The matchmaking paradox: a statistical explanation

The matchmaking paradox: a statistical explanation

Iddo I Eliazar[1] and Igor M Sokolov[2]

[1] Department of Technology Management, Holon Institute of Technology, PO Box 305, Holon 58102, Israel
[2] Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, D-12489 Berlin, Germany

E-mail: eliazar@post.tau.ac.il and igor.sokolov@physik.hu-berlin.de

Abstract
Medical surveys regarding the number of heterosexual partners per person yield different female and male averages—a result which, from a physical standpoint, is impossible. In this paper we term this puzzle the 'matchmaking paradox', and establish a statistical model explaining it. We consider a bipartite graph with N male and N female nodes ($N \gg 1$), and B bonds connecting them ($B \gg 1$). Each node is associated a random 'attractiveness level', and the bonds connect to the nodes randomly—with probabilities which are proportionate to the nodes' attractiveness levels. The population's average bonds-per-nodes B/N is estimated via a sample average calculated from a survey of size n ($n \gg 1$). A comprehensive statistical analysis of this model is carried out, asserting that (i) the sample average well estimates the population average if and only if the attractiveness levels possess a finite mean; (ii) if the attractiveness levels are governed by a 'fat-tailed' probability law then the sample average displays wild fluctuations and strong skew—thus providing a statistical explanation to the matchmaking paradox.

PACS number: 02.50.−r

1. Introduction

In a *Nature* News and Views comment titled 'Sexual behaviour unsurveyed', *Nature* editor J Maddox stated that [1]: '*The figures so far show that the average number of heterosexual partners of men in the course of a lifetime is 11.0 and of women 2.9*'. In a response to this comment, physicist S J Gurman referred to the afore-mentioned statistics and wrote [2]: '*A heterosexual union is analogous to a heteronuclear chemical bond, and the total number must be the same if viewed from the male or female end*'. In this paper we term the discrepancy between the male statistics (11.0) and the female statistics (2.9), the '*matchmaking paradox*', and present a statistical explanation of this paradox.

In the context of heterosexual relations, a given human population can be regarded as a bipartite graph consisting of $2N$ nodes and B bonds[3]. The nodes represent the population of men and women—N male nodes consisting the left side of the bipartite graph and N female nodes consisting the graph's right side. The bonds represent heterosexual partnerships— a bond connecting male i to female j indicating a heterosexual partnership between these individuals. Clearly, the average number of heterosexual partners—observed either from the male or female side of the bipartite graph—is given by the ratio

$$\rho = \frac{B}{N}. \tag{1}$$

Hence, from the entire-population perspective, Gurman's assertion is correct: the male and female statistics of average heterosexual partners must coincide.

Yet, in order to obtain the population average ρ one has to hold a census of the entire population—which, in practice, seldom happens. In reality, the population average ρ is estimated via a sample survey. Namely, a subset of size n of the male (female) population is sampled at random. If X represents the random number of bonds connected to the sample-subset's nodes, then the average number of heterosexual partners—with respect to the sample survey—is given by the ratio

$$\hat{\rho} = \frac{X}{n}. \tag{2}$$

Hence, in reality the statistical figure we obtain is the sample average $\hat{\rho}$ rather than the population average ρ.

From a statistical perspective, the matchmaking paradox boils down to the difference between the sample average $\hat{\rho}$ and the population average ρ. If the sample average $\hat{\rho}$ well estimates the population average ρ—Gurman's assertion is indeed correct. However, if the sample average $\hat{\rho}$ displays significant random fluctuations, then very different statistical figures can be obtained—thus providing a statistical explanation to the matchmaking paradox.

In this research we consider an 'attractiveness model' for the random connection of the bonds to the male and female sides of the bipartite graph: each member of the population is assigned a random 'attractiveness level', and the B bonds connect males and females according to their attractiveness levels. Namely, the probability that a specific bond connects between male i and female j is proportional to the product of their attractiveness levels. The male and female attractiveness levels are assumed to be independent and identically distributed (iid) positive-valued random variables, and the bonds are assumed to connect independently of each other.

The attractiveness model turns out to provide a comprehensive statistical explanation to the matchmaking paradox—yielding scenarios in which the sample average $\hat{\rho}$ is either close to the population average ρ, or fluctuates significantly far from it. The detailed description of the attractiveness model, as well as its preliminary statistical analysis, is presented in Section 2. The statistical analysis of the sample average $\hat{\rho}$ follows in sections 3 and 4. Informally speaking, the statistical analysis asserts that the sample average $\hat{\rho}$ is close to the population average ρ if and only if the attractiveness levels possess a finite mean.

- If the attractiveness levels possess a finite mean, then the sample average $\hat{\rho}$ well estimates the population average ρ—see section 3.
- If the attractiveness levels are governed by a 'fat-tailed' probability law, then the sample average $\hat{\rho}$ displays universal fluctuations and fails to estimates the population average ρ—see section 4.

---

[3] Namely, the bipartite graph is composed of two sides—the left and the right side—each side consisting of N nodes. And, the bonds connect nodes on different sides of the bipartite graph—each bond connected to exactly one node on each side.

The sharp phase transition taking place when passing from the finite-mean case to the infinite-mean case is much reminiscent and analogous to the phenomena of *ergodicity breaking* reported in various physical systems—in which sample means, which usually behave regularly, behave anomalously and display universal fluctuations [3–5].

The 'attractiveness levels' of human populations are, in effect, their underlying 'fitness landscapes'. The natural emergence of fat-tailed probability laws in 'fitness landscapes' can be explained by several theoretical models including self-organized criticality [6, 7], and statistical resilience to random perturbations [8]. For other analysis methods of bipartite graphs arising in the context of social interactions, the readers are referred to [9].

2. Setting and preliminary analysis

In this section we (i) describe, in detail, the 'attractiveness model' noted in the introduction; (ii) conduct a preliminary statistical analysis of the sample average $\hat{\rho}$. Henceforth, we consider only one side—either the male (left side) or the female (right side)—of the bipartite graph representing the heterosexual partnerships.

2.1. The attractiveness model

Consider a male (female) population of size $N \gg 1$, its members labeled $k = 1, \ldots, N$. Each member k of the population has a positive-valued attractiveness level A_k. The population's average attractiveness level is $\bar{A} = (A_1 + \cdots + A_N)/N$, and the relative attractiveness of member k is $A_k/\bar{A}$.

A fixed number $B \gg 1$ of bonds are connected to the population members. Each bond— independently of all other bonds—chooses at random a single population member and connects to it. The probability π_k of a bond choosing member k and connecting to it is proportional to the member's relative attractiveness $A_k/\bar{A}$. Since the probabilities $\{\pi_k\}_{k=1}^{N}$ need to sum up to unity we have $\pi_k = A_k/(A_1 + \cdots + A_N)$.

A subset of size $n \gg 1$ is sampled at random from the entire population—the sample size being a fraction f of the population size

$$f = \frac{n}{N}. \tag{3}$$

With no loss of generality, we take the sample to be the subset $\{1, \ldots, n\}$. The probability of a bond to connect to the sample—i.e. to one of the population members $k = 1, \ldots, n$—is given by

$$P = \pi_1 + \cdots + \pi_n = \frac{A_1 + \cdots + A_n}{A_1 + \cdots + A_N}. \tag{4}$$

Consequently, the number of bonds X connected to the subset $\{1, \ldots, n\}$ is random and binomially distributed: $X \sim \mathbf{Bin}(B, P)$, namely,

$$\Pr(X = x) = \frac{B!}{x!(B-x)!} P^x (1-P)^{B-x} \tag{5}$$

$(x = 0, 1, \ldots, B)$.

The average number of bonds per population member is given by the population average ρ—defined in equation (1). Within the sample-subset, however, the average number of bonds per population member is given by the sample average $\hat{\rho}$—defined in equation (2). Viewed from the male and female sides of the bipartite graph, the attractiveness model is, in fact, a random allocation model. Indeed, a fixed 'budget' of size B is randomly allocated amongst a population of size N. The average budget per population member is given by the ratio

$\rho = B/N$. Yet within a population subset of size n the average budget per population member is given by the ratio $\hat{\rho} = X/n$—the random variable X denoting the budget allocated to the subset.

The attractiveness levels $\{A_k\}_{k=1}^{N}$ are henceforth assumed to be iid. copies of a generic—random and positive-valued—attractiveness level A. As we shall show, the statistical behavior of the sample average $\hat{\rho}$ depends on the probability law of the attractiveness level A. And, the statistical behavior of the sample average $\hat{\rho}$ displays phase transitions as (i) the variance of the attractiveness level A changes from finite to infinite; (ii) the mean of the attractiveness level A changes from finite to infinite.

2.2. Preliminary analysis of the sample average

Since the number of bonds B is large ($B \gg 1$), and since the number of bonds X connected to the sample-subset—given the random probability P—is **Bin**(B, P) distributed, then given the random probability P, the Gaussian CLT [10] implies that the random ratio X/B is well approximated by a Gaussian law with mean P and variance $P(1 - P)/B$. Making use of the ratio-relations of equations (1) and (3) we arrive at the following, P-conditional, Gaussian approximation of the sample average:

$$\hat{\rho} \approx \frac{\rho}{f} P + \frac{1}{\sqrt{N}} \left(\frac{\sqrt{\rho}}{f} \sqrt{P(1 - P)} \right) Z_1,$$ (6)

where Z_1 is a Gaussian random variable with zero mean and unit variance.

From equation (6) it is evident that the distribution of the random probability P—given by equation (4)—governs the behavior of the sample average $\hat{\rho}$. The statistics of the random probability P, in turn, display markedly different statistical behaviors in the following scenarios.

- *The regular Gaussian scenario*: the attractiveness level A possesses both a finite mean $\mu < \infty$ and a finite variance $\sigma^2 < \infty$.
- *The regular Lévy scenario*: the probability tail of the attractiveness level A follows a power-law decay with exponent $1 < \alpha < 2$—implying a finite mean $\mu < \infty$ and an infinite variance $\sigma^2 = \infty$.
- *The anomalous Lévy scenario*: the probability tail of the attractiveness level A follows a power-law decay with exponent $0 < \alpha < 1$—implying both an infinite mean $\mu = \infty$ and an infinite variance $\sigma^2 = \infty$.

The anomalous Lévy scenario is what we termed in both the abstract and the introduction the case of 'fat-tailed' probability laws. We turn now to analyze these three different scenarios, and explore how they affect the Gaussian approximation of equation (6).

3. Analysis of the regular scenarios

In this section we study the regular Gaussian and Lévy scenarios—in which both the attractiveness level A possesses a finite mean.

3.1. The regular Gaussian scenario

In the regular Gaussian scenario the attractiveness level A possesses both a finite mean $\mu < \infty$ and a finite variance $\sigma^2 < \infty$. In this scenario the Gaussian CLT [10] provides the following approximation to the random probability P:

$$P \approx f + \frac{1}{\sqrt{N}} \left(\frac{\sigma}{\mu} \sqrt{f(1 - f)} \right) Z_2,$$ (7)

where Z_2 is a Gaussian random variable with zero mean and unit variance. Combining together the Gaussian approximations of equation (6), and the Gaussian approximations of equation (7), we arrive at the following Gaussian approximation of the sample average:

$$\hat{\rho} \approx \rho + \frac{1}{\sqrt{N}} \cdot \sqrt{\left(\frac{1-f}{f}\right)\left(\rho + \left(\rho\frac{\sigma}{\mu}\right)^2\right)} \cdot Z, \tag{8}$$

where Z is Gaussian random variable with zero mean and unit variance.

The Gaussian approximation of equation (8) asserts that the sample average $\hat{\rho}$ well estimates the population average ρ. Specifically, the fluctuations of the sample average $\hat{\rho}$ around the population average ρ are additive, their amplitudes are of order $O(1/\sqrt{N})$, and their probability law is Gaussian. A detailed derivation of equations (7) and (8) is given in the appendix.

3.2. The regular Lévy scenario

In the regular Lévy scenario the probability tails of the attractiveness level A follow a power-law decay with exponent $1 < \alpha < 2$—implying a finite mean $\mu < \infty$ and an infinite variance $\sigma^2 = \infty$. In this scenario the Lévy CLT [10] provides the following approximation to the random probability P:

$$P \approx f + \frac{1}{N^{1-1/\alpha}} \cdot \frac{1}{\mu}(f(1-f)^\alpha + (1-f)f^\alpha)^{1/\alpha} \cdot L, \tag{9}$$

where L is a two-sided Lévy random variable with exponent α and skew

$$\beta = \frac{f^{\alpha-1} - (1-f)^{\alpha-1}}{f^{\alpha-1} + (1-f)^{\alpha-1}}. \tag{10}$$

Namely, L is a real-valued random variable whose probability law is characterized by the Fourier transform:

$$\langle\exp(i\theta L)\rangle = \exp\left(-|\theta|^\alpha \cdot \left(1 - i\beta\,\mathrm{sign}(\theta)\tan\left(\frac{\pi}{2}\alpha\right)\right)\right) \tag{11}$$

(θ real).

Combining together the Gaussian approximation of equation (6), and the Lévy approximation of equation (9), we arrive at the following Lévy approximation of the sample average:

$$\hat{\rho} \approx \rho + \frac{1}{N^{1-1/\alpha}} \cdot \frac{\rho}{\mu f}(f(1-f)^\alpha + (1-f)f^\alpha)^{1/\alpha} \cdot L, \tag{12}$$

The Lévy approximation of equation (12) asserts that the sample average $\hat{\rho}$ well estimates the population average ρ. Specifically, the fluctuations of the sample average $\hat{\rho}$ around the population average ρ are additive, their amplitudes are of order $O(1/N^{1-1/\alpha})$, and their probability law is Lévy (with exponent α and skew β given by equation (10)). A detailed derivation of equations (9) and (12) is given in the appendix.

4. Analysis of the anomalous scenario

In this section we study the anomalous Lévy scenario—in which the attractiveness level A is governed by a 'fat-tailed' probability law. Namely, the probability tail of the attractiveness level A follows a power-law decay with exponent $0 < \alpha < 1$—implying both an infinite mean $\mu = \infty$ and an infinite variance $\sigma^2 = \infty$.

4.1. The anomalous Lévy scenario

In the anomalous Lévy scenario the Lévy CLT [10] provides the following approximation to the random probability P:

$$P \approx \frac{f^{1/\alpha}}{f^{1/\alpha} + (1 - f)^{1/\alpha} R}, \tag{13}$$

where R is the quotient of two independent one-sided Lévy random variables with exponent $0 < \alpha < 1$. Namely, R is a positive-valued random variable whose probability law is characterized by the Laplace transform [11]:

$$\langle \exp(-\theta R) \rangle = \sum_{m=0}^{\infty} \frac{(-\theta^{\alpha})^m}{\Gamma(1 + \alpha m)} \tag{14}$$

($\theta \geqslant 0$). The right-hand side of equation (14) is the Mittag–Leffler function of index α, evaluated at the point $-\theta^{\alpha}$ [12].

Combining together the Gaussian approximation of equation (6), and the Lévy approximation of equation (13), we arrive at the following Lévy approximation of the sample average:

$$\hat{\rho} \approx \rho \cdot \xi, \tag{15}$$

where

$$\xi = \frac{f^{(1/\alpha)-1}}{f^{1/\alpha} + (1 - f)^{1/\alpha} R}. \tag{16}$$

The Lévy approximation of equation (15) asserts that the sample average $\hat{\rho}$ is not a good estimate of the population average. Specifically, the fluctuations of the sample average $\hat{\rho}$ around the population average ρ are multiplicative and of order $O(1)$. A detailed derivation of equations (13) and (15) is given in the appendix.

4.2. Fluctuation analysis

In fact, in the anomalous Lévy scenario the sample average $\hat{\rho}$ admits wild fluctuations of the population average ρ. This follows from the statistical behavior of the random multiplicative factor ξ defined in equation (16), and which assumes values in the range $0 < \xi < 1/f$. Indeed, a comprehensive statistical analysis carried out in [11] yields the following facts regarding the random multiplicative factor ξ:

(i) mean:

$$\langle \xi \rangle = 1; \tag{17}$$

(ii) variance:

$$\mathbf{Var}[\xi] = (1 - \alpha) \frac{1 - f}{f}; \tag{18}$$

(iii) lower-bound probability asymptotics:

$$\Pr(\xi < \delta) \underset{\delta \to 0}{\sim} \frac{\sin(\pi \alpha)}{\pi \alpha} \frac{1 - f}{f^{1-\alpha}} \cdot \delta^{\alpha}; \tag{19}$$

(iv) upper-bound probability asymptotics:

$$\Pr\left(\xi > \frac{1}{f} - \delta\right) \underset{\delta \to 0}{\sim} \frac{\sin(\pi \alpha)}{\pi \alpha} \frac{f^{1+\alpha}}{1 - f} \cdot \delta^{\alpha}. \tag{20}$$

Equations (19) and (20) assert that the random multiplicative factor ξ concentrates a high probability at the boundaries of its range—hence exhibiting wild fluctuations. Moreover, since the random variable R is the quotient of two iid random variables, its median is unity: $\mathbf{Med}[R] = 1$. Combining this observation with equation (16) implies, in turn, that the median of the random multiplicative factor ξ is given by

$$\mathbf{Med}[\xi] = \frac{f^{(1/\alpha)-1}}{f^{1/\alpha} + (1-f)^{1/\alpha}} \begin{cases} < \langle \xi \rangle & \text{if} \quad f < 1/2; \\ = \langle \xi \rangle & \text{if} \quad f = 1/2; \\ > \langle \xi \rangle & \text{if} \quad f > 1/2. \end{cases} \tag{21}$$

In reality, the sample sizes (n) are significantly smaller than the population size (N)—implying a small sample fraction $f = n/N$. In such a case the variance $\mathbf{Var}[\xi]$ is significantly larger than the mean ($\langle \xi \rangle = 1$), and the median $\mathbf{Med}[\xi]$ is significantly smaller than the mean ($\langle \xi \rangle = 1$). Indeed, setting $f = 1/(m+1)$ we obtain that

$$\mathbf{Var}[\xi] = (1-\alpha)m \quad \text{and} \quad \mathbf{Med}[\xi] = \frac{1+m}{1+m^{1/\alpha}} \approx \frac{1}{m^{(1/\alpha)-1}}. \tag{22}$$

Namely, in the prevalent cases of small sample fractions the random multiplicative factor ξ displays wild fluctuations, and is also strongly skewed to the left (i.e. strongly skewed toward the origin 0). For example, if $\alpha = 1/2$ and $m = 500$ then the variance is 250 times larger than the mean, and the median is 500 times smaller than the mean.

5. Conclusions

This research embarked from the 'matchmaking paradox': empirical evidence indicating different male and female averages of heterosexual partners—evidence which, from a physical standpoint, is impossible. To resolve this paradox we modeled heterosexual partnerships by a bipartite graph consisting of N male nodes, N female nodes, and B bonds connecting the male and female nodes (each bond connecting exactly one male node to one female node). The scattering of the bonds was considered random: each node was associated with an attractiveness level, and the bonds were connected to the nodes randomly—the probability of a bond to connect between male i and female j being proportional to the product of their attractiveness levels. The male and female attractiveness levels were assumed to be iid random variables (governed by possibly the different male and female probability laws), and the bonds were assumed to connect independently of each other. Viewed from both the male and female sides of the bipartite graph, the aforementioned attractiveness model is a random allocation model: allocating a 'budget' of B bonds amongst a population of size N—based on the attractiveness levels of the population members.

The average number of bonds per node—i.e. the average number of male (female) heterosexual partnerships, or the average budget per population member—is given by the ratio $\rho = B/N$. The goal of this research was to conclude whether or not the population average can be well estimated by a sample average $\hat{\rho}$—the average number of bonds per node calculated from a random sample of n male (female) nodes. Conducting a comprehensive statistical analysis of the sample average $\hat{\rho}$ we categorized three different statistical scenarios.

The regular Gaussian scenario: in which the attractiveness level A possesses both a finite mean and a finite variance. In this scenario the sample average $\hat{\rho}$ well estimates the population average ρ: the fluctuations of the sample average $\hat{\rho}$ around the population average ρ are additive, their amplitudes are of order $O(1/\sqrt{N})$, and their probability law is Gaussian.

The regular Lévy scenario: in which the probability tail of the attractiveness level A follows a power-law decay with exponent $1 < \alpha < 2$. In this scenario, the attractiveness level

A possesses a finite mean and an infinite variance, and the sample average $\hat{\rho}$ well estimates the population average ρ: the fluctuations of the sample average $\hat{\rho}$ around the population average ρ are additive, their amplitudes are of order $O(1/N^{1-1/\alpha})$, and their probability law is Lévy.

The anomalous Lévy scenario: in which the probability tail of the attractiveness level A follows a power-law decay with exponent $0 < \alpha < 1$. In this scenario, the attractiveness level A possesses both an infinite mean and an infinite variance, and the sample average $\hat{\rho}$ fails to estimate the population average ρ: the fluctuations of the sample average $\hat{\rho}$ around the population average ρ are multiplicative and wild. Moreover, in the case of small sample fractions $f = n/N$, the sample average $\hat{\rho}$ is also strongly skewed toward 0.

These results provide us with a panoramic statistical explanation of the matchmaking paradox—pinpointing the cases in which this paradox does not arise, the cases in which it does arise and the boundaries between these cases. From a more general perspective, these results provide us with a statistical distinction between regular and anomalous 'attractiveness-based' allocations.

Acknowledgments

Appendix

A.1. The regular Gaussian scenario: derivation of equations (7) and (8)

Note that the random probability of equation (4) can be rewritten in the form

$$P = \frac{1}{1 + \frac{A_{n+1} + \cdots + A_N}{A_1 + \cdots + A_n}}. \tag{A.1}$$

Further note that

$$\frac{A_{n+1} + \cdots + A_N}{A_1 + \cdots + A_n} = \frac{N-n}{n} \frac{\frac{A_{n+1} + \cdots + A_N}{N-n}}{\frac{A_1 + \cdots + A_n}{n}}. \tag{A.2}$$

The random variables $\{A_k\}_{k=1}^N$ are iid with finite mean and variance; $n \gg 1$ and $N - n \gg 1$. Hence, the Gaussian CLT [10] implies that

$$\frac{A_1 + \cdots + A_n}{n} \approx \mu + \frac{\sigma}{\sqrt{n}} \xi_1, \tag{A.3}$$

and

$$\frac{A_{n+1} + \cdots + A_N}{N-n} \approx \mu + \frac{\sigma}{\sqrt{N-n}} \xi_2, \tag{A.4}$$

where ξ_1 and ξ_2 are independent Gaussian random variables with zero mean and unit variance.

Substituting equations (A.3) and (A.4) into equation (A.2), while using equation (3), yields

$$\begin{aligned}
\frac{A_{n+1} + \cdots + A_N}{A_1 + \cdots + A_n} &\approx \frac{1-f}{f} \frac{\mu + \frac{\sigma}{\sqrt{1-f}\sqrt{N}} \xi_2}{\mu + \frac{\sigma}{\sqrt{f}\sqrt{N}} \xi_1} \\
&= \frac{1-f}{f} \frac{1 + \left(\frac{\sigma}{\mu\sqrt{1-f}} \xi_2\right) \frac{1}{\sqrt{N}}}{1 + \left(\frac{\sigma}{\mu\sqrt{f}} \xi_1\right) \frac{1}{\sqrt{N}}}.
\end{aligned} \tag{A.5}$$

Moreover,

$$
\frac{1 + \left(\frac{\sigma}{\mu\sqrt{1-f}}\xi_2\right)\frac{1}{\sqrt{N}}}{1 + \left(\frac{\sigma}{\mu\sqrt{f}}\xi_1\right)\frac{1}{\sqrt{N}}} \approx \left(1 + \left(\frac{\sigma}{\mu\sqrt{1-f}}\xi_2\right)\frac{1}{\sqrt{N}}\right)\left(1 - \left(\frac{\sigma}{\mu\sqrt{f}}\xi_1\right)\frac{1}{\sqrt{N}}\right)
$$

$$
\approx 1 + \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu\sqrt{1-f}}\xi_2 - \frac{\sigma}{\mu\sqrt{f}}\xi_1\right)
$$

$$
= 1 + \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\left(\frac{\xi_2}{\sqrt{1-f}} - \frac{\xi_1}{\sqrt{f}}\right)\right)
$$

$$
= 1 - \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\frac{Z_2}{\sqrt{f(1-f)}}\right), \tag{A.6}
$$

where Z_2 is a Gaussian random variable with zero mean and unit variance.

Substituting equation (A.6) into equation (A.5) gives

$$
\frac{A_{n+1} + \cdots + A_N}{A_1 + \cdots + A_n} \approx \frac{1-f}{f}\left(1 - \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\frac{Z_2}{\sqrt{f(1-f)}}\right)\right). \tag{A.7}
$$

In turn, substituting equation (A.7) into equation (A.1) yields

$$
P \approx \frac{1}{1 + \frac{1-f}{f}\left(1 - \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\frac{Z_2}{\sqrt{f(1-f)}}\right)\right)} = \frac{f}{1 - \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\sqrt{\frac{1-f}{f}}\right)Z_2}
$$

$$
\approx f\left(1 + \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\sqrt{\frac{1-f}{f}}\right)Z_2\right). \tag{A.8}
$$

Hence, we conclude that the random probability P is approximated by

$$
P \approx f + \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\sqrt{f(1-f)}\right)Z_2, \tag{A.9}
$$

where Z_2 is a Gaussian random variable with zero mean and unit variance.

Substituting equation (A.9) into equation (6) we arrive at

$$
\hat{\rho} \approx \frac{\rho}{f}\left(f + \frac{1}{\sqrt{N}}\left(\frac{\sigma}{\mu}\sqrt{f(1-f)}\right)Z_2\right) + \frac{1}{\sqrt{N}}\left(\frac{\sqrt{\rho}}{f}\sqrt{f(1-f)}\right)Z_1
$$

$$
\approx \rho + \frac{1}{\sqrt{N}}\sqrt{\frac{1-f}{f}}\left(\sqrt{\rho}Z_1 + \rho\frac{\sigma}{\mu}Z_2\right), \tag{A.10}
$$

where Z_1 and Z_2 are independent Gaussian random variables with zero mean and unit variance. (The Gaussian random variable Z_1 captures the randomness stemming from binomial scattering of the bonds, and is independent of the attractiveness levels. On the other hand, the Gaussian random variable Z_2 captures the randomness of the random probability P—which, in turn, stems from the attractiveness levels—and is independent of the bonds' binomial scattering.) Finally, equation (A.10) implies that

$$
\hat{\rho} \approx \rho + \frac{1}{\sqrt{N}} \cdot \sqrt{\left(\frac{1-f}{f}\right)\left(\rho + \left(\rho\frac{\sigma}{\mu}\right)^2\right)} \cdot Z, \tag{A.11}
$$

where Z is Gaussian random variable with zero mean and unit variance.

A.2. The regular Lévy scenario: derivation of equations (9) and (12)

Recall that a random variable ξ is a two-sided Lévy random variable with exponent α ($1 < \alpha < 2$) and skew β ($-1 \leqslant \beta \leqslant 1$) if it is characterized by the Fourier transform [10]:

$$\langle \exp(i\theta\xi) \rangle = \exp\left(-|\theta|^{\alpha} \cdot \left(1 - i\beta \, \mathrm{sign}(\theta) \tan\left(\frac{\pi}{2}\alpha \right) \right) \right) \tag{A.12}$$

(θ real).

Set $\gamma = 1 - 1/\alpha$, and note that $1 < \alpha < 2$ implies that $0 < \gamma < 1/2$. The random variables $\{A_k\}_{k=1}^{N}$ are iid with power-law decay of their probability tails—governed by the exponent α ($1 < \alpha < 2$); $n \gg 1$, and $N - n \gg 1$. Hence, the Lévy CLT [10] implies that

$$\frac{A_1 + \cdots + A_n}{n} \approx \mu + \frac{1}{n^{\gamma}}\xi_1 \tag{A.13}$$

and

$$\frac{A_{n+1} + \cdots + A_N}{N - n} \approx \mu + \frac{1}{(N-n)^{\gamma}}\xi_2, \tag{A.14}$$

where ξ_1 and ξ_2 are independent two-sided Lévy random variables with exponent α and skew $\beta = 1$.

Substituting equations (A.13) and (A.14) into equation (A.2), while using equation (3), yields

$$\frac{A_{n+1} + \cdots + A_N}{A_1 + \cdots + A_n} \approx \frac{1-f}{f} \frac{\mu + \frac{1}{N^{\gamma}(1-f)^{\gamma}}\xi_2}{\mu + \frac{1}{N^{\gamma}f^{\gamma}}\xi_1}$$

$$= \frac{1-f}{f} \frac{1 + \left(\frac{1}{\mu(1-f)^{\gamma}}\xi_2 \right)\frac{1}{N^{\gamma}}}{1 + \left(\frac{1}{\mu f^{\gamma}}\xi_1 \right)\frac{1}{N^{\gamma}}}. \tag{A.15}$$

Moreover

$$\frac{1 + \left(\frac{1}{\mu(1-f)^{\gamma}}\xi_2 \right)\frac{1}{N^{\gamma}}}{1 + \left(\frac{1}{\mu f^{\gamma}}\xi_1 \right)\frac{1}{N^{\gamma}}} \approx \left(1 + \left(\frac{1}{\mu(1-f)^{\gamma}}\xi_2 \right) \frac{1}{N^{\gamma}} \right)\left(1 - \left(\frac{1}{\mu f^{\gamma}}\xi_1 \right) \frac{1}{N^{\gamma}} \right)$$

$$\approx 1 + \frac{1}{N^{\gamma}}\frac{1}{\mu}\left(\frac{1}{(1-f)^{\gamma}}\xi_2 - \frac{1}{f^{\gamma}}\xi_1 \right). \tag{A.16}$$

And, substituting equation (A.16) into equation (A.15) gives

$$\frac{A_{n+1} + \cdots + A_N}{A_1 + \cdots + A_n} \approx \frac{1-f}{f}\left(1 - \frac{1}{N^{\gamma}}\frac{1}{\mu}\left(\frac{1}{f^{\gamma}}\xi_1 - \frac{1}{(1-f)^{\gamma}}\xi_2 \right) \right). \tag{A.17}$$

In turn, substituting equation (A.17) into equation (A.1) yields

$$P \approx \frac{1}{1 + \frac{1-f}{f}\left(1 - \frac{1}{N^{\gamma}}\frac{1}{\mu}\left(\frac{1}{f^{\gamma}}\xi_1 - \frac{1}{(1-f)^{\gamma}}\xi_2 \right) \right)}$$

$$= \frac{f}{1 - \frac{1}{N^{\gamma}}\frac{1-f}{\mu}\left(\frac{1}{f^{\gamma}}\xi_1 - \frac{1}{(1-f)^{\gamma}}\xi_2 \right)}$$

$$\approx f + \frac{1}{N^{\gamma}}\frac{f(1-f)}{\mu}\left(\frac{1}{f^{\gamma}}\xi_1 - \frac{1}{(1-f)^{\gamma}}\xi_2 \right). \tag{A.18}$$

Moreover, equation (A.12) implies that

$$\frac{1}{f^{\gamma}}\xi_1 - \frac{1}{(1-f)^{\gamma}}\xi_2 \stackrel{\text{Law}}{=} \left(\frac{1}{f^{\alpha-1}} + \frac{1}{(1-f)^{\alpha-1}} \right)^{1/\alpha} \cdot L \tag{A.19}$$

(the sign $\overset{\text{Law}}{=}$ denoting equality in law), where L is a two-sided Lévy random variable with exponent α and skew

$$\beta = \frac{f^{\alpha-1} - (1-f)^{\alpha-1}}{f^{\alpha-1} + (1-f)^{\alpha-1}}. \tag{A.20}$$

Hence, we conclude that the random probability P is approximated by

$$P \approx f + \frac{1}{N^\gamma} \cdot \left(\frac{f(1-f)}{\mu} \right) \left(\frac{1}{f^{\alpha-1}} + \frac{1}{(1-f)^{\alpha-1}} \right)^{1/\alpha} \cdot L$$

$$= f + \frac{1}{N^\gamma} \cdot \frac{1}{\mu} (f(1-f)^\alpha + (1-f)f^\alpha)^{1/\alpha} \cdot L, \tag{A.21}$$

where L is a two-sided Lévy random variable with exponent α and skew β given by equation (A.20).

Finally, substituting equation (A.19) into equation (6)—while taking into account that $\gamma = 1 - 1/\alpha < 1/2$—we arrive at

$$\hat{\rho} \approx \frac{\rho}{f} \left(f + \frac{1}{N^\gamma} \frac{1}{\mu} (f(1-f)^\alpha + (1-f)f^\alpha)^{1/\alpha} L \right) + \frac{1}{\sqrt{N}} \left(\frac{\sqrt{\rho}}{f} \sqrt{f(1-f)} \right) Z_1$$

$$\approx \rho + \frac{1}{N^{1-1/\alpha}} \cdot \frac{\rho}{\mu f} (f(1-f)^\alpha + (1-f)f^\alpha)^{1/\alpha} \cdot L, \tag{A.22}$$

where L is a two-sided Lévy random variable with exponent α and skew β given by equation (A.20).

A.3. The anomalous Lévy scenario: derivation of equations (13) and (15)

Recall that a random variable ξ is a one-sided Lévy random variable with exponent α ($0 < \alpha < 1$) if it is characterized by the Laplace transform [10]:

$$\langle \exp(i\theta\xi) \rangle = \exp(-\theta^\alpha) \tag{A.23}$$

($\theta \geqslant 0$).

The random variables $\{A_k\}_{k=1}^N$ are iid with power-law decay of their probability tails—governed by the exponent α ($0 < \alpha < 1$); $n \gg 1$ and $N - n \gg 1$. Hence, the Lévy CLT [10] implies that

$$\frac{A_1 + \cdots + A_n}{n^{1/\alpha}} \approx \xi_1 \tag{A.24}$$

and

$$\frac{A_{n+1} + \cdots + A_N}{(N-n)^{1/\alpha}} \approx \xi_2, \tag{A.25}$$

where ξ_1 and ξ_2 are independent one-sided Lévy random variables with exponent α.

Equations (A.24) and (A.25) imply that

$$\frac{A_{n+1} + \cdots + A_N}{A_1 + \cdots + A_n} \approx \left(\frac{N-n}{n} \right)^{1/\alpha} \frac{\xi_2}{\xi_1}. \tag{A.26}$$

Setting $R = \xi_2/\xi_1$, and substituting equation (A.26) into equation (A.1)—while using equation (3)—yields

$$P \approx \frac{f^{1/\alpha}}{f^{1/\alpha} + (1-f)^{1/\alpha}R}. \tag{A.27}$$

Finally, substituting equation (A.27) into equation (6) we arrive at

$$
\hat{\rho} \approx \frac{\rho}{f} \left(\frac{f^{1/\alpha}}{f^{1/\alpha} + (1-f)^{1/\alpha} R} \right) + \frac{1}{\sqrt{N}} \left(\frac{\sqrt{\rho}}{f} \sqrt{P(1-P)} \right) Z_1
$$
$$
\approx \rho \cdot \left(\frac{f^{(1/\alpha)-1}}{f^{1/\alpha} + (1-f)^{1/\alpha} R} \right), \tag{A.28}
$$

where R is the quotient of two iid one-sided Lévy random variables with exponent α $(0 < \alpha < 1)$.

References

[1] Maddox J 1989 *Nature* **341** 181
[2] Gurman S J 1989 *Nature* **342** 12
[3] Bel G and Barkai E 2005 *Phys. Rev. Lett.* **94** 240602
[4] Lubelski A, Sokolov I M and Klafter J 2008 *Phys. Rev. Lett.* **100** 250602
[5] He Y, Burov S, Metzler R and Barkai E 2008 *Phys. Rev. Lett.* **101** 058101
[6] Bak P 1996 *How Nature Works* (New York: Springer)
[7] Jensen H J 1998 *Self-Organized Criticality* (Cambridge: Cambridge University Press)
[8] Eliazar I and Klafter J 2009 *Phys. Rev.* E **79** 011103
[9] Mitrovic M and Tadic B 2009 http://arxiv.org/abs/0910.2849
[10] Feller W 1971 *An Introduction to Probability Theory and Its Applications* vol 2, 2nd edition (New York: Wiley)
[11] Eliazar I 2005 *Physica* A **356** 207
[12] Erdélyi A, Magnus W, Oberhettinger F and Tricomi F G 1955 *Higher Transcendental Functions* (the Bateman project) vol 3 (New York: McGraw-Hill)