J. Appl. Prob. **11**, 437–444 (1974)
Printed in Israel

A POPULATION BIRTH-AND-MUTATION PROCESS, I: EXPLICIT DISTRIBUTIONS FOR THE NUMBER OF MUTANTS IN AN OLD CULTURE OF BACTERIA

BENOIT MANDELBROT, *IBM Thomas J. Watson Research Center, Yorktown Heights, New York*

Abstract

Luria and Delbrück (1943) have observed that, in old cultures of bacteria that have mutated at random, the distribution of the number of mutants is extremely long-tailed. In this note, this distribution will be derived (for the first time) exactly and explicitly. The rates of mutation will be allowed to be either positive or infinitesimal, and the rate of growth for mutants will be allowed to be either equal, greater or smaller than for non-mutants. Under the realistic limit condition of a very low mutation rate, the number of mutants is shown to be a stable-Lévy (sometimes called "Pareto Lévy") random variable, of maximum skewness β, whose exponent α is essentially the ratio of the growth rates of non-mutants and of mutants. Thus, the probability of the number of mutants exceeding the very large value m is proportional to $m^{-\alpha-1}$ (a behavior sometimes referred to as "asymptotically Paretian" or "hyperbolic"). The unequal growth rate cases $\alpha \neq 1$ are solved for the first time. In the $\alpha = 1$ case, a result of Lea and Coulson is rederived, interpreted, and generalized. Various paradoxes involving divergent moments that were encountered in earlier approaches are either absent or fully explainable.

The mathematical techniques used being standard, they will not be described in detail, so this note will be primarily a collection of results. However, the justification for deriving them lies in their use in biology, and the mathematically unexperienced biologists may be unfamiliar with the tools used. They may wish for more details of calculations, more explanations and Figures. To satisfy their needs, a report available from the author upon request has been prepared. It will be referred to as Part II.

STOCHASTIC PROCESSES; BIRTH-AND-MUTATION PROCESS; NUMBERS OF MUTANTS

1. Introduction

Let the bacteria in a culture grow, and sometimes mutate, at random, for a long time. In an occasional culture, the number of mutants will be enormous, which means that "typical values", such as the moments or the most probable value, give a very incomplete description of the overall distribution. Also, when the same mutation experiment is replicated many times, the number of mutants in one replica, which chanced to be the most active, may exceed by orders of magnitude the sum of the numbers of mutants in the aggregate of all other replicas. Luria and Delbrück (1943), who first observed the above facts, also outlined an explanation that has played a critical role in the birth of molecular biology: if and when the

Received in revised form 18 December 1973.

first mutation occurs very early in an experiment, the advantage of primogeniture is so great that the clone to which it gives rise has time to grow to a very much larger size than either any other clone in the same replica, or than the largest clone grown in a more typical replica in which no early mutation happened to be included.

Interest in expressing this explanation quantitatively, by describing the full distribution of the numbers of mutants, first peaked around 1950 (Lea and Coulson (1949), Kendall (1952), Armitage (1952), (1953) and Bartlett (1966)), but the solutions advanced were not definitive. Several investigators only calculated moments. Also, rates of growth were always assumed to be the same for mutants and non-mutants, and the rate of mutation to be very small. Kendall's work, on the other hand, was so general that it may include in principle the results to be described, but, because of its generality, it lacked explicitness.

In the present note, the whole distribution will be described, for the first time, under assumptions that seem both sufficiently general to be realistic and sufficiently special for the solution to be exact and near explicit, in the sense that the Laplace transform of the distribution is given in closed analytic form.

The extreme statistical variability characteristic of the Luria and Delbrück experiment is also found in other biological experiments in progress; one may therefore hope that a careful study of the earliest and simplest such problem would provide guidance in dealing with new cases when very erratic behavior is unavoidable, and in avoiding them when possible and thus achieving better estimates of such quantities as rates of mutation.

2. Preliminary: assumptions and some known distributions

Assumptions

(A) At time $t = 0$ the culture includes no mutant but includes a large number b_0 of non-mutants of a single kind.

(B) Between times t and $t + dt$, a bacterium has the probability mdt of mutating.

(C) Back mutation is possible.

(D) Neither the mutants nor the non-mutants die.

(E) The rate of mutation m is so small that one can consider each mutation as statistically independent of all others.

(F) Mutants and non-mutants multiply at rates that may be different. The scale of time is so selected that, between the instants t and $t + dt$, the probability of division is gdt for a mutant and dt for a non-mutant.

Non-mutants. A bacterium that mutates may be considered by its non-mutant brethren as having died, so $N(t, m)$, defined as the number of non-mutant bacteria at the instant t, follows the well-known "simple birth and death process" (see, e.g.,

Feller (1968), p. 454). When $b_0 \gg 1$, the variation of $N(t, m)$ is to a good approximation deterministic,

$$N(t, m) \sim EN(t, m) \sim b_0 e^{t(1-m)}.$$

Non-random clones. Denote by $K(t, m)$ the number of "clones" at the instant t (a clone being the progeny of one mutation). From Assumption (E), $K(t, m)$ is so small relative to $N(t, m)$ that different mutations can be considered statistically independent, so $K(t, m)$ is a Poisson random variable of expectation

$$m \int_0^t N(s, m)\, ds \;=\; b_0 m(1-m)^{-1}[e^{t(1-m)} - 1].$$

Random clones. Denote by $Y(t, m, g)$ the number of mutants in a clone selected at random (each possibility having the same probability) among the clones that have developed from mutations that occurred between the instants 0 and t. The distribution of $Y(t, m, g)$ will be seen to depend on its parameters through the combinations e^{gt} and $\alpha = (1-m)/g$; since eventually we shall let $m \to 0$, α nearly reduces to the ratio of growth rates, $1/g$. One can prove that, after a finite t,[1]

$$\Pr\{Y(t, m, g) \geqq y\} \;=\; \alpha[1 - e^{-t(1-m)}]^{-1} \int_1^{\exp(gt)} v^{\alpha-y}(v-1)^{y-1} dv.$$

In the case $\alpha = 1$, this yields explicitly

$$\Pr\{Y(t, m, g) \geqq y\} \;=\; y^{-1}[1 - e^{-gt}]^{y-1}.$$

The generating function (g.f.) of Y, denoted by $\hat{Y}$, equals

$$\hat{Y}(b, t, m, g) \;=\; \alpha[1 - e^{-t(1-m)}] \int_1^{\exp(gt)} v^{-\alpha-1}\{[v(e^b - 1) + 1]^{-1} dv\}.$$

As $t \to \infty$, while m and g are kept constant, Y tends to a limit random variable $Y(\alpha)$ that only depends on α. When $\alpha = 1$,

$$\Pr\{Y(1) = y\} \;=\; \int_0^1 v(1-v)^{y-1}\, dv \;=\; \frac{1}{y(y+1)},$$

a result known to Lea and Coulson (1949). For all α,

$$\Pr\{Y(\alpha) \geqq y\} \;=\; \alpha \frac{\Gamma(\alpha)\Gamma(y)}{\Gamma(\alpha + y)}.$$

---

[1] The formulae in the remainder of this Section restate some results obtained by Yule (1924), in a paper which is known to have introduced the birth process, but has otherwise been almost completely neglected. Yule treated a nominally different problem: our "growth" was his "increase in the number of species in one genus", our "mutation" was his "starting of a new genus." His work and the term "Yule distribution" had owed part of their limited notoriety to several papers by H. A. Simon, who sought to modify Yule's argument to obtain equally strong results from less strong assumptions. This attempt has proved a failure.

For large y,

$$\Pr\{Y(\alpha) \geqq y\} \sim \Gamma(\alpha+1)y^{-1-\alpha}.$$

The $Y(\alpha)$ thus constitutes a form of asymptotically "hyperbolic" or "Pareto" random variable of exponent α. The population moment $EY^h(\alpha)$ is finite if $h < \alpha$ but infinite if $h \geqq \alpha$. For example, the expectation of $Y(\alpha)$ is finite if and only if $\alpha > 1$ and the variance is finite if and only if $\alpha > 2$. Infinite moments are a vital part of the present problem.

3. The total number of mutants

$M(t, m, g)$ will denote the number of mutant bacteria at the instant t. Thus, $M(0, m, g) = 0$, and

$$M(t, m, g) = \sum_{k=1}^{K(t, m, g)} Y_k(t, m, g).$$

Denote its g.f. by $\hat{M}(b, t, m, g)$; since K is a Poisson random variable of expectation EK, $\log \hat{M}(b, t, m, g) = EK[\hat{Y}(b, t, m, g) - 1]$.

Since the distributions of K and Y both depend on t (and are therefore interrelated) one cannot apply the standard theorems concerning the limit behavior of sums (Gnedenko and Kolmogoroff (1954), Feller (1966)), but the special analysis that is required is straightforward. An approximate formal application of the standard theorems, by first letting the Y converge to the $Y(\alpha)$ and then adding K of them, would be unjustified, but some of its results nevertheless remain applicable. (Some of the paradoxes encountered in the analyses *circa* 1950 are related to cases where inversion of limit procedures is unjustified.) One correct formal inference concerns the correct standardizing choices of a scale factor $S(K)$ and a location factor $L(K)$, so as to ensure that the probability distribution of $R = S(K)[M - L(K)]$ tends to a non-degenerate limit as $K \to \infty$. These are as follows:

$$\begin{aligned}
&\alpha > 2: L(K) = EY; && S(K) = (EK)^{-1/2} \\
&1 < \alpha < 2: L(K) = EY; && S(K) = (EK)^{-1/\alpha} \\
&\alpha = 1: L(K) = \log EK; && S(K) = (EK)^{-1/\alpha} \\
&\alpha < 1: L(K) = 0; && S(K) = (EK)^{-1/\alpha}
\end{aligned} \Bigg\} = De^{-gt}.$$

Here, we denote

$$D = [b_0 m(1-m)^{-1}]^{-1/\alpha}.$$

The limits are as follows.

The case $\alpha > 2$. Here, $\lim_{t\to\infty}(EK)^{-\frac{1}{2}}(M - EM)$ can be shown to be Gaussian. Nothing original!

The case $\alpha < 1$. Here, $\lim_{t\to\infty}(EK)^{-1/\alpha}\sum_{k=1}^{K} Y_k$ can be shown to have a g.f. equal to

$$\hat{R}(b, \infty, m, g) = \exp\left[\alpha \int_0^D bw^{-\alpha}(bw+1)^{-1}dw\right].$$

The corresponding limit r.v. – call it $R(\infty, \alpha, D)$ – seems to appear for the first time in the present context. Its being non-degenerate (not reduced either 0 or ∞) confirms that the above standardization was well chosen. Moreover, near $b = 0$, $\hat{R}(b, \infty, m, g)$ has a good expansion in Taylor series, so all moments of $R(\infty, \alpha, D)$ converge. However, this convergence has limited significance because, in actual practice, m is extremely small and D is extremely large, so the moments of $R(\infty, \alpha, D)$ are themselves enormous and tell us very little about the distribution of $R(\infty, \alpha, D)$. On the other hand, as had been realized by Luria and Delbrück, the birth and mutation process is illuminated by a sort of "diagonal" procedure whereby, while t is increased, m and/or b_0 change in such a way that $D \to \infty$ while $g > 1$ to ensure that α remains between 0 and 1. If so, the function $\hat{R}$ tends towards

$$\exp\left[-\alpha b^{\alpha} \int_0^{\infty} z^{-\alpha}(1+z)^{-1}dz\right] = \exp[-b^{\alpha}\alpha\pi/\sin(\alpha\pi)],$$

which is an unfamiliar form of an expression well known in the literature; namely the g.f. if a stable random variable of maximal skewness $\beta = 1$, which is positive (Gnedenko and Kolmogoroff (1954), Feller (1966)). It is also the limit one would have obtained for $K \to \infty$ by first letting $Y \to Y(\alpha)$ and then considering the similarly standardized sum of K independent random variables of the form $Y(\alpha)$.

In the limit, all the moments of order $h > \alpha$ (including all integer moments) diverge. As a practical consequence, the statistical estimation of m and g from values of M is both complicated and unreliable. Traditionally, statistics had relied heavily on sample averages, but when the population averages are infinite, the behavior of the sample averages is extremely erratic, and one must absolutely avoid any method that involves them.

The case $1 < \alpha < 2$. Here, $\lim_{t\to\infty}(EK)^{-1/\alpha}\sum_{k=1}^{K}[Y_k - EY_k]$ can be shown to have the g.f.

$$\exp\left[\alpha\int_0^D b^2 w^{-\alpha+1}(bw+1)^{-1}dw\right].$$

As $D \to \infty$, this function tends towards

$$\exp[-b^{\alpha}\alpha\pi/\sin(\alpha\pi)],$$

which is again the g.f. of a stable random variable of exponent α and maximal skewness, i.e., of the limit of a similarly standardized sum of K independent ran-

dom variables of the form $Y(\alpha)$. The theory of these limits is well known, but their shape is not; see Mandelbrot (1960), Mandelbrot and Zarnfaller (1959).

The case $\alpha = 1$. Here, $\lim_{t\to\infty}(EK)^{-1}\sum_{k=1}^{K}[Y_k - \log EK]$ can be shown to have the g.f.

$$\exp\,[b\log b + b\log(1 + 1/b\,D)].$$

As $D \to \infty$, this function tends towards $\exp[b\log b]$, corresponding to the stable density of exponent $\alpha = 1$ and maximal skewness $\beta = 1$, sometimes called the "asymmetric Cauchy" density. It has been derived (but not identified) in Lea and Coulson (1949), which concerns the case when the mutation rate m is small, and the growth rates for the mutants and the non-mutants are equal, so that $\alpha \sim 1$.

4. The total number of bacteria and the degree of concentration

Designate by $B(t, m, g) = N(t, m) + M(t, m, g)$ the number of bacteria of either kind at the instant t. In the straightforward special case $g = 1$, the function $B(t, m, g)$ follows a "simple birth process" or "Yule process"; see Feller (1968). When $b_0 \gg 1$, the growth of B is for all practical purposes deterministic and exponential, meaning that $B(t) \sim b_0 e^t$.

In the cases $g \neq 1$, things are much more complex, but much of the story is told by the orders of magnitude for large t: $M(t, m, g) \sim e^{gt}$ and $N(t, m, g) \sim e^{(1-m)t}$.

When $\alpha < 1$, $B(t, m, g) \sim M(t, m, g)$, meaning that the mutants — which we know are subject to very large fluctuations — become predominant.

When $\alpha > 1$, $B(t, m, g) \sim e^{(1-m)t}$ with little relative fluctuation, the random factor that multiplies t being nearly the same "as if" there had been no mutation. Thus, the dependence of B upon g is asymptotically eliminated.

Now examine the "degree of concentration" of the mutants, namely the ratio ρ of the number of mutants in the largest of the K clones in a replication, divided by the total number of mutants in the other clones of this replication.

It was discovered by Luria and Delbrück that an alternative ratio can be quite large; namely the number of mutants in the largest among H replications, divided by the sum of the number of mutants in the other of the replications. It can be shown that the above two ratios follow the same distribution, so it will suffice to study the first, beginning with two extreme cases.

Let mutation bring in so great a competitive disadvantage and such decrease in the growth rate that $\alpha \gg 2$. Then, the number of young and small clones increases much faster than the size of the single oldest clone in an experiment. Therefore, it is conceivable that a negligible proportion of mutants will be descended from this oldest clone, and the more so from any other single clone. This expectation is indeed confirmed. We know that if $\alpha > 2$ the quantity $M(t, m, g)$ tends towards a Gaussian limit, so the contribution of any individual addend Y_k to their sum is indeed negligible.

Let, on the contrary, mutation bring in great competitive advantage and such increase in the growth rate that $\alpha \ll 1$. Then, the size of the oldest clone in an experiment (corresponding to the earliest mutation) grows much faster than the number of fresh clones. It is conceivable therefore that the largest clone in an experiment be comparable in size with the sum of all the other clones. An appreciable proportion of the mutants could descend from the single largest clone. This expectation is indeed confirmed in two different ways. First, it has been shown by Darling (1952) (see also Feller (1966), p. 439, problem 20), that if $\alpha < 1$ the ratio ρ does not tend to zero as $K \to \infty$. Rather, its distribution tends to a non-degenerate limit, and $E(1/\rho)$ has the non-degenerate limit $\alpha/(1-\alpha)$. As α varies from 0 to 1, it varies from 0 to ∞. That is, when mutation brings enormous increase in growth rate so that the value of α is very small, $1/\rho$ is nearly 0 on the average, and the limit value of ρ for large K is often very large. When, on the contrary, mutation brings very slight advantage, so that α is very nearly 1, $1/\rho$ is very large on the average and ρ tends to be small. But its values can be seen to be widely scattered, and large values are not unlikely.

The limits described by the preceding theorem are attained asymptotically, rapidly when α is small, but very slowly when α is near 1. Thus, in the Lea and Coulson case corresponding to $\alpha = 1$, the value of K must be very large for ρ to become negligible. For ordinary values of K, the typical value of ρ is non-negligible, and the dispersion of ρ around this typical value is very wide, so that the original argument of Luria and Delbrück is justified.

A different aspect of concentration, attacked by Mandelbrot (1960), concerns the distribution of ρ if the replica is known to be "very large", or if the largest clone it contains is known to be very large. In that case, ρ is likely to be nearly one (see Mandelbrot (1960), p. 96, or Feller (1966), p. 279, problem 27).

References

ARMITAGE, P. (1952) The statistical theory of bacterial populations subject to mutation. *J. R. Statist. Soc.* B **14,** 1–40.

ARMITAGE, P. (1953) Statistical concepts in the theory of bacterial mutation. *J. Hygiene* **51,** 162–184.

BARTLETT, M. S. (1966) *An Introduction to Stochastic Processes.* 2nd. ed. Cambridge University Press.

DARLING, D. A. (1952) The influence of the maximum term in the addition of independent random variables. *Trans. Amer. Math. Soc.* **73,** 95–107.

FELLER, W. (1966) *An Introduction to Probability Theory and Its Applications.* Vol. II. Wiley, New York.

FELLER, W. (1968) *An Introduction to Probability Theory and Its Applications.* Vol. I, 3rd. ed. Wiley, New York.

GNEDENKO, B. V. AND KOLMOGOROFF, A. N. (1954) *Limit Distributions for Sums of Independent Random Variables.* (Translated by K. L. Chung) Addison-Wesley, Reading, Mass.

KENDALL, D. G. (1952) Les processus stochastiques de croissance en biologie. *Ann. Inst. Poincaré* **13,** 43–108.

LEA, D. E. AND COULSON, C. A. (1949) The distribution of the number of mutants in bacterial populations. *J. Genetics* **49,** 264–285.

LURIA, S. E. AND DELBRÜCK, M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, **28**, 491-511.

MANDELBROT, B. (1960) The Pareto-Lévy law and the distribution of income. *Internat. Economic Rev.* **1**, 79–106 and **4** (1963) 111–115.

MANDELBROT, B. AND ZARNFALLER, F. (1959) Five place tables of certain stable distributions. *Research Report* RC–421, IBM Research Center. (Available from the first named author, as part of a revised reprint of the above 1960 paper.)

YULE, G. U. (1924) A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Philos. Trans. Roy. Soc. London* B **213**, 21–87.