Analysis of the Luria-Delbrück Distribution Using Discrete Convolution Powers. *J. Appl. Probab.* **1992**, *29*, 255–267. [CrossRef]
8. Oprea, M.; Kepler, T.B. Improved Inference of Mutation Rates: II. Generalization of the Luria-Delbrück Distribution for Realistic Cell-Cycle Time Distributions. *Theor. Popul. Biol.* **2001**, *59*, 49–59. [CrossRef] [PubMed]
9. Rosche, W.A.; Foster, P.L. Determining Mutation Rates in Bacterial Populations. *Methods* **2000**, *20*, 4–17. [CrossRef]
10. Sarkar, S.; Ma, W.T.; Sandri, G.v.H. On Fluctuation Analysis: A New, Simple and Efficient Method for Computing the Expected Number of Mutants. *Genetics* **1992**, *85*, 173–179. [CrossRef]
11. Wu, X.; Strome, E.D.; Meng, Q.; Hastings, P.J.; Plon, S.E.; Kimmel, M. A Robust Estimator of Mutation Rates. *Mutat. Res.* **2009**, *661*, 101–109. [CrossRef]
12. Wu, X.; Zhu, H. Fast Maximum Likelihood Estimation of Mutation Rates Using a Birth-Death Process. *J. Theor. Biol.* **2015**, *366*, 1–7. [CrossRef] [PubMed]

13. Xiong, X.; Boyett, J.M.; Webster, R.G.; Stech, J. A Stochastic Model for Estimation of Mutation Rates in Multiple-replication Proliferation Processes. *J. Math. Biol.* **2009**, *59*, 175–191. [CrossRef] [PubMed]
14. Drake, J.W. *The Molecular Basis of Mutation*; Holden-Day: San Francisco, CA, USA, 1970.
15. Stewart, F.M. Fluctuation Tests: How Reliable Are the Estimates of Mutation Rates? *Genetics* **1994**, *137*, 1139–1146. [CrossRef] [PubMed]
16. Zheng, Q. Progress of a Half Century in the Study of the Luria–Delbrück Distribution. *Math. Biosci.* **1999**, *162*, 1–32. [CrossRef]
17. Zheng, Q. The Luria-Delbrück Protocol Is Still the Most Practical. *J. Theor. Biol.* **2015**, *386*, 188–190. [CrossRef]
18. Niccum, B.A.; Poteau, R.; Hamman, G.E.; Varada, J.C.; Dshalalow, J.H.; Sinden, R.R. On an Unbiased and Consistent Estimator for Mutation Rates. *J. Theor. Biol.* **2012**, *300*, 360–367. [CrossRef]
19. Campbell, A. Synchronization of Cell Division. *Bacteriol. Rev.* **1957**, *21*, 263–272. [CrossRef]
20. Chang, Z.; Shen, Y.; Lang, Q.; Zheng, H.; Tokuyasu, T.A.; Huang, S.; Liu, C. Microfluidic Synchronizer Using a Synthetic Nanoparticle-Capped Bacterium. *ACS Synth. Biol.* **2019**, *8*, 962–967. [CrossRef]
21. Helmstetter, C.E.; Cummings, D.J. Bacterial Synchronization by Selection of Cells at Division. *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 767–774. [CrossRef]
22. Helmstetter, C.E. A Ten-Year Search for Synchronous Cells: Obstacles, Solutions, and Practical Applications. *Front. Microbiol.* **2015**, *6*, 238. [CrossRef]
23. Anderson, P.A.; Pettijohn, D.E. Synchronization of Division in *Escherichia Coli*. *Science* **1960**, *131*, 1098. [CrossRef]
24. Kepes, F.; Kepes, A. Automatic synchronization of growth of "*Escherichia coli*" (author's transl). *Ann. Microbiol.* **1980**, *131*, 3–16.
25. Kepes, F.; Kepes, A. Freeze Preservation of Synchrony in Cultures of Enterobacteriaceae Synchronized by Continuous Phasing in Phosphate-Limited Media. *Biotechnol. Bioeng.* **1984**, *26*, 1288–1293. [CrossRef]
26. Kubitschek, H.E. Linear Cell Growth in *Escherichia Coli*. *Biophys. J.* **1968**, *8*, 792–804. [CrossRef]
27. Ling, H.; Chang, M.W. A Novel Synchronization Approach Using Synthetic Magnetic *Escherichia Coli*. *Synth. Syst. Biotechnol.* **2019**, *4*, 130–131. [CrossRef] [PubMed]
28. Noack, S.; Klöden, W.; Bley, T. Modeling Synchronous Growth of Bacterial Populations in Phased Cultivation. *Bioprocess Biosyst. Eng.* **2008**, *31*, 435–443. [CrossRef] [PubMed]
29. Shehata, T.E.; Marr, A.G. Synchronous Growth of Enteric Bacteria. *J. Bacteriol.* **1970**, *103*, 789–792. [CrossRef]
30. Wallden, M.; Fange, D.; Lundius, E.G.; Baltekin, Ö.; Elf, J. The Synchronization of Replication and Division Cycles in Individual *E. Coli* Cells. *Cell* **2016**, *166*, 729–739. [CrossRef]
31. Mortelmans, K.; Riccio, E.S. The Bacterial Tryptophan Reverse Mutation Assay with *Escherichia Coli* WP2. *Mutat. Res.* **2000**, *455*, 61–69. [CrossRef]
32. Hamel, A.; Roy, M.; Proudlock, R. The bacterial reverse mutation test. In *Genetic Toxicology Testing*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 79–138. [CrossRef]
33. Lee, H.; Popodi, E.; Tang, H.; Foster, P.L. Rate and Molecular Spectrum of Spontaneous Mutations in the Bacterium *Escherichia Coli* as Determined by Whole-Genome Sequencing. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2774–E2783. [CrossRef]
34. Szafrańska, A.K.; Junker, V.; Steglich, M.; Nübel, U. Rapid Cell Division of Staphylococcus Aureus during Colonization of the Human Nose. *BMC Genom.* **2019**, *20*, 229. [CrossRef] [PubMed]
35. Schaaff, F.; Reipert, A.; Bierbaum, G. An Elevated Mutation Frequency Favors Development of Vancomycin Resistance in Staphylococcus Aureus. *Antimicrob. Agents Chemother.* **2002**, *46*, 3540–3548. [CrossRef] [PubMed]
36. Juan, P.A.; Attaiech, L.; Charpentier, X. Natural Transformation Occurs Independently of the Essential Actin-like MreB Cytoskeleton in Legionella Pneumophila. *Sci. Rep.* **2015**, *5*, 16033. [CrossRef] [PubMed]
37. Woo, P.C.Y.; To, A.P.C.; Lau, S.K.P.; Yuen, K.Y. Facilitation of Horizontal Transfer of Antimicrobial Resistance by Transformation of Antibiotic-Induced Cell-Wall-Deficient Bacteria. *Med. Hypotheses* **2003**, *61*, 503–508. [CrossRef]
38. Cafini, F.; Nguyen, L.T.T.; Higashide, M.; Román, F.; Prieto, J.; Morikawa, K. Horizontal Gene Transmission of the *Cfr* Gene MRSA *Enterococcus*: Role of *Staphylococcus epidermidis* as a reservoir and alternative pathway for the spread of linezolid resistance. *J. Antimicrob. Chemother.* **2016**, *71*, 587–592. [CrossRef]
39. Stanczak-Mrozek, K.I.; Laing, K.G.; Lindsay, J.A. Resistance Gene Transfer: Induction of Transducing Phage by Sub-Inhibitory Concentrations of Antimicrobials Is Not Correlated to Induction of Lytic Phage. *J. Antimicrob. Chemother.* **2017**, *72*, 1624–1631. [CrossRef]
40. Dshalalow, J.H. *Real Analysis: An Introduction to the Theory of Real Functions and Integration*; Studies in Advanced Mathematics; Chapman & Hall: Boca Raton, FL, USA, 2001.
41. Dshalalow, J.H. *Foundations of Abstract Analysis*, 2nd ed.; Springer-Verlag: New York, NY, USA, 2014. [CrossRef]
42. Athreya, K.B.; Ney, P.E. *Branching Processes*; Number, B., Ed.; 196 in Die Grundlehren Der Mathematischen Wissenschaften in Einzeldarstellungen Mit Besonderer Berücksichtigung Der Anwendungsgebiete; Springer-Verlag: Berlin, Germany, 1972.
43. Kimmel, M.; Axelrod, D. *Branching Processes in Biology*, 2nd ed.; Interdisciplinary Applied Mathematics; Springer-Verlag: New York, NY, USA, 2015. [CrossRef]