1 Problem Statement and Parameters

We consider a continuous-time birth–death–mutation–immigration process for a cell population with possibly multiple "states."

- $\lambda > 0$: Per-capita **division (birth)** rate.
- $\delta \geq 0$: Per-capita **death** rate.
- $p \in [0, 1]$: Probability that each new-born cell is a *new state* (mutation) rather than inheriting the parent state.
- $\gamma \geq 0$: **Immigration** rate. Two major models:
  - **Source-state model:** All immigrants enter one designated "source" state.
  - **Direct-state model:** Each immigrant arrives in a *distinct* new state.
- $d \in \{1, 2, \dots\} \cup \{\infty\}$: total number of possible states. If $d < \infty$, label depletion arises after all states are used.

1.1 Key Rates

$r \;=\; \lambda - \delta$ (net growth of total population), $\quad s \;=\; (1-p)\,\lambda - \delta$ (net growth of one fixed state).

2 Source-State Model: Step-by-Step Derivations

2.1 Mean-Field ODE for Total Population

Let $N(t)$ be the total population at time t. A simple first-moment approximation yields

$$\frac{d}{dt}\,\langle N(t)\rangle \;=\; (\lambda - \delta)\,\langle N(t)\rangle \;+\; \gamma \;=\; r\,\langle N(t)\rangle \;+\; \gamma.$$

Hence

$$\langle N(t)\rangle = \begin{cases} N(0)\,\mathrm{e}^{rt} \;+\; \dfrac{\gamma}{r}\,\bigl(\mathrm{e}^{rt} - 1\bigr), & r \neq 0, \\ N(0) \;+\; \gamma\,t, & r = 0. \end{cases}$$

2.2 Rate of New-State Formation

Each cell divides at rate λ, so births occur at $\lambda\,N(t)$. Since a fraction p are new states, the *expected* new-state formation rate is

$$p\,\lambda\,\langle N(t)\rangle.$$

In the infinite-states setting, there is no shortage of labels. In finite-states, label depletion is handled via a depletion ODE (later).

2.3 Individual State Growth Rate s

If a state does not mutate further, it has birth rate $(1-p)\,\lambda$ and death rate δ. Thus

$$s \;=\; (1-p)\,\lambda \;-\; \delta.$$

This is the state-level net rate.

2.4 Case 1: $s > 0,\ \gamma = 0$ (Supercritical, No Immigration)

Derivation of $n(k) \propto k^{-\,s/r}$.

(i) $\langle N(t)\rangle \approx e^{r\,t}$ since $\gamma = 0$ and $r = \lambda - \delta > 0$.

(ii) New states appear at rate $p\,\lambda\,e^{r\,t}$.

(iii) A new state formed at time $\tau \in [0,t]$ grows approximately like $e^{s\,(t-\tau)}$ if we ignore extinction for the moment.

(iv) It will exceed some threshold n at final time t precisely if $\tau \le t - \frac{\ln n}{s}$.

(v) Hence the expected number of states $\ge n$ is:

$$S(n) \;=\; \int_0^{t-\frac{\ln n}{s}} p\,\lambda\,e^{r\,\tau}\,d\tau \;=\; \frac{p\,\lambda}{r}\left[e^{r\left(t-\frac{\ln n}{s}\right)} \;-\; 1\right] \;\propto\; e^{r\,t} \times n^{-\,\frac{r}{s}}.$$

So $S(n) \approx \text{const} \times n^{-r/s}$. If the rank $k \approx S(n)$, then $n \approx k^{-\,s/r}$. Therefore

$$\boxed{n(k) \;\propto\; k^{-\,\frac{s}{r}} \quad \text{(no immigration, supercritical).}}$$

2.5 Case 2: $s \le 0,\ \gamma = 0$ (Subcritical/Critical, No Immigration)

No persistent large states form in expectation; rank–abundance quickly becomes trivial. No further derivation needed for a nonzero tail.

2.6 Case 3: $s > 0,\ \gamma > 0$ (Supercritical, With Immigration)

- $\langle N(t)\rangle =$ solution of $d\langle N\rangle/dt = r\langle N\rangle + \gamma$.
- For large t, $\langle N(t)\rangle \sim e^{r\,t}$ if $r > 0$.
- The new-state formation rate is $p\,\lambda\,\langle N(t)\rangle$. Repeating the integral argument (as in Case 1) gives $S(n) \sim n^{-\,r/s}$. Thus the rank–abundance exponent $-\,s/r$ is unchanged by γ.

$$\boxed{n(k) \;\propto\; k^{-\,\frac{s}{r}} \quad (\text{supercritical}, \gamma > 0, \text{same exponent } -\,\tfrac{s}{r}).}$$

2.7 Case 4: $s < 0,\ \gamma > 0,\ p > 0$ (Subcritical Source + Mutation)

Step-by-step derivation for the $k^{-\frac{|s|^2}{p\lambda\gamma}}$ tail.

(i) The **source state** (label 0) is fed by immigration γ. - Mean-field: a purely subcritical branching at rate $(1-p)\lambda - \delta = s < 0$ would die out if isolated. - But immigration γ adds a constant inflow. For large time, the source stabilizes near $N_0 = \gamma/|s|$.

(ii) Each **division** in the source state produces a new mutant with probability p. So new mutants arise at rate $R_{\text{new}} = p\,\lambda\,N_0 = \dfrac{p\,\lambda\,\gamma}{|s|}$.

(iii) **Subcritical mutant lineages.** Once a mutant arises, it evolves with net rate $s < 0$. Deterministically it would shrink as $\mathrm{e}^{s\,a}$. Stochastically, it can briefly fluctuate above $\mathrm{e}^{s\,a}$.

(iv) **Age distribution approach.** Mutants appear at a constant Poisson rate R_{new}. The *age* a of a mutant state is the time since it appeared. At equilibrium, the distribution of mutant ages is $f(a) = R_{\text{new}}\,e^{-R_{\text{new}}\,a}$ for $a \geq 0$.

(v) A lineage with age a has an *expected* size $\mathrm{e}^{s\,a} \approx \mathrm{e}^{-|s|\,a}$. The probability it is $\geq n$ is roughly determined by the condition $\mathrm{e}^{-|s|\,a} \geq n \implies a \leq \frac{\ln(1/n)}{|s|}$. Integrating the fraction of lineages that are young enough to exceed n:

$$S(n) \;=\; R_{\text{new}} \int_0^{\frac{\ln(1/n)}{|s|}} e^{-\,R_{\text{new}}\,a}\,\mathrm{d}a \;=\; R_{\text{new}}\,\frac{1}{R_{\text{new}}}\left[1 - \mathrm{e}^{-\,R_{\text{new}}\,\frac{\ln(1/n)}{|s|}}\right] \;\approx\; \frac{R_{\text{new}}}{|s|}\,n^{-\frac{R_{\text{new}}}{|s|}},$$

since $\mathrm{e}^{-\,R_{\text{new}}\,\frac{\ln(1/n)}{|s|}} = n^{-\frac{R_{\text{new}}}{|s|}}$. Recalling $R_{\text{new}} = p\,\lambda\,\gamma/|s|$, we get

$$S(n) \;\propto\; n^{-\frac{p\,\lambda\,\gamma/|s|}{|s|}} \;=\; n^{-\frac{p\,\lambda\,\gamma}{|s|^2}}.$$

Hence, if rank $k \approx S(n)$, then

$$n(k) \;\propto\; k^{-\frac{|s|^2}{p\,\lambda\,\gamma}},$$

and the source state ($k = 1$) remains at $\gamma/|s|$. Thus:

$$\boxed{n(1) = \frac{\gamma}{|s|}, \quad n(k \geq 2) \;\propto\; k^{-\frac{|s|^2}{p\,\lambda\,\gamma}}, \quad \text{(subcritical source + mutation)}.}$$

2.8 Case 4a: $s < 0,\ \gamma > 0,\ p = 0$

No mutants arise ($p = 0$), so there is *only* the source, at mean size $\gamma/|s|$. One-state population.

3 Direct-State Model: Step-by-Step Derivations

Now immigration itself *directly* seeds distinct states. Typically set $p = 0$ (no further splits).

3.1 Subcritical ($s < 0$): Logarithmic Decay in Rank

Derivation of $n(k) \propto \ln\bigl[C/(k-1)\bigr]/(2\,|s|)$.

(i) γ is the immigration rate, each arrival starts a new state with net rate $s < 0$.

(ii) For one subcritical birth–death process with net rate $s < 0$, the probability of reaching size $\geq n$ before going extinct is asymptotically $\mathrm{e}^{-2\,|s|\,n}$ (in the usual gambler's-ruin or branching approximation).

(iii) Over a time window $[0, T]$, the expected number of states that ever reach size $\geq n$ is roughly

$$\gamma\, T\, \mathrm{e}^{-2|s|\,n}.$$

Denote that by $S(n)$.

(iv) Invert $k \approx S(n)$ to solve for $n(k)$. So if $k \approx \gamma\, T\, \mathrm{e}^{-2|s|\,n}$, then $\mathrm{e}^{-2|s|\,n} \approx k/(\gamma T) \implies -2|s|\,n \approx \ln\bigl[k/(\gamma T)\bigr]$. Hence

$$n \;\approx\; \frac{1}{2|s|}\,\ln\Bigl(\frac{\gamma T}{k}\Bigr).$$

Write it as

$$n(k) \;\propto\; \frac{1}{2|s|}\,\ln\Bigl(\frac{C}{k}\Bigr),$$

where $C \propto \gamma\, T$. Often one sees it in a form with $\ln\bigl[C/(k-1)\bigr]$. So

$$\boxed{n(k) \;\propto\; \frac{1}{2|s|}\,\ln\Bigl(\frac{C}{k}\Bigr), \quad \text{(direct-state, subcritical).}}$$

(v) Because these states drift to extinction, $\frac{n(1)}{n(2)}$ often stays near 1 – no single state dominates.

3.2 Supercritical ($s > 0$): Exponential Decay in Rank

Derivation of $n(k) \propto \mathrm{e}^{-(s/\gamma)\,k}$.

(i) Immigration at rate γ seeds new states. The k-th state arrives at time $\tau_k \approx k/\gamma$.

(ii) A supercritical state (net rate $s > 0$) grows like $\mathrm{e}^{s\,(t-\tau_k)}$.

(iii) At some observation time $t \gg \tau_k$, the size of the k-th state is

$$n(k) \;=\; \exp\Bigl[s\,(t - \tfrac{k}{\gamma})\Bigr] \;=\; e^{s\,t} \times e^{-(s/\gamma)\,k}.$$

Ignoring the overall factor $e^{s\,t}$, we see an exponential decrease in rank k. Hence

$$\boxed{n(k) \;\propto\; e^{-\frac{s}{\gamma}\,k}, \quad \text{(direct-state, supercritical).}}$$

(iv) Consequently, $\frac{n(1)}{n(2)} = e^{s/\gamma}$ (big ratio).

4 Finite d (Label Depletion)

If only d states exist, once all are used, no further new states can appear. A mean-field depletion ODE:

$$\frac{d}{dt}D(t) \;=\; -\,p\,\lambda\,\frac{D(t)}{d}\,\langle N(t)\rangle,$$

where $D(t)$ is the count of unused labels. Eventually $D(t) \to 0$. This *truncates* any infinite-state tail at rank $k = d$.

5 Results: Rank–Abundance and Transition Ratios (Fully Derived)

5.1 Rank–Abundance Table (All Cases)

Table 1: Rank–abundance relationships for $d = \infty$ (before label depletion).

Regime	**Structure**	**Rank–Abundance** $n(k)$
$s > 0,\ \gamma = 0$	Power law	$n(k)\ \propto\ k^{-\frac{s}{r}}$
$s > 0,\ \gamma > 0$	Power law (long-term) Source + power law (transient)	$n(k)\ \propto\ k^{-\frac{s}{r}}$
$s < 0,\ \gamma > 0,\ p > 0$	Source + single power law	$n(1) = \frac{\gamma}{\lvert s\rvert},\quad n(k \geq 2)\ \propto\ k^{-\frac{\lvert s\rvert^2}{p\,\lambda\,\gamma}}$
$s < 0,\ \gamma > 0,\ p = 0$	Single state	$n(1) = \frac{\gamma}{\lvert s\rvert},\quad n(k > 1) = 0$
Direct-state ($s < 0$)	Logarithmic decay	$n(k)\ \propto\ \frac{1}{2\lvert s\rvert}\ \ln\!\left(\frac{C}{k}\right)$
Direct-state ($s > 0$)	Exponential decay	$n(k)\ \propto\ \exp\!\left[-\frac{s}{\gamma}\,k\right]$

All these are derived in §2–§5 above.

5.2 Transition-Ratio Table (All Cases)

Every ratio above appears immediately by plugging in the explicit forms for $n(k)$ from the rank–abundance derivations.

Table 2: Top-rank ratio: $\frac{n(1)}{n(2)}$ in each regime. Every expression is derived below.

Regime	$\dfrac{n(1)}{n(2)}$	**Derivation**
Source-state, supercritical $(s > 0)$	$2^{\,s/r}$	$n(k) \propto k^{-\,s/r} \implies \frac{n(1)}{n(2)} = \frac{1^{-\,s/r}}{2^{-\,s/r}} = 2^{\,s/r}$.
Source-state, subcritical $(s < 0,\ \gamma > 0,\ p > 0)$	$\frac{\frac{\gamma}{\lvert s\rvert}}{C \cdot 2^{-\alpha}} = \frac{\gamma}{\lvert s\rvert\, C}\, 2^{\frac{\lvert s\rvert^2}{p\,\lambda\,\gamma}}$	$n(1) = \gamma/\lvert s\rvert,\ n(2) = C \cdot 2^{-\alpha}$. Hence $\frac{n(1)}{n(2)} = \frac{\gamma/\lvert s\rvert}{C\,2^{-\alpha}} = \frac{\gamma}{\lvert s\rvert\, C}\, 2^{\alpha}$.
Direct-state, subcritical $(s < 0)$	≈ 1	All states remain small (logarithmic rank curve). No single large state stands out, so $\frac{n(1)}{n(2)} \approx 1$.
Direct-state, supercritical $(s > 0)$	$\mathrm{e}^{\,s/\gamma}$	$n(k) \propto \exp[-(s/\gamma)\,k]$. Then $\frac{n(1)}{n(2)} = \dfrac{e^{-s/\gamma}}{e^{-2\,s/\gamma}} = e^{\,s/\gamma}$.