The Probability of Monogenism

by Ulrich Utiger*

ABSTRACT: According to various studies on mitochondrial DNA, which is only inherited via the maternal part, all modern humans are descended from a primordial mother who lived in East Africa around 200,000 years ago, earning her the nickname “mitochondrial Eve”. This concept of monogenism stands in opposition to most evolutionary biologists, claiming that it is very unlikely that mitochondrial Eve was the only mother of our species, in other words, that polygenism is the more likely scenario. Here, I verify this claim using probability theory and empirical data about hunter-gatherers.

* Ulrich Utiger is an independent researcher who earned a master’s degree in theoretical physics from the University of Fribourg, Switzerland. He was a lecturer at the HFTM College in Switzerland and is retired since 2018.
Updates to this article can be downloaded from historycycles.org/monogenism.pdf.
The Mathematica notebook codes.nb is available from historycycles.org/codes.zip.
Contact: ulrich.utiger@gmail.com; DOI: 10.6084/m9.figshare.25745754

Last updated on Mai, 2024

Introduction

The view that all humans descend from a single ancestor couple prevailed during centuries until Africa and later America were discovered and colonized. In the 18th and 19th centuries, the most abominable and pseudoscientific theories emerged to justify the exploitation of peoples considered to be inferior human races, which happened even before Charles Darwin, who surprisingly was a monogenist. His views nevertheless accelerated this emergence of scientific racism as Darwin held that there were hierarchically distinct races due to natural selection, even though he distanced himself from slavery (Cohen 1980 ch. 8; Jackson & Weidman 2004 chs. 2-3).

Around the same time, in 1868, the first fossil bones of modern humans, dated to around 30 ka, were found in Cro-Magnon (France), which is why our species is often named after this site. Subsequently, other fossils were unearthed, predominantly in Europe but later also elsewhere, which led to a wide spectrum of evolutionary proposals (Sykes 2001 p. 114; Klein 2009 pp. 617-619). According to the multi-regionalists, our ancestor *Homo erectus* evolved independently into different modern humans on different continents, thus producing different races (Reichholf 1998 p. 13). However, the oldest fossils dated to around 200 ka were subsequently excavated in eastern Africa, which eventually led to the out-of-Africa theory, standing in contradiction to these views favoring multi-regionalism (Klein 2009 p. 738-739; Bräuer 2015 pp. 2300-2307).

In support of this date and origin, Cann et al. (1987) showed that all modern humans descend from one African mother who lived about 200’000 years ago. The article led to a heated debate as it was a blow for the then still predominant view of multi-regionalism. Their results were confirmed by other genetic studies based on mitochondrial DNA, which is only inherited from the maternal part (Klein 2009 pp. 615, 631-638). In the following years, the view that modern humans emerged only in Africa and replaced *Homo erectus* on the whole planet was finally adopted. Our African ancestor was even nicknamed *mitochondrial Eve*, although evolutionary biologists distance themselves from this sobriquet inspired by Genesis.

However, the concept of polygenism was not abolished at all (Brown 1990 ch. 3; Henke & Tattersall 2015 p. 2303). In fact, there is the consensus among scholars that it is very unlikely that mitochondrial Eve was the only original mother of our species, arguing that during her lifetime there may have been thousands of other mothers who were just as modern, but that their lineages died out (Brown 1990 pp. 108-111; Ayala 1995; Foley 1995 p. 129; Reichholf 1998 p. 20; Sykes 2001 p. 277; Sykes 2003 pp. 136-138; Klein 2009 p. 634; Dawkins & Wong 2017 p. 60).

As reported by Callaway (2017), fossils from Jebel Irhoud (Morocco) were discovered and dated to about 40 ka in 1960 and the following year. They were then considered to belong to an African form of Neanderthals. However, when excavations in 2017 at the same site by another team allowed a new dating to about 315 ka, Hublin et al. (2017) assigned them to *H. sapiens*, thus shifting the origin of our species about 100 thousand years into the past. Jean-Jaques Hublin and colleagues support the view that the emergence of *H. sapiens* was gradually and not restricted to a small region (like eastern Africa) but involved the whole African continent, Callaway citing Hublin:

> Until now, the common wisdom was that our species emerged probably rather quickly somewhere in a 'Garden of Eden' that was located most likely in sub-Saharan Africa. I would say the Garden of Eden in Africa is probably Africa — and it's a big, big garden.

So here we see resurfacing the old concept of polygenism, to which still many other paleoanthropologists adhere (Gibbons 2017; Stringer & Galway-Witham 2017; Scerri et al. 2018). However, Meneganzin et al. (2022) argue that, according to Ernst Mayr (1954 pp. 157-180) speciation is most likely to occur in small populations and thereby reject the Pan-African multi-regional hypothesis of Hublin et al. (2017). They also hold that speciation is not gradual but a rapid process, referring to the punctuated equilibrium concept of Gould and Eldredge, and criticize the assignment of the Jebel Irhoud material to modern humans as the skull is visibly elongated like that of a Neanderthal. The same is criticized by María Martínon-Torres, a paleoanthropologist at University College London, stressing that the remains from Jebel Irhoud lack features that characterize our species, such as a prominent chin and forehead. Also, Jeffrey Schwartz of the University of Pittsburgh, Pennsylvania, objects that too many different-looking fossils have been lumped together with our species (Callaway 2017).

This brief overview shows that evolutionary biologists often base their conclusions on their biased intuition, favoring the Darwinian concept of polygenism, which is indeed a corollary to the theory of natural selection as a species-forming mechanism. Francisco J. Ayala (1995), for instance, links monogenism to the probability of a bottleneck in human evolution, leaving behind a single ancestor pair, which is indeed very unlikely. In other words, if there were several lineages in the beginning, it would be very likely that all present humans descend from several lineages, since their extinction in a bottleneck is very unlikely. Now, if Ayala had an unbiased intuition, he would arrive to the complementary conclusion that it is very unlikely that there have been several lineages in the beginning, knowing that at present there is only one lineage.

Whether there has been a bottleneck or not is the incorrect question to ask because this presupposes that a bottleneck is the only possibility in order for humanity to stem from single parents. We are going to see that monogenism can indeed also emerge from a large population. So the right question to ask is how likely it is that the whole of humanity stems from a single ancestor pair, *knowing* that all present humans stem from a single mother. Together with an estimate of the possible population

size in the region and at the time of the mitochondrial Eve, this is the only knowledge we have at our disposal.

The basic mathematical problem is as follows: let *t* be the total number of women who lived in Africa at the time when modern humans are supposed to have diverged from our ancestors. It is not necessary to know the extend of this region, nor the time of divergence and whether the whole population was in reproductive contact or not, which is why the region could in principle be extended to the whole planet and the number *t* to its whole population. Let then be *n* the hypothetical number of maternal lineages from which humanity is supposed to descend. The claim from evolutionary biologists is that this number might be greater than one because the possibility exists that $n-1$ of them died out up to the present. So this number lies in the interval $1 \le n \le t$.

Now, in the absence of any knowledge except an approximate *t*, every number *n* has equal probability. This can be compared to a police lineup of *t* persons among whom are *n* suspected criminals. From the point of view of the aligned persons, they exactly know whether they are guilty or not. From the point of view of the witness behind the one-way mirror, not every person is equally suspicious because he/she knows something about the alleged criminals, possibly having seen their faces or heard their voices. For someone who only knows their number, the probability that a certain person is a criminal is n/t (if all suspected are guilty) because *n* is the number of outcomes and *t* the total number of outcomes. From the point of view of the police, the situation is still different because it does not know with certainty whether all *n* suspected are guilty. But for someone who knows nothing about the *t* persons, every number *n* is equally probable. This shows that probability is not absolute but depends on the knowledge of someone who calculates it.

The situation of complete ignorance except for the number *t* can be compared to throwing a die. Each number *n* of $t = 6$ possible numbers rolled is just as likely as any other. In other words:

$$P(n) = \frac{1}{6} \text{ for all } n \text{ between 1 and } t = 6$$

Let now be the event A_n = "*n* lineages existed in the beginning", which is short for "the existence of *n* maternal lineages in the beginning of modern humans' divergence from their ancestors". If our knowledge is insufficient to attribute a higher or lower probability to a certain number of lineages, the situation is analog to throwing a dice:

$$P(A_n) = \frac{1}{t} \;\forall\; 1 \le n \le t \tag{1}$$

So this is a uniform probability distribution. Let now be B_m = "*m* lineages have survived after *v* generations". This knowledge changes the probability of A_n, in other words, $P(A_n) \neq P(A_n \mid B_m)$. As $m = 1$ in present times, we will be able to calculate $P(A_1 \mid B_1)$ from this and compare it to $P(A_{n>1} \mid B_1)$ within the limits of *t*, which is what we want to know in the end. The A_n are a complete system of mutually incompatible events as "*i* lineages were present" AND "*j* lineages were present" cannot occur at the same time, supposing that $i \neq j$. This implies that

$$\sum_{n=1}^{t} P(A_n) = 1$$

which is effectively the case because of equation 1. It is evident that B_m can only occur if previously A_n has occurred, under the condition that $0 \le m \le n$. This is a necessary condition for the law of total probability, according to which we have

$$P(B_m) = \sum_{n=1}^{t} P(A_n) P(B_m \mid A_n)$$

With equation 1 and the condition $0 \le m \le n$, this becomes

$$P(B_m) = \frac{1}{t} \sum_{n=m}^{t} P(B_m \mid A_n) \tag{2}$$

What we finally want to know, we get from Bayes' theorem using equations 1 and 2:

$$P(A_n \mid B_m) = \frac{P(A_n) P(B_m \mid A_n)}{P(B_m)} = \frac{\frac{1}{t} P(B_m \mid A_n)}{\frac{1}{t} \sum_{k=m}^{t} P(B_m \mid A_k)} = \frac{P(B_m \mid A_n)}{\sum_{k=m}^{t} P(B_m \mid A_k)} \tag{3}$$

So what we have to calculate for this purpose is $P(B_m \mid A_n)$, which will be done in the following sections.

Fertility Rate

The first quantity we need is the probability b_i that a woman will give birth to i daughters of reproductive age during her lifetime. For this purpose, we will use the data of Bentley et al. (1993) about hunter-gatherers because our ancestors are likely to have adopted this way of life. However, these data indicate the fertility rate, including both girls and boys. So we still need to transform this rate into a rate limited to girls. Let's label the fertility rate related to both sexes x_k and start with $i = 0$. For briefness, I will refer to kids rather than children and girls rather than daughters. The probability that a woman has no girls is

$$f_0 = P($$
the woman has 0 kids or she has 1 kid and it is a boy or… or
the woman has k kids and they are all boys or… or
the woman has d kids and they are all boys) $= \sum_{k=0}^{d} a_k$

with

$$a_k = P(\text{the woman has } k \text{ kids and they are all boys})$$
$$= P(\text{the woman has } k \text{ kids})\, P(\text{they are all boys}) = x_k\, y^k$$

using the relation $P(A \text{ and } B) = P(A) P(B \mid A)$. This is a general formula for an arbitrary event B depending on A. In the current context, A = "the woman has k kids" and B = "they are all boys", which depends on the number k of kids of the mother. This relation will be used extensively below but without being mentioned every time for the sake of briefness. On the other hand, y is the sex ratio at birth, which varies from one nation to another and through history, depending on many factors. So this is a somewhat arbitrary parameter, but it doesn't significantly influence the end results, as we are going to see. Here it is taken

$$y = \frac{105\,\text{boys}}{100\,\text{girls}} \approx 0.51$$

according to Coale 1972 p. 15. So

$$f_0 = \sum_{k=0}^{d} x_k y^k = x_0 + \sum_{k=1}^{d} x_k y^k$$

With this formula, x_0 is not multiplied with y in the special case "the woman has 0 kids", as required. Then,

$$f_1 = P(\text{the woman has 1 kid and it is a girl or… or the woman has } k \text{ kids and among them is a girl or… or the woman has } d \text{ kids and among them is a girl}) = \sum_{k=1}^{d} a_k$$

with

$$\begin{aligned} a_k &= P(\text{the woman has } k \text{ kids and among them is a girl}) \\ &= P(\text{the woman has } k \text{ kids})\ P(\text{among them is a girl}) \end{aligned}$$

where

$$\begin{aligned} &P(\text{among them is a girl}) \\ &= P(\text{the } 1^{\text{st}} \text{ kid is a girl and all others boys or… or the } i^{\text{th}} \text{ kid is a girl and all others boys or… or the } k^{\text{th}} \text{ kid is a girl and all others boys}) \\ &= k\,P(1 \text{ of } k \text{ kids is a girl and all others boys}) \\ &= k\,(1-y)\,y^{k-1} \end{aligned}$$

Therefore

$$f_1 = \sum_{k=1}^{d} a_k = \sum_{k=1}^{d} k\,x_k\,(1-y)\,y^{k-1}$$

Analogously

$$f_2 = P(\text{the woman has 2 kids and they are 2 girls or… or the woman has } k \text{ kids and among them are 2 girls or… or the woman has } d \text{ kids and among them are 2 girls}) = \sum_{k=2}^{d} a_k$$

with

$$\begin{aligned} a_k &= P(\text{the woman has } k \text{ kids and among them are 2 girls}) \\ &= P(\text{the woman has } k \text{ kids})\ P(\text{among them are 2 girls}) \end{aligned}$$

where

P(among them are 2 girls)

$$= \binom{k}{2} P(\text{2 of } k \text{ kids are girls and the others boys})$$

$$= \binom{k}{2} (1-y)^2 \, y^{k-2}$$

and

$$\binom{k}{2} = \frac{k!}{2!(k-2)!}$$

is the number of all combination to choose 2 girls out of k. Therefore

$$f_2 = \sum_{k=2}^{d} a_k = \sum_{k=2}^{d} \binom{k}{2} x_k \, (1-y)^2 \, y^{k-2}$$

From this, the general formula can easily be guessed:

$$f_i = \sum_{k=i}^{d} \binom{k}{i} x_k \, (1-y)^i \, y^{k-i} \tag{4}$$

Now, the normalized data from Bentley et al. (1993 p. 274) are reproduced in figure 1. They are based on a reduced sample, therefore lacking some smoothness. As we need an integer k for the probability x_k that during her lifetime a woman brings to birth k children reaching the reproduction age, it is necessary to average the bars. In order to keep the central peak, I will sum the bars at 5.75 and 6.25, the middle of which is 6. The bars at the left and the right of the peak are merged analogously. The bars at the outer left and right keep the same height but get the fertility rates of 3 and 10, respectively. This will smooth the data a bit and one gets integers for TF.

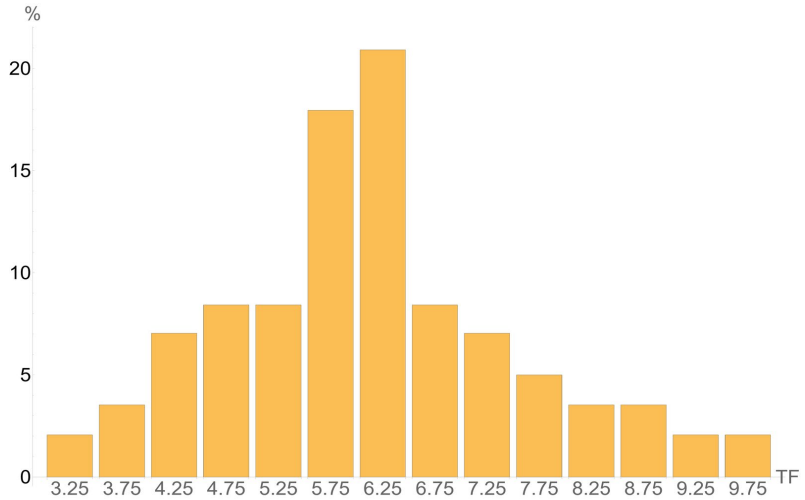

Figure 1: Percentage distribution of the total fertility (TF) of natural fertility populations (reproduced from Bentley et al. (1993 p. 274)).

Next, I will try to find a probability distribution that fits the data approximately. This will allow us to guess the probabilities for the rates 0, 1 and 2 on the left as well as for the rates 11 and 12 on the right. I have tried out several distributions. It seems that a Cauchy distribution comes closest to the data. As the original data are normalized only over the interval from 3.25 to 9.25, including the guessed rates on the left and right will necessitate renormalizing all data. The result can be seen in figure 2.

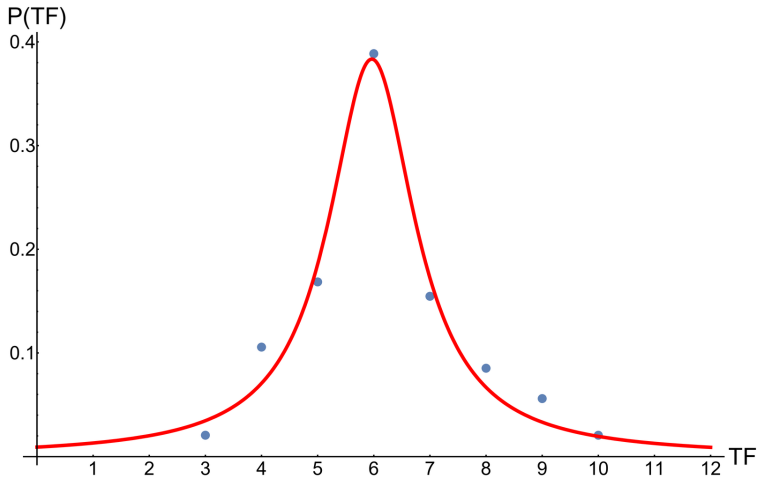

Figure 2: The smoothed Bentley data (blue points) approximated by a Cauchy distribution (red curve).

The x_k are finally obtained using the data points of the fitted Cauchy distribution and normalizing them in the interval from 0 to 12. Putting these x_k into formula 4, one finds the mean value

$$\overline{f} = \sum_{i=1}^{d} i\, f_i$$

From this mean, one finds the $p = \overline{f}/d$ for the binomial distribution

$$b_i = \binom{d}{i} p^i (1-p)^{d-i} \tag{5}$$

because its mean value is $\overline{b} = d\, p$. As can be seen from figure 3, the data thus obtained (red points) fit very well the f_i (blue points). Now, as can been also is that the probabilities of the rates 10 to 12 are almost 0. Since we are interested to keep d as small as possible because the time to calculate formula 12 increases exponentially with increasing d, we can reduce d from 12 to 9 without significantly altering the accuracy because the original data are also approximate, not only because they are derived from a small population sample but also because we don't exactly know under what conditions our first ancestors lived. What matters is that the data form some kind of a bell instead of a uniform distribution.

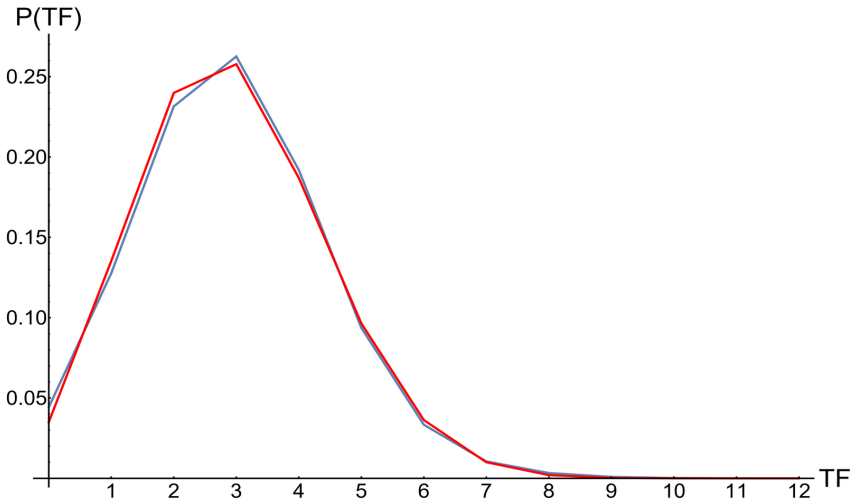

Figure 3: The Bentley data approximated by a Cauchy distribution yield the x_k, which are put into formula 4. The data thus obtained (blue points) fit very well a binomial distribution (red points). Here, TF is the total fertility rate limited to the birth of girls.

This reducing can be done in different ways. I prefer to sum the rates from 9 to 12, which then yields the probability that a woman brings about 9 to 12 children. However, after this transformation, we don't have anymore a binomial distribution. So it must then be converted back into such a distribution by determining a slightly different *p* from the mean value in the same way as explained above. The reason why it is advantageous to have a binomial distribution is because

$$\sum_{S}\prod_{i=1}^{r} b_{s_i} = \binom{d\,r}{s} p^s (1-p)^{d\,r-s}$$

if the b_i are a binomial distribution of the same *p* and *d* as in formula 5. This relation will be used in formula 7 and allow for a much faster calculation. It is given without proof. The reader can try it out by putting in some values to verify its correctness (see codes.nb).

One Lineage

The calculation of $P(B_m \mid A_n)$ in equation 3 has to be done in two steps. First, I will suppose that there was only one original mother and then apply this result to the case of several mothers in the next section. I will elaborate the probability P_v that this only lineage dies out at generation *v* and the probability $Q_{v,s}$ that the lineage has *s* girls after *v* generations, s being greater than zero. Indeed, $P_v = Q_{v,0}$ but this is a special case that has to be distinguished from $s > 0$. I also label b_s the probability that a woman brings about *s* girls during her lifetime, under the condition that they reach the age of reproduction. As discussed in the previous section, the b_s are a binomial distribution based on empirical data about hunter-gatherers. The maximal possible number of daughters will be labeled *d*. From this follows that

$$P_1 = b_0 \text{ and } Q_{1,s} = b_s \tag{6}$$

For $v > 1$, let's proceed step by step and consider the next case $v = 2$, labeling *r* the number of girls of the preceding generation and *s* the number of girls of the actual generation:

$$\begin{aligned} P_2 = P(&\text{the mother has 1 girl and this girl has 0 girls or… or}\\ &\text{the mother has } r \text{ girls and these girls have together 0 girls or… or}\\ &\text{the mother has } d \text{ girls and these girls have together 0 girls}) = \sum_{r=1}^{d} a_r \end{aligned}$$

where

$$\begin{aligned} a_r &= P(\text{the mother has } r \text{ girls and these girls have together 0 girls})\\ &= P(\text{the mother has } r \text{ girls})\, P(\text{these girls have together 0 girls}) \end{aligned}$$

and

$$P(\text{the mother has } r \text{ girls}) = b_r = Q_{1,r}$$

according to equation 6, while

$$\begin{aligned} &P(\text{these girls have together 0 girls})\\ &\quad = P(\text{the 1}^{\text{st}} \text{ girl has no girl and… and the } r^{\text{th}} \text{ girl has no girl}) = b_0^r \end{aligned}$$

So

$$P_2 = \sum_{r=1}^{d} a_r = \sum_{r=1}^{d} Q_{1,r}\, b_0^r$$

The probability that the 2nd generation has together 1 girl is

$$Q_{2,1} = P($$
the mother has 1 girl and this girl has 1 girl or… or
the mother has r girls and these girls have together 1 girls or… or
the mother has d girls and these girls have together 1 girls$) = \sum_{r=1}^{d} a_r$

where

$$\begin{aligned} a_r &= P(\text{the mother has } r \text{ girls and these girls have together 1 girls}) \\ &= P(\text{the mother has } r \text{ girls})\ P(\text{these girls have together 1 girls}) \\ &= Q_{1,r}\ P(\text{these girls have together 1 girls}) \end{aligned}$$

with

$$\begin{aligned} &P(\text{these girls have together 1 girl}) \\ &= P(\text{the } 1^{\text{st}} \text{ girl has 1 girl and the other } r-1 \text{ have none or… or} \\ &\quad \text{the } r^{\text{th}} \text{ girl has 1 girl and the other } r-1 \text{ have none}) = r\, b_1\, b_0^{r-1} \end{aligned}$$

So

$$Q_{2,1} = \sum_{r=1}^{d} a_r = \sum_{r=1}^{d} Q_{1,r}\, r\, b_1\, b_0^{r-1}$$

Things get more complicated in the general case where the girls of the mother have together s girls. We get again

$$Q_{2,s} = \sum_{r=1}^{d} a_r$$

analogously to the former case but with

$$\begin{aligned} a_r &= P(\text{the mother has } r \text{ girls and these girls have together } s \text{ girls}) \\ &= P(\text{the mother has } r \text{ girls})\ P(\text{these girls have together } s \text{ girls}) \\ &= Q_{1,r}\ P(\text{these girls have together } s \text{ girls}) \\ &= Q_{1,r}\ P(\text{the } 1^{\text{st}} \text{ girl has } s_1 \text{ girls and… and the } r^{\text{th}} \text{ girl has } s_r \text{ girls or… or}) \\ &= Q_{1,r} \sum_{S} \prod_{i=1}^{r} b_{s_i} \end{aligned}$$

where

$$S = \left\{ s_i \in \mathbb{N} \mid s_1 + \ldots + s_r = s;\ 0 \le s_i \le d \right\}$$

Thereby

$$Q_{2,s} = \sum_{r=1}^{d} a_r = \sum_{r=1}^{d} Q_{1,r} \sum_{S} \prod_{i=1}^{r} b_{s_i}$$

To better understand this formula, let's illustrate it with the example $d = 2$ and $s = 1$. Since r is iterated from 1 to d, we have 2 cases:

$r = 1$ yields the only combination $(s_1 = 1) \Rightarrow a_1 = Q_{1,1}\, b_1 = b_1\, b_1$

$r = 2$ yields the combinations $(s_1 = 0, s_2 = 1)$ and $(s_1 = 1, s_2 = 0)$

$\Rightarrow a_2 = Q_{1,2}\left(b_0\, b_1 + b_1\, b_0\right) = 2\, b_0\, b_1\, b_2$

$\Rightarrow Q_{2,1} = a_1 + a_2 = b_1\left(b_1 + 2\, b_0\, b_2\right)$

So this yields the same result as if it were calculated with the formula above. One can show analogously that

$$P_3 = \sum_{r=1}^{d^2} Q_{2,r}\, b_0^r \text{ and } Q_{3,s} = \sum_{r=1}^{d^2} Q_{2,r} \sum_{S} \prod_{i=1}^{r} b_{s_i}$$

from which it is easy to guess the general recursive formulas

$$P_v = \sum_{r=1}^{d^{v-1}} Q_{v-1,r}\, b_0^r \text{ and } Q_{v,s} = \sum_{r=1}^{d^{v-1}} Q_{v-1,r} \sum_{S} \prod_{i=1}^{r} b_{s_i}$$

where $Q_{0,1} = 1$ because this is the probability that there is only one original mother, which by definition is 100% certain. On the other hand, $Q_{0,r>1} = 0$ because there is only one original mother, but this case never occurs in the formulas. Under the special conditions $d = 0$ and $v = 1$, these formulas can be tested analytically. Without using the formula, one should get $P_1 = 1$ because this is the probability that the first generation dies out, which is certain if the original mother is unable to bring about daughters. Using the formula, r goes from 1 to 0^0, which is normally undetermined. However, it is only the exponent $v - 1$ that is exactly 0, while d must not imperatively be an exact integer because it is the maximal number of daughters a woman can bring about. In a hypothetical society being hit by some rare event, which would reduce the fertility to almost zero, it would be calculated by the number of women still capable of reproducing a single daughter divided by the number of all women between a certain age, which would yield a number close to but greater than zero. So it could also be just a very small real number. Thereby, 0^0 must be treated as $\lim_{d \to 0} d^0 = 1$, which yields

$$P_1 = \sum_{r=1}^{d^0 = 1} Q_{0,r}\, b_0^r = Q_{0,1}\, b_0^1 = b_0 = 1 \text{ as required.}$$

If $d = 0$ and $v > 1$, then one should get $P_{v>1} = 0$ because the first generation certainly has already diet out. So there can be no further generation. Using the formula in this case, r goes from 1 to $0^{v-1} = 0$, in other words, the sum is equal to zero, which means that $P_{v>1} = 0$ as expected. For $Q_{v,s}$ one can argue similarly: if $d = 0$ and $v = 1$, then one should obtain $Q_{1,s} = 0$ because the original mother is unable to reproduce. Using the formula, one gets

$$Q_{1,s} = \sum_{r=1}^{d^0 = 1} Q_{0,r} \sum_{S} \prod_{i=1}^{r} b_{s_i} = Q_{0,1} \sum_{S} \prod_{i=1}^{1} b_{s_i} = \sum_{S} \prod_{i=1}^{1} b_{s_i}$$

with $S=\{s_1=s\}$. But $s_1=0$ because of $0\le s_1\le d$, which is why one gets $Q_{1,0}=b_0=1$. This seems to be in contradiction with the requirement that $Q_{1,s}=0$. However, $Q_{1,0}$ is not the probability that the lineage survives but dies out, in other words, $Q_{1,0}=P_1$, so this amounts to the same as above. For $v>1$, $Q_{1,s}=0$ because r goes from 1 to 0, as explained above.

For a given r of the previous generation, the maximal number of s of the actual generation is equal to $d\,r$. Therefore if $s>d\,r$, $S=\{\}$ and thereby $\sum\prod b_{s_i}=0$. So there is no need to calculate these cases. To exclude them, we set

$$s\le d\,r\Rightarrow r\ge\frac{s}{d}$$

As this does not always yield an integer, it has to be rounded up, finally getting

$$Q_{v,s}=\sum_{r=\lceil s/d\rceil}^{d^{v-1}}Q_{v-1,r}\sum_{S}\prod_{i=1}^{r}b_{s_i} \tag{7}$$

The maximal number of daughters at generation v is d^v. Therefore, the probability that a lineage survives, in other words, that there is at least one daughter present at generation v, is

$$Q_v=\sum_{s=1}^{d^v}Q_{v,s} \tag{8}$$

Let now $X=v$ be the random event "the lineage dies out at generation v", which can take any positive integer. Therefore,

$$1=\sum_{v=1}^{\infty}P(X=v)=P(X\le v)+P(X>v)\;\Rightarrow\;P(X>v)=1-P(X\leqslant v)$$

where $P(X>v)$ is the probability that the lineage does not die out at the 1st generation nor at the 2nd… nor at the vth. In other words, $P(X>v)=\mathrm{Q}_v$. On the other hand, $P(X\le v)$ is the probability that the lineage dies out at the 1st generation or at the 2nd… or at the vth. In other words, $P(X\le v)=\sum P_k$, from which one gets the more rapid alternative formula

$$Q_v=1-\sum_{k=1}^{v}P_k$$

It will be interesting to test equation 3 in the case where the $b_i=u=1/(d+1)$ are a uniform distribution $P(i)=u$ for 0 to d. Its mean value is

$$\sum_{i=0}^{d}i\,P(i)=\sum_{i=1}^{d}i\,P(i)=u\frac{d(d+1)}{2}=\frac{1}{d+1}\frac{d(d+1)}{2}=\frac{d}{2}$$

whereas the mean for the binomial distribution is close to 3 according to figure 3. So $d=6$ is a realistic value when comparing results using both distributions. With this, one gets

$$\prod_{i=1}^{r}b_{s_i}=u^r$$

On the other hand, $\sum_S$ will still sum over $S=\{s_i\in\mathbb{N}\mid s_1+\ldots+s_r=s;\,0\le s_i\le d\}$, but there is the formula

$$f(r,s,d)=\sum_{i=0}^{\left\lfloor \frac{s}{d+1} \right\rfloor}(-1)^i\binom{r}{i}\binom{r+s-(d+1)i}{r-1}$$

that calculates the number of these combinations directly. In addition, there is also a recursive function that is still faster. These formulas are given without proof. The reader can just verify them by trying out with some arbitrarily chosen parameters (see codes.nb). With this formula we finally get

$$P_v=\sum_{r=1}^{d^{v-1}}Q_{v-1,r}\,u^r \text{ and } Q_{v,s}=\sum_{r=\lceil s/d\rceil}^{d^{v-1}}Q_{v-1,r}f(r,s,d)u^r \tag{9}$$

These formulas can be tested numerically with Mathematica. When using a uniform distribution for the b_i, this can be done for different d. As shown by figure 4, there is neat agreement between the numerical values and the analytical results. What can be observed also is that P_v rapidly converges to zero for increasing v. This convergence is all the more rapid as d increases. This comes as no surprise because the probability that a lineage dies out is only likely for the first generations. Once a population has reached a certain size after some generations, it is nearly impossible that it dies out by a bottleneck. When low fertility rates are used (low d), the probability that the lineage dies out is higher, implying a slower convergence. For Q_v, we have the inverse situation. For increasing d, it is no surprise that the likelihood that the lineage survives increases and also rapidly converges to certain values.

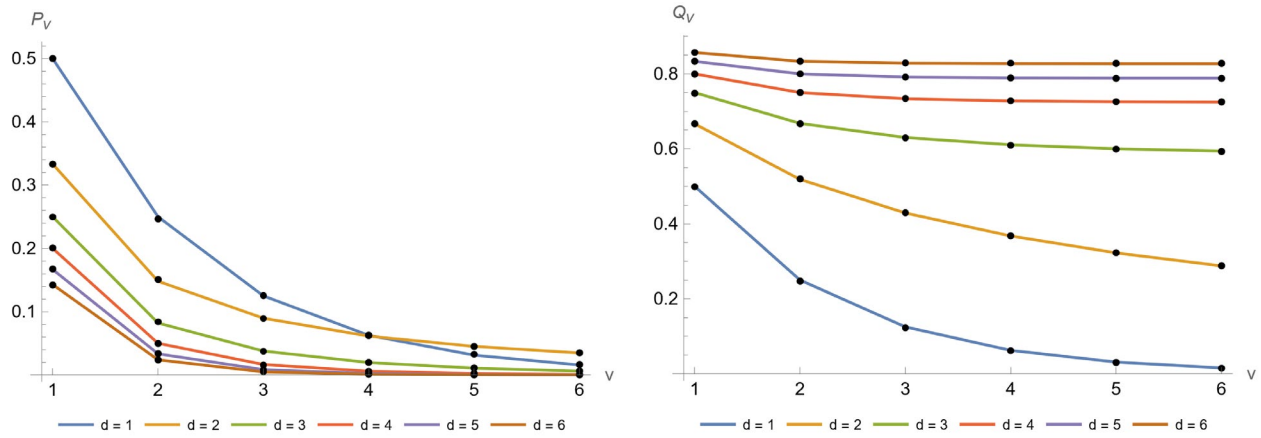

Figure 4: There is neat agreement between the numerical results (black points) and the analytical solutions (colored joined points), using a uniform distribution for P_v and Q_v.

Using a binomial distribution for the b_i, the behavior is almost the same as for a uniform distribution in the case $d=6$, as shown by figure 5, even though here the probability of survival is somewhat higher. Another result we will need is as follows: let A_v = "a lineage has survived until generation v" and B_v = "the lineage dies out at generation v", getting

$$P_v=P(A_{v-1}\text{ and }B_v)=P(A_{v-1})P(B_v\mid A_{v-1})\Rightarrow$$
$$P(B_v\mid A_{v-1})=\frac{P(A_{v-1}\text{ and }B_v)}{P(A_{v-1})}=\frac{P_v}{Q_{v-1}} \tag{10}$$

Analogously, one has

$$Q_v = P\left(A_{v-1} \text{ and } A_v\right) = P\left(A_{v-1}\right) P\left(A_v \mid A_{v-1}\right) \Rightarrow$$
$$P\left(A_v \mid A_{v-1}\right) = \frac{P\left(A_{v-1} \text{ and } A_v\right)}{P\left(A_{v-1}\right)} = \frac{Q_v}{Q_{v-1}} \tag{11}$$

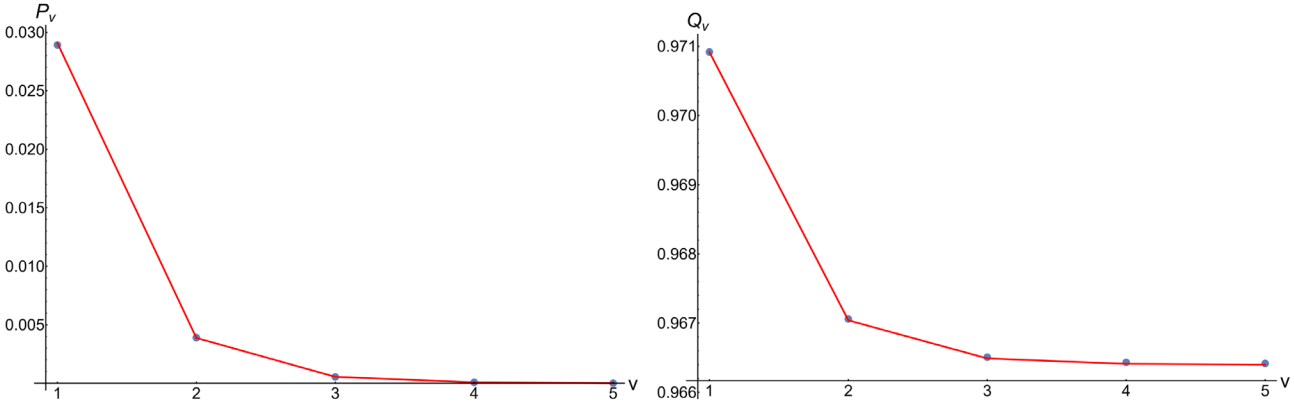

Figure 5: Again a neat agreement between the numerical result (blue points) and the analytical solution (red joined points), using a binomial distribution for P_v and Q_v.

Several Lineages

In this section, I will calculate the probability $R_{v,m}$ that $0 \le m \le n$ lineages are still present at generation v, knowing that there were $n \ge 1$ in the beginning. As explained in the introduction, this is indeed the probability $P(B_m \mid A_n)$ we are looking for. It is relabeled here for the sake of simplicity and clarity. Let's proceed again step by step and begin with $v = 1$ and $m = 0$:

$$\begin{aligned} R_{1,0} &= P(\text{the } 1^{\text{st}} \text{ mother has no girl and}\ldots \text{ and the } n^{\text{th}} \text{ mother has no girl}) \\ &= P(\text{a mother has no girl})^n = P_1^n \end{aligned}$$

With $m = 1$ one gets

$$\begin{aligned} R_{1,1} &= P(\text{the } 1^{\text{st}} \text{ mother has at least 1 girl and the others none or}\ldots\text{or} \\ &\quad \text{the } n^{\text{th}} \text{ mother has at least 1 girl and the others none}) \\ &= n\, P(\text{any mother has at least 1 girl and the others none}) \\ &= n\, P(\text{any mother has at least 1 girl})\, P(n-1 \text{ mothers have no girl}) \\ &= n\, Q_1 P_1^{n-1} \end{aligned}$$

For $m = 2$ we have

$$\begin{aligned} R_{1,2} &= P(\text{the } 1^{\text{st}} \text{ and } 2^{\text{nd}} \text{ mother have at least one girl and } n-2 \text{ none or}\ldots \text{ or} \\ &\quad \text{the } 1^{\text{st}} \text{ and the } 3^{\text{rd}} \text{ mother have at least one girl and } n-2 \text{ none or}\ldots \text{ or}) \\ &= \binom{n}{2} Q_1^2 P_1^{n-2} \end{aligned}$$

From this it is easy to guess the formula for an arbitrary m and $v = 1$:

$$R_{1,m} = \binom{n}{m} Q_1^m P_1^{n-m}$$

Let's now proceed analogously for $v = 2$:

$$R_{2,0} = P($$
1 lineage of the 1st gen. has survived and dies out at the 2nd or… or
k lineages of the 1st gen. have survived and all die out at the 2nd or … or
n lineages of the 1st gen. have survived and all die out at the 2nd$) = \sum_{k=1}^{n} a_k$

with

$$\begin{aligned} a_k &= P(k \text{ lineages of the 1st gen. have survived and all die out at the 2nd}) \\ &= P(k \text{ lineages of the 1st gen. have survived}) \, P(\text{all } k \text{ die out at the 2nd}) \\ &= R_{1,k} \, P(\text{all } k \text{ lineages die out at the 2nd gen.}) \\ &= R_{1,k} \, P(\text{any lineage dies out at the 2nd gen.})^k \end{aligned}$$

P(any lineage dies out at the 2nd gen.) = P_2/Q_1 according to equation 10 because it is a conditional probability, knowing that the lineage survived at the precedent generation, finally getting

$$a_k = R_{1,k}\left(\frac{P_2}{Q_1}\right)^k \text{ and } R_{2,0} = \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} R_{1,k}\left(\frac{P_2}{Q_1}\right)^k$$

For $m = 1$ we have

$$R_{2,1} = P($$
1 lineage of the 1st gen. has survived and also survives at the 2nd or… or
k lineages of the 1st gen. have survived and only the 1st survives at the 2nd or … or
k lineages of the 1st gen. have survived and only the k^{th} survives at the 2nd or … o
n lineages of the 1st gen. have survived and only the 1st survives at the 2nd or … or
n lineages of the 1st gen. have survived and only the n^{th} survives at the 2nd or … or)
$$= \sum_{k=1}^{n} a_k$$

with

$$a_k = P($$
k lineages of the 1st gen. have survived and only the 1st survives at the 2nd or … or
k lineages of the 1st gen. have survived and only the k^{th} survives at the 2nd)
$$\begin{aligned} &= k \, P(k \text{ lineages of the 1st gen. have survived}) \, P(1 \text{ of } k \text{ lineages survives at the 2nd}) \\ &= k \, R_{1,k} \, P(1 \text{ of } k \text{ lineages survives, the others dying out at the 2nd gen.}) \\ &= k \, R_{1,k} \, P(1 \text{ lineage survives at the 2nd gen.}) \, P(\text{the others dying out at the 2nd gen.}) \\ &= k \, R_{1,k} \, P(1 \text{ lineage survives at the 2nd gen.}) \, P(\text{any lineage dies out at the 2nd gen.})^{k-1} \end{aligned}$$

P(1 lineage survives at the 2nd gen.) and P(any lineage dies out at the 2nd gen.) are again conditional probabilities, knowing that all k lineages survived at the 1st generation, which are given by equations 10 and 11, yielding

$$a_k = k \, R_{1,k} \frac{Q_2}{Q_1}\left(\frac{P_2}{Q_1}\right)^{k-1} \text{ and } R_{2,1} = \sum_{k=1}^{n} a_k = \frac{Q_2}{Q_1} \sum_{k=1}^{n} k \, R_{1,k}\left(\frac{P_2}{Q_1}\right)^{k-1}$$

The case $m = 2$ is only slightly more complicated:

$$R_{2,2} = P($$

2 lin. have survived at the 1st gen. and those have survived at the 2nd gen. or… or
k lin. have survived at the 1st gen. and 2 of those have survived at the 2nd gen. or… or
n lin. have survived at the 1st gen. and 2 of those have survived at the 2nd gen.) $= \sum_{k=2}^{n} a_k$

with

$$\begin{aligned} a_k &= P(k \text{ lineages have survived at the } 1^{\text{st}} \text{ gen. and 2 of those have survived at the } 2^{\text{nd}} \text{ gen.}) \\ &= P(k \text{ lineages have survived at the } 1^{\text{st}} \text{ gen.})\, P(2 \text{ of } k \text{ lineages have survived at the } 2^{\text{nd}} \text{ gen.}) \\ &= R_{1,k}\, P(2 \text{ of } k \text{ lineages have survived at the } 2^{\text{nd}} \text{ gen. and } k-2 \text{ died out}) \\ &= R_{1,k} \binom{k}{2} \left(\frac{Q_2}{Q_1}\right)^2 \left(\frac{P_2}{Q_1}\right)^{k-2} \Rightarrow R_{2,2} = \sum_{k=2}^{n} a_k = \left(\frac{Q_2}{Q_1}\right)^2 \sum_{k=2}^{n} \binom{k}{2} R_{1,k} \left(\frac{P_2}{Q_1}\right)^{k-2} \end{aligned}$$

From this, it is straightforward to guess the general formula for an arbitrary *m*:

$$R_{2,m} = \left(\frac{Q_2}{Q_1}\right)^m \sum_{k=m}^{n} \binom{k}{m} R_{1,k} \left(\frac{P_2}{Q_1}\right)^{k-m}$$

From here, the general case for an arbitrary *v* is also evident:

$$R_{v,m} = \left(\frac{Q_v}{Q_{v-1}}\right)^m \sum_{k=m}^{n} \binom{k}{m} R_{v-1,k} \left(\frac{P_v}{Q_{v-1}}\right)^{k-m} \tag{12}$$

where $R_{0,m} = 0$ for $0 \le m < n$ and $R_{0,n} = 1$ because there are certainly n original mothers in the beginning. Analogously $Q_0 = 1$ because this is the probability that there is a single original mother. However, this formula is only valid for $m > 0$. For $m = 0$, we need to use the following formula to prevent that *k* starts with 0.

$$R_{v,0} = \sum_{k=1}^{n} R_{v-1,k} \left(\frac{P_v}{Q_{v-1}}\right)^k$$

$R_{v,m}$ can be tested numerically for different *d*, using a uniform distribution (see codes.nb). The results for $n = 4$ can be seen in figure 6, which shows that if *d* is low, the probability that only one lineage survives is higher than the others for the first generations. This is because the likelihood of a lineage becoming extinct is high when *d* is low. Thereby, it is more likely that the lineages die out early. This situation slowly inverses for *d* increasing. What is also very well visible again is the convergence of $R_{v,m}$ for increasing *v*.

When using a binomial distribution, the disparity of $R_{v,m}$ from very low to very high is even more manifest such that a logarithmic plot must be used to distinguish the data. As shown by figure 7, if *m* is increased from 1 to $n = 5$, the highest probability occurs if $m = n$ (left). In other words, the likelihood of a bottleneck is very low for high fertility rates, as mentioned in the introduction. A similar behavior occurs if $m = 1$ and *n* is increased (right), having again the highest probability if $m = n$.

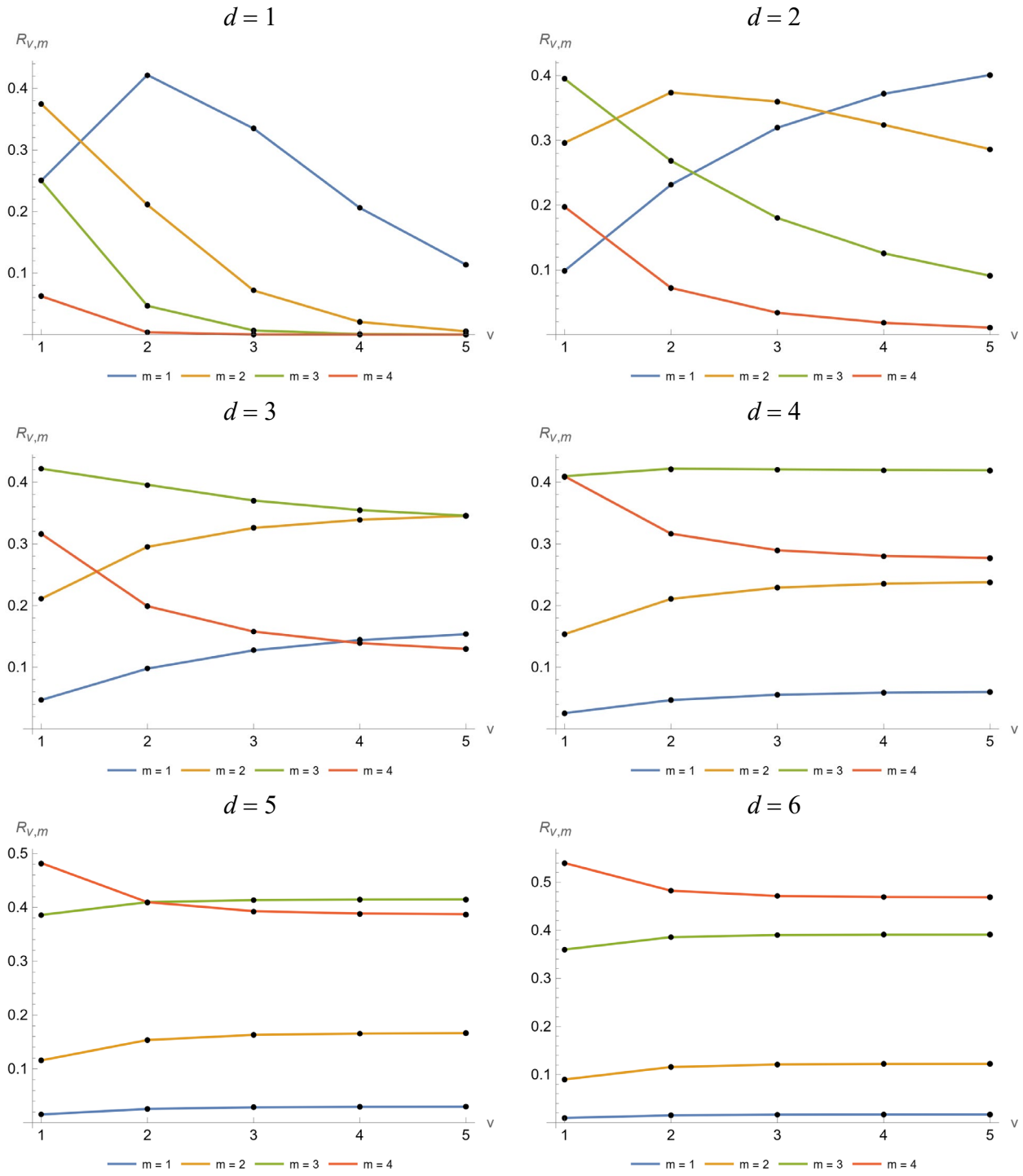

Figure 6: The probability $R_{v,m}$ that m lineages survive after v generations, knowing that there were $n = 4$ in the beginning and using a uniform distribution. The black points are the numerical results, while the colored joined points represent the analytical solution.

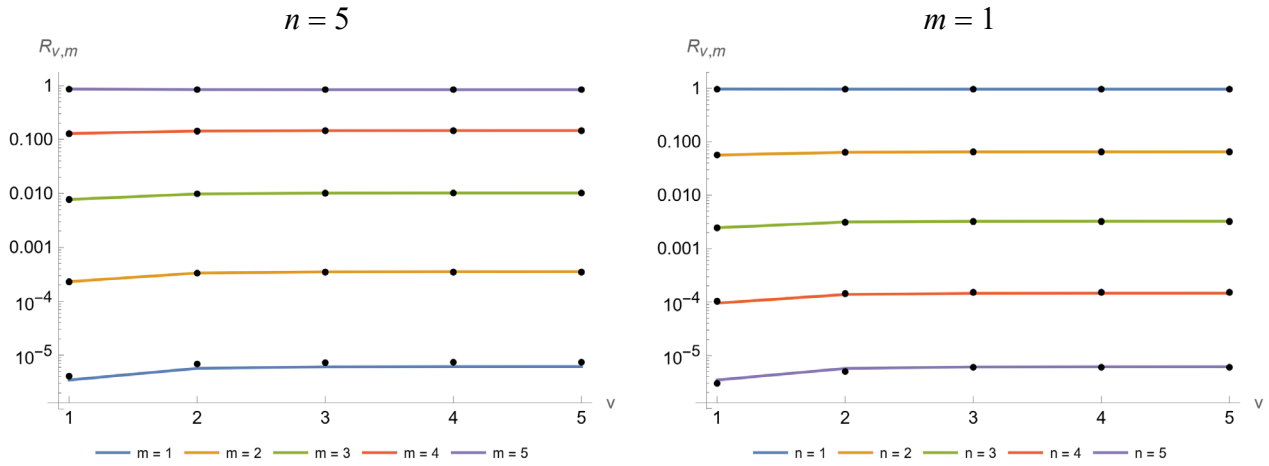

Figure 7: The probability $R_{v,m}$ that m lineages survive after v generations, using a binomial distribution. **Left:** $n = 5$ and m is increased. For $m = 1$, the numerical result (black points) is somewhat inaccurate because this case rarely occurs in the algorithm, which is why its relative frequency is approximate. In addition, the error is amplified because of the logarithmic plot. **Right:** $m = 1$ and n is increased.

Results

We are now able to calculate and discuss equation 3, that is, the probability $P(A_n \mid B_m)$ that there have been n lineages in the beginning, knowing that there are m at present. As there is indeed only one lineage at present, $m = 1$ will always be tacitly used. I will first take a uniform distribution for the b_i because it is interesting to see how this probability behaves for different d, which is difficult to show if they are a binomial distribution. I will therefore label

$$U = P(A_n \mid B_m) = \frac{P(B_m \mid A_n)}{\sum_{k=m}^{t} P(B_m \mid A_k)}$$

where $P(B_m \mid A_n) = R_{v,m}$ according to equation 12, which depends on the parameters v, m, n, t and d. These are too numerous to be able to discuss every interdependence between them. Nevertheless, as can be seen from figure 8, a general trend valid for every d is that if v increases, U_v decreases for small n and increases for greater n. What can be observed too is that the U_v seem to converge to a limit for every n, which will be discussed using a binomial distribution. This convergence is all the more rapid for d increasing from 1 to more realistic values up to 6. The most important aspect is that $U_{n=1}$ has the highest probabilities for every v and $d > 2$. Not only does it have the highest probabilities but it is also largely superior to 50%, which means that $U_{n=1} > U_{n>1}$. In other words, if it can be confirmed that the U_v really converge to a limit for increasing v, the probability that originally there was only 1 lineage is superior to the probability that there were more than 1 lineage.

Now that we know that $U_{n=1} > 50\%$ for $d = 6$ and $t = 7$ we can set $n = 1$ and check whether this is also the case for greater t. As shown by figure 9, U converges rapidly if t is increased from 2 to 8 for the same $d = 6$, the curves becoming almost indistinguishable for $t > 4$. Apparently, the convergence of $R_{v,m}$ if n is increased (fig. 7 right) also leads to the convergence of the sum in the denominator of U. What can also be seen from figure 9 is that U converges for increasing v, which is indeed a confirmation of what has been observed in the previous sections. This is very fortunate because a calculation of U over thousands of generations from the present of the mitochondrial Eve to today becomes unnecessary. In the contrary case, such a calculation would be impossible not only with a PC but also with a supercomputer.

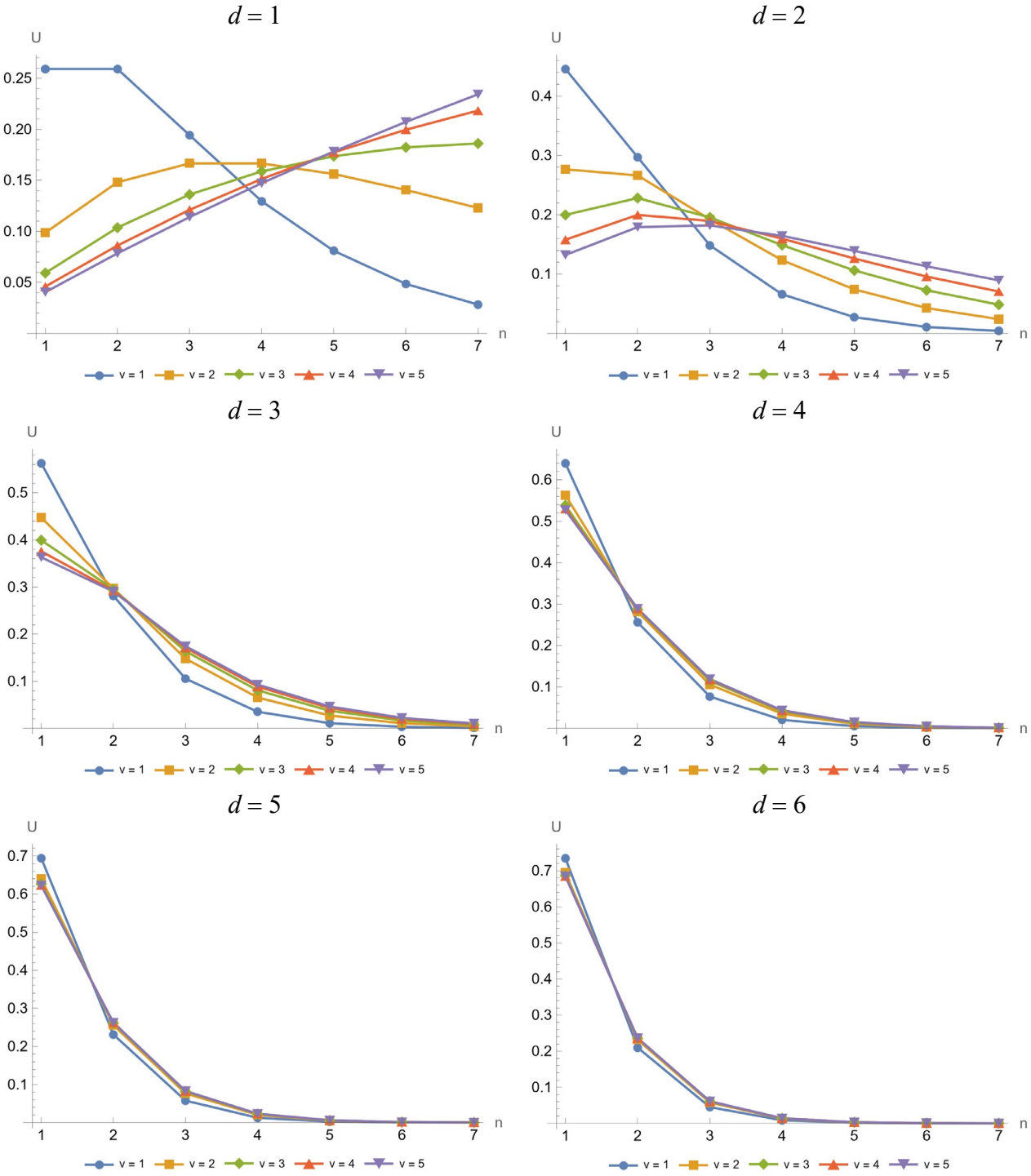

Figure 8: The probability U that there were initially n lineages, knowing that at present there is only one and using a uniform distribution for the b_i.

Let us now see if this convergence regarding both t and v also occurs with a binomial distribution. In this case, I will label $W = P(A_n \mid B_m)$, the letter W standing for a weighted distribution. As can be seen from figure 10, the convergence is even more rapid for both t and v, which allows for the conclusion that W is over 90% for very great t and v.

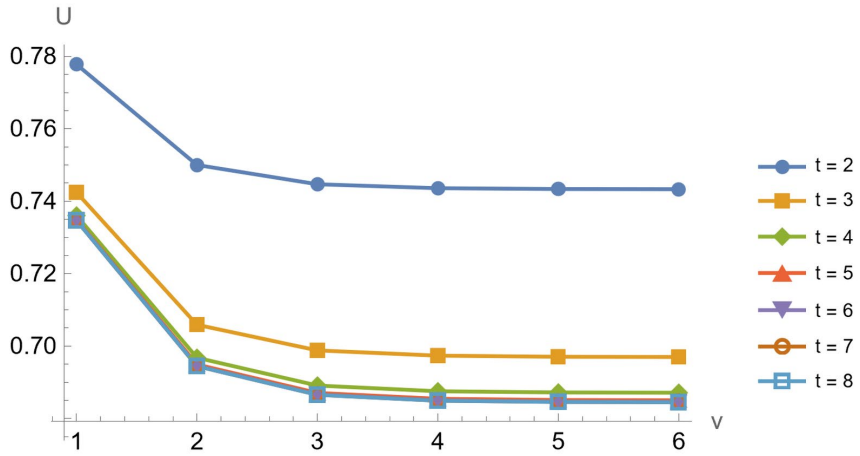

Figure 9: *U* converges rapidly for increasing *t* and *v* in the case where $d = 6$ and $n = 1$.

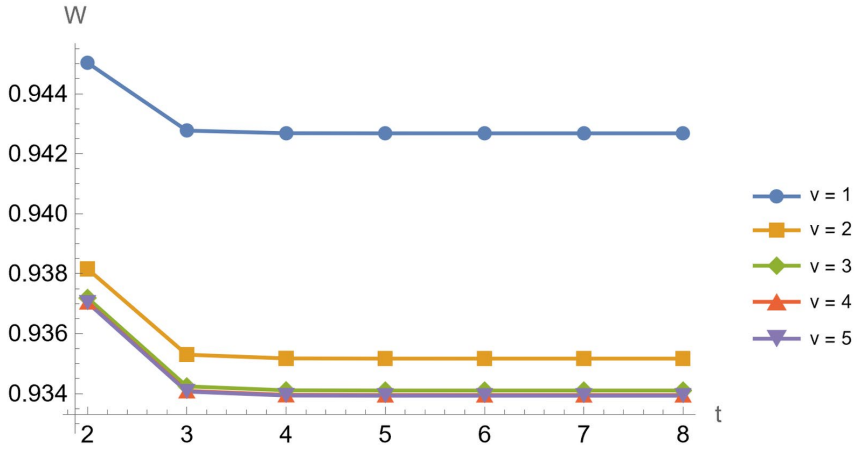

Figure 10: In the case of a binomial distribution, *W* also converges rapidly for increasing *t* and *v*, from which one can infer that *W* is over 90% for any *t* and *v*.

So the final statement is that the unlikeliness of a mitochondrial Eve, as claimed by most evolutionary biologists, is clearly refuted. What still can be discussed is how it is biologically (not mathematically) possible that monogenism can apparently emerge from a very large population without having to go through a bottleneck, an extinction of the other lineages being very unlikely.

When reasoning in evolutionary terms, one could resort to allopatric speciation theorized by Ernst Mayr (1954 pp. 157-180), according to which speciation cannot happen in a large population due to constant mixing up of the gene pool. Therefore, this concept predicts that only when a small group of individuals splits from the main group and remains separated for a long time due to a geographic barrier, it acquires enough genetic changes to become a new species. However, as I have shown (Utiger 2020), there is by far not enough time for even the most favorable case of such a scenario, that is, a single couple split from the main group.

So something very special must have happened at the time of the mitochondrial Eve, something that cannot be understood in evolutionary terms, which seems to baffle many scholars, beating around the bush instead of putting a proper name to the mystery. Often, they stick to Gould and Eldredge's "punctuated equilibrium", as mentioned in the introduction. Tattersall (2009) calls it "a single change in gene regulation", while Dawkins (2006 p. 230) correctly dares to declare that "the only alternative explanation… is divine creation", only to reject his statement immediately in the next sentence, without giving further justifications…

References

Ayala, F. J. (1995). The Myth of Eve: Molecular Biology and Human Origins. *Science, 270*(5244), 1930-1936.

Bentley, G. R., Goldberg, T., & Jasieńska, G. Z. (1993). The Fertility of Agricultural and Non-Agricultural Traditional Societies. *Population Studies, 47*(2), 269-281.

Bräuer, G. (2015). Origin of Modern Humans. In W. Henke, & I. Tattersall (Eds.), *Handbook of Paleoanthropology* (2nd ed., pp. 2299-2330). Springer.

Brown, M. H. (1990). *The Search For Eve.* New York: Harper & Row.

Callaway, E. (2017). Oldest Homo Sapiens Fossil Claim Rewrites Our Species' History. *Nature, 546*, 289-293.

Cann, R. L., Stoneking, M., & Wilson, A. C. (1987). Mitochondrial DNA and Human Evolution. *Nature, 325*(6099), 31-36.

Coale, A. J. (1972). *The Growth and Structure of Human Populations: A Mathematical Investigation.* Princeton University Press.

Cohen, W. B. (1980). *The French Encounter With Africans: White Response to Blacks, 1530-1880.* Indiana University Press.

Dawkins, R. (2006). *The Blind Watchmaker.* London: Penguin Books.

Dawkins, R., & Wong, Y. (2017). *An Ancestor's Tale: A Pilgrimage to the Dawn of Life.* London: Weidenfeld & Nicolson.

Foley, R. (1995). *Humans Before Humanity.* Oxford: Blackwell Publishers Ltd.

Gibbons, A. (2017). *World's Oldest Homo Sapiens Fossils Found in Morocco*. Retrieved March 2024, from Science: https://www.science.org/content/article/world-s-oldest-homo-sapiens-fossils-found-morocco

Hublin, J.-J., Ben-Ncer, A., Bailey, S. E., Freidline, S. E., Neubauer, S., Skinner, M. M., . . . Gunz, P. (2017). New Fossils From Jebel Irhoud, Morocco and the Pan-African Origin of Homo Sapiens. *Nature, 546*(7657), 289-292.

Jackson, J., & Weidman, N. (2004). *Race, Racism, and Science: Social Impact and Interaction.* ABC-CLIO.

Klein, R. G. (2009). *The Human Career: Human Biological and Cultural Origins* (3rd ed.). The University of Chicago Press.

Mayr, E. (1954). Change of Genetic Environment and Evolution. In J. Huxley, A. Hardy, & E. Ford (Eds.), *Evolution as a Process* (pp. 157-180). London: Allen & Unwin.

Meneganzin, A., Pievani, T., & Manzi, G. (2022). Pan-Africanism Vs. Single-Origin of Homo Sapiens: Putting the Debate in the Light of Evolutionary Biology. *Evolutionary Anthropology: Issues, News, and Reviews, 31*(4), 199-212.

Reichholf, J. H. (1998). *Das Rätsel der Menschwerdung: Die Entstehung des Menschen im Wechselspiel mit der Natur* (4th ed.). München: Deutscher Taschenbuch Verlag.

Scerri, E. M., Thomas, M. G., Manica, A., Gunz, P., Stock, J. T., Stringer, C., . . . d'Errico, F. (2018). Did Our Species Evolve in Subdivided Populations Across Africa, and Why Does It Matter? *Trends in Ecology & Evolution, 33*(8), 582-594.

Stringer, C., & Galway-Witham, J. (2017). On the Origin of Our Species. *Nature, 546*(7657), 212-214.

Sykes, B. (2001). *The Seven Daughters of Eve.* New York: W. W. Norton & Company.

Sykes, B. (2003). *Adam's Curse: A Future Without Men.* Bantam Press.

Tattersall, I. (2009). Human Origins: Out of Africa. *Proceedings of the National Academy of Sciences, 106*(38), 16018-16021.

Utiger, U. (2020, April). *The Probability of a Nucleotide Sequence Formed by Random Mutations.* Retrieved from http://historycycles.org/nucleotides.pdf