Supplementary Information

Extinction probability of an epidemic

To explore the extinction probability and extinction time of an epidemic, we calculate them based on three kinds of distribution: a) Poisson distribution, *i.e.*, $V(k) = E(k)$; b) power law distribution, *i.e.*, $V(k) > E(k)$; c) binomial distribution, *i.e.*, $V(k) < E(k)$. The calculation method is as follows:

The probability generating function of k (progeny number of an individual) is defined as

$$G_k(s) = \sum_{k=0}^{\infty} p_k s^k = E(s^K)$$

where p_k is the probability distributon of k. Note $G_k(1) = 1$, $G_k(0) = p_0$. Specifically, when k follows Poisson distribution [k~Poisson(λ)], its probability generating function is

$$G_k(s) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} s^k = e^{\lambda(s-1)}$$

When k follows binomial distribution [K~Bin(n, p)], its probability generating function is

$$G_k(s) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} s^k = (1-p+ps)^n$$

The probability that the process is extinct by the n-th generation is

$$u_n = P(N(n) = 0 | N(0) = 1) \quad (A1)$$

where $N(n)$ is the population size at n-th generation. $N(0)$ is just the initial population size. Now $u_n \leq 1$, and $u_n \leq u_{n+1}$ – *i.e.*, $\{u_n\}$ is a bounded monotonic sequence [1]. So

$$u = \lim_{n \to \infty} u_n$$

exists and is called the probability of ultimate extinction. The average extinction time with $N(0) = 1$ can be calculated as follows:

$$T = \sum_{t=1}^{\infty} t \times u_t \quad (A2)$$

Because of the independence,

$$P(N(n) = 0 | N(0) = k) = u_n^k \quad (A3)$$

Recall that p_k is the probability distributon of *k*, we can have the following theorem

$$u_n = \sum_{k=0}^{\infty} p_k (u_{n-1})^k \quad (A4)$$

Furthermore, when $n \to \infty$, we can know that u is the smallest non-negative root of following equation [1]

$$G_K(u) = \sum_{k=0}^{\infty} p_k u^k = u \quad (A5)$$

To explore the extinction probability of an epidemic, we calculate the extinction probability based on three kinds of distribution: a) Poisson distribution, *i.e*., *V*(*k*) = *E*(*k*); b) power law distribution, *i.e*., *V*(*k*) > *E*(*k*); c) binomial distribution, *i.e*., *V*(*k*) < *E*(*k*).

a) Poisson distribution

If *k* follows Poisson distribution [*k*~Poisson(λ)], its probability generating function is

$$G_k(s) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} s^k = e^{\lambda(s-1)}$$

we can obtain its ultimate extinction probability by obtaining the smallest non-negative root of following equation:

$$G_k(u) = e^{\lambda(u-1)} = u \to u = \frac{LambertW\left(-\lambda e^{-\lambda}\right)}{\lambda} \quad (A6)$$

where $LambertW$ is the *Lambert W* function [2]. Here we let *V*(*k*) = *E*(*k*) = 0.5, 1.0, 1.05, 1.1, 1.2, 1.3, 1.4, 1.5, 2, 3, 4, 5. The extinction probability can be obtained by Eq. (A6). The results are shown in Table S1.

b) Binomial distribution

For binomial distribution [*k*~Bin(*n*, *p*)], its probability generating function is

$$G_k(s) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} s^k = (1 - p + ps)^n$$

And its ultimate extinction probability is just the smallest non-negative root of following equation:

$$G_k(u) = (1 - p + pu)^n = u \quad (A7)$$

we let $p = q = 0.5$, $E(k) = 0.5, 1, 2, 3, 4, 5$. And then $V(k) = 0.5E(k)$. Use Eq. (A7), we can calculate the extinction probabilities of binomial distribution (Table S2)

c) Power law distribution

Just as Ruan *et al.* [3] suggest, we let the distribution of k follow the Zipf's law ([4]; a discrete format of power law distribution):

$$P(k = i; c, M) = \frac{1/(i+1)^c}{H_{M,c}}, i = 0,1, \dots, M-1$$

where $H_{M,c} = \sum_{m=1}^{M} 1/m^c$. The mean and variance of k are

$$E(k) = \frac{H_{M,c-1}}{H_{M,c}} - 1$$

$$V(k) = \frac{H_{M,c-2}}{H_{M,c}} - \frac{H_{M,c-1}^2}{H_{M,c}^2}$$

Based on the definition above, we can calculate the extinction probability according to Eq. (A4) and Eq. (A5). We let $V(k) = 5E(k)$ or $10E(k)$, and $E(k)$ ranges from 0.5 to 5. In addition, we do a simulation to test the accuracy of the analytical equations. The simulation results are consistent with the analytical results (Table S3 and S4).

Modeling the viral evolution in H_0 with different parameter sets

In main text, we simulate the viral evolution in H_0 at PL_0 with following parameter set: power law distribution with $V(k) = 5E(k)$, initial $E(k) = 1.1$ (Fig. 4). Here we expand the parameter set: (a) power law distribution with $V(k) = 10E(k)$, initial $E(k) = 1.1$, see Fig. S1; (b) poisson distribution with $V(k) = E(k)$, initial $E(k) = 1.1$, see Fig. S2. The patterns of Fig. 4, Fig. S1 and Fig. S2 are very similar. All of them show a biomodal distribution when $z = 3$, but not when $z = 1$ or 0.3. The reason for this bimodal distribution is that most of the invasions with $E(k) = 1.1$ would fail, which corresponds to the rapidly failed cases. The remaining would become extinct when the host population develops herd immunity, which takes more generations (denoted as G) in this simulation. Only in the latter cases would N_{inf} be large enough to yield adaptive mutations for the virus to overcome the immune suppression. The main difference amont them is G will turn to be larger as $V(k)$ increases. This is because as $V(k)$ increases, the time of extinction of fixation is shorten [5].

References

[1] Grimmett G, Stirzaker DR. Probability and random processes. Oxford: Oxford Univ. Press, 2009

[2] Corless RM, Gonnet GH, Hare DEG, et al. On the lambertw function. Advances in Computational Mathematics, 1996, 5: 329-359

[3] Ruan Y, Luo Z, Tang X, et al. On the founder effect in covid-19 outbreaks – how many infected travelers may have started them all? National Science Review, 2020, nwaa246

[4] Zipf GK. Human behavior and the principle of least effort. Oxford, England: Addison-Wesley Press, 1949

[5] Crow JF, Kimura M. An introduction to population genetics theory. New York, Evanston and London: Harper & Row, Publishers, 1970

Supplementary tables and figures

Table S1. Extinction probability of Poisson distribution with *V*(*k*) = *E*(*k*)

	***E*(*k*)=0.5**	**1**	**1.05**	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**2**	**3**	**4**	**5**
***N*(0)=1**	1	1	0.91	0.82	0.69	0.58	0.49	4.2E-01	2.0E-01	6.0E-02	2.0E-02	7.0E-03
2	1	1	0.82	0.68	0.47	0.33	0.24	1.7E-01	4.1E-02	3.5E-03	3.9E-04	4.9E-05
3	1	1	0.74	0.56	0.32	0.19	0.12	7.3E-02	8.4E-03	2.1E-04	7.8E-06	3.4E-07
4	1	1	0.67	0.46	0.22	0.11	0.06	3.0E-02	1.7E-03	1.3E-05	1.5E-07	2.4E-09
5	1	1	0.61	0.38	0.15	0.06	0.03	1.3E-02	3.5E-04	7.5E-07	3.1E-09	1.7E-11

Table S2. Extinction probability of binomial distribution with *V*(*k*) = 0.5*E*(*k*)

	***E*(*k*)=0.5**	**1**	**1.5**	**2**	**3**	**4**	**5**
***N*(0)=1**	1.00	1.00	2.4E-01	8.7E-02	1.7E-02	4.0E-03	9.9E-04
2	1.00	1.00	5.6E-02	7.6E-03	3.0E-04	1.6E-05	9.8E-07
3	1.00	1.00	1.3E-02	6.7E-04	5.2E-06	6.5E-08	9.7E-10
4	1.00	1.00	3.1E-03	5.8E-05	9.0E-08	2.6E-10	9.6E-13
5	1.00	1.00	7.3E-04	5.1E-06	1.6E-09	1.1E-12	9.5E-16

Table S3. Extinction probability of power law distribution with *V*(*k*) = 10*E*(*k*)

	***E*(*k*)=0.5**	**1**	**1.05**	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**2**	**3**	**4**	**5**
***N*(0)=1(simulation)**	1.00	1.00	1.00	0.98	0.96	0.93	0.92	0.89	0.80	0.63	0.52	0.39
***N*(0)=1(analytical)**	1.00	1.00	0.99	0.98	0.96	0.93	0.91	0.90	0.78	0.63	0.51	0.41
2	1.00	1.00	0.98	0.96	0.92	0.87	0.84	0.80	0.61	0.39	0.26	0.17
3	1.00	1.00	0.97	0.94	0.88	0.82	0.77	0.72	0.47	0.25	0.13	0.07
4	1.00	1.00	0.96	0.92	0.84	0.76	0.70	0.65	0.37	0.15	0.07	0.03
5	1.00	1.00	0.95	0.91	0.81	0.71	0.64	0.58	0.29	0.10	0.03	0.01

Table S4. Extinction probability of power law distribution with *V*(*k*) = 5*E*(*k*)

	***E*(*k*)=0.5**	**1**	**1.05**	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**2**	**3**	**4**	**5**
***N*(0)=1(simulation)**	1.00	1.00	0.98	0.97	0.92	0.88	0.86	0.82	0.69	0.47	0.36	0.37
***N*(0)=1(analytical)**	1.00	1.00	0.98	0.96	0.92	0.89	0.85	0.82	0.67	0.48	0.36	0.35
2	1.00	1.00	0.96	0.92	0.85	0.79	0.72	0.67	0.45	0.23	0.13	0.12
3	1.00	1.00	0.94	0.88	0.78	0.70	0.61	0.55	0.30	0.11	0.05	0.04
4	1.00	1.00	0.92	0.85	0.72	0.63	0.52	0.45	0.20	0.05	0.02	0.02
5	1.00	1.00	0.90	0.82	0.66	0.56	0.44	0.37	0.14	0.03	0.01	0.01

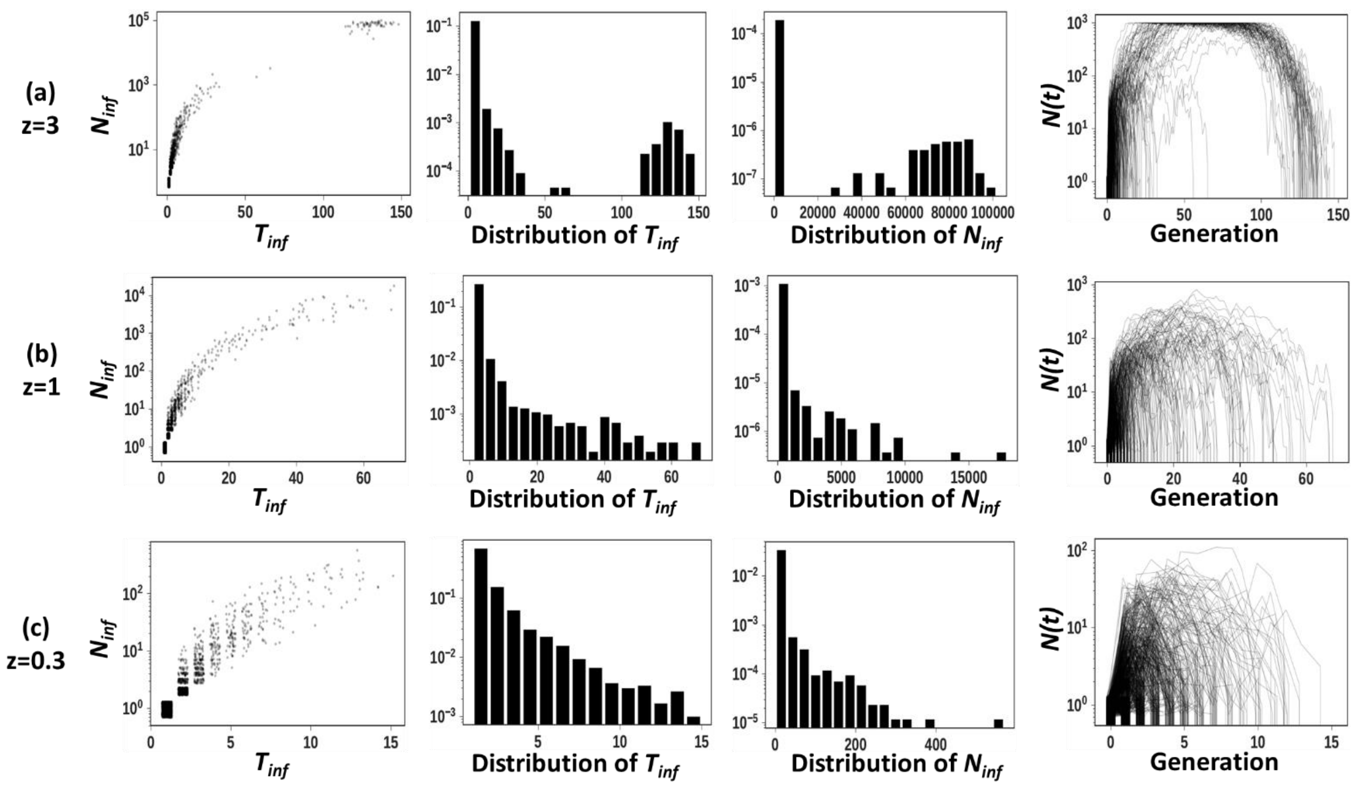

Fig. S1. The distribution of T_{inf} and N_{inf} after viral invasion. A virus (*i.e.*, $N(0) = 1$) with initial $E(k) = 1.1$ invades human population (population size is 1000). The distribution of k follows a power law distribution with $V(k) = 10E(k)$. T_{inf} is the duration between the initial infection and ultimate extinction. N_{inf} is the total number of infected individuals summed over all generations. Due to the prior build-up of herd immunity, $E(k)$ will decrease at 3 different rates (z = 3, 1, or 0.3) as shown in Fig. 4a – 4c, respectively. For each parameter set, 3000 repeats are simulated. Each repeat is portrayed by a thin line in the last panel

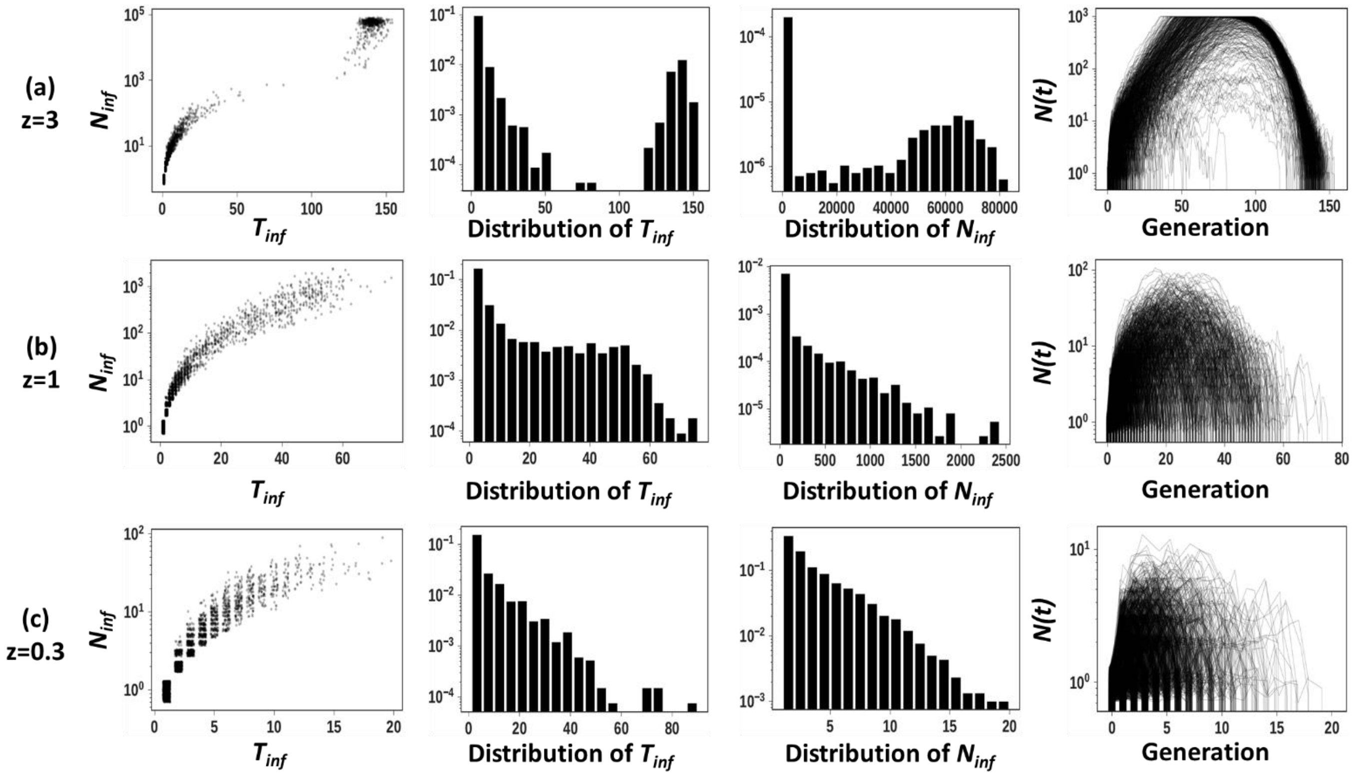

Fig. S2. same as Fig. S1 except that the distribution of k follows a Poisson distribution with $V(k) = E(k)$.