arXiv:2601.06433v1 [physics.comp-ph] 10 Jan 2026

根据信息输入统计研究内部时间的减慢与饱和：响应系统的最小模型

Tatsuaki Tsuruyama[1,2,*]

[1] *Department of Physics, Tohoku University, Sendai 980-8578, Japan*
[2] *Department of Drug Discovery Medicine, Kyoto University, Kyoto 606-8501, Japan*（日期：2026年1月13日）

我们考察一种响应系统，它会根据从外部到达的信息输入更新其内部状态。本文将截至某一给定时刻至少出现过一次的编码“种类数量”定义为内部时间，并分析内部时间的推进方式如何由信息输入统计（到达率和编码分布）决定。当到达遵循泊松过程时，内部时间的平均推进速度随时间单调减小；若编码种类数量有限，则最终会趋近上限并达到饱和。因此，在较长的时间尺度上，内部时间相对于物理时间会变得更短。对于均匀编码分布，我们给出内部时间与物理时间之间对应关系的闭式，并表明在后期，“内部时间推进一步”所需的物理时间会增加。作为辅助量，我们用条件熵量化仅知内部时间时“哪些编码已经出现过”仍存在的不确定性，并给出均匀情形下的单峰性和最大化时间，以及非均匀情形下的上界、等号成立条件和与上界之差的表达式。最后，我们还提出一种推广方法，为每个编码赋予权重（描述长度），使内部时间按照输入中的信息量递增。

I. 引言

即使经过相同的物理时间，经验的更新或学习的进展也未必恒定。在从外部到来的信息输入中，尚未观察到的编码越来越少时，更新的机会本身也会减少；更新速度随之减慢，并最终趋于饱和。本文将这一现象表述为一个根据输入信息统计来更新内部状态的“响应系统”最小模型，并严谨地推导出内部时间相对于物理时间会减慢；若编码的种类数量有限，内部时间便会趋于饱和。

在心理学和认知科学中，关于时间评估的研究不胜枚举 [1, 7], ，但本文并不旨在直接解释这些测量指标。相反，在一个简单的设定中，信息输入由有限种编码到达，而响应系统只保留最简记录；我们表明，根据输入统计（到达率、编码分布和权重），内部时间的推进速度必然会随时间减慢。因此，本文的核心结论是：“响应系统所计量的内部时间会根据输入统计而减慢，并在编码种类有限时趋于饱和。”也就是说，内部时间并非对物理时间的简单映射，而是由输入统计决定的响应量；在较长的时间尺度上，内部时间会相对短于物理时间。

* tsuruyam@kuhp.kyoto-u.ac.jp

a. 响应系统的最简性 本文所说的响应系统，是指每当信息输入到达时就更新其内部状态，并为后续响应保留最简记录的系统。本文将所保留的记录限定为“迄今观察到的编码种类”，并将这些种类的数量称为内部时间。正因为这种最简性，内部时间的减慢与饱和便以一种可根据输入统计直接计算的形式显现出来。

b. 结构安排 在 §II 中，我们定义模型；在 §III 中，我们将内部时间的减速与饱和作为主要结果呈现。在 §IV 中，我们介绍内部时间的加权推广。在 §V 中，我们梳理物理时间与内部时间之间的对应关系，并阐明内部时间在后续阶段会变得更难推进。在 §V D 0 a 中，我们引入条件熵作为辅助量，并给出均匀情形下的闭式、单峰性和最大化时间，以及非均匀情形下的上界、等号成立条件、KL表达式和近似计算。最后，在 §VII 中，我们讨论其意义、局限性及拓展方向。

II. 模型

在本文中，我们将外部输入的种类称为**编码**。我们将一个变量称为**已观测代码列表**，它概括截至某一给定时间至少被观测到一次的编码，且不重复，并记为 $S(t)$。我们将已观测代码列表中所含编码的数量称为**内部时间**，并记为 $\tau(t)$。到达时间由**到达过程**描述，每次到达时都会观测到一个编码。本文所指的输入

统计量包括：(i) 到达率λ （单位时间内到达次数的均值），以及 (ii) 编码分布 $p=(p_1,\ldots,p_K)$（每种编码被观测到的相对频率）。作为进一步的推广，我们引入 (iii) 依赖于代码的权重ℓ_i （描述长度）。

A. 信息输入：到达过程和编码

信息输入由到达时间序列$\{t_n\}_{n\geq 1}$ 和代码序列$\{X_n\}_{n\geq 1}$组成。编码的可能取值是有限的：

$$\mathcal{X}=\{1,2,\ldots,K\}.$$

设截至时间 t 的到达次数为

$$N(t):=\#\{n:t_n\leq t\}.$$

作为标准设定，我们假设 $N(t)$是强度为 $\lambda>0$的泊松过程，并且 X_n 是服从分布 $p=(p_1,\ldots,p_K)$的独立同分布随机变量。关于点过程的一般理论，参见 [4, 9]。

B. 响应系统的内部状态：已观测代码列表和内部时间

定义变量如下，以概括截至时间 t 至少观测到一次的代码，且不重复：

$$S(t):=\{X_1,\ldots,X_{N(t)}\}\subseteq\mathcal{X}$$

（此处我们使用数学记号 $\{\cdot\}$）。这等价于指示每个代码 i 是否已被观测到的比特序列：

$$I(t)=(I_1(t),\ldots,I_K(t)) \tag{1}$$

$$I_i(t):=\mathbf{1}\{i\in S(t)\}\in\{0,1\}. \tag{2}$$

将内部时间定义为

$$\tau(t):=|S(t)|=\sum_{i=1}^{K}I_i(t).$$

$\tau(t)$ 是一个取整数值的随机过程，表示“观测到的代码种类数量”。

a. 内部时间的含义 内部时间 $\tau(t)$并不是像到达计数 $N(t)$ 那样，用来计数 “ 输入到达了多少次 ” 的量，而是用来计数“观察到了多少种编码”的量。即使到达次数相同，如果反复观察到的是同一种编码， $\tau(t)$也几乎不会增加。这一区别正是后文所示减速与饱和的直接原因。

III. 主要结果：内部时间的减速与饱和

本节将说明，在泊松到达的情况下，内部时间 $\tau(t)$ 会减速，并且当 K有限时会达到饱和。这里的结论是关于响应系统中内部时间均值行为的严格结果。后文关于辅助量的部分，则独立于这一主要结果进行介绍。

A. 一般分布（泊松到达）

假设到达服从泊松过程（速率为 λ），且代码序列独立同分布（分布为 p）。属于代码 i 的到达计数 $N_i(t)$彼此独立，并满足

$$N_i(t)\sim\mathrm{Poisson}(\lambda p_i t)$$

（泊松分裂； [4, 9]）。因此

$$\mathbb{P}(i\in S(t))=\mathbb{P}(N_i(t)\geq 1)=:q_i(t)=1-e^{-\lambda p_i t}.$$

因此

$$\mathbb{E}[\tau(t)]=\sum_{i=1}^{K}q_i(t)=\sum_{i=1}^{K}\left(1-e^{-\lambda p_i t}\right), \tag{3}$$

并求导，

$$\frac{d}{dt}\mathbb{E}[\tau(t)]=\lambda\sum_{i=1}^{K}p_i e^{-\lambda p_i t}. \tag{4}$$

定理 1 （减速与饱和（一般情形 p））。在泊松到达下，平均推进速度$d\mathbb{E}[\tau$（t）$]/dt$ 随 t 单调递减，并且

$$\lim_{t\to\infty}\frac{d}{dt}\mathbb{E}[\tau(t)]=0$$

成立。此外 $\mathbb{E}[\tau$（t）$]\leq K$ 并且

$$\lim_{t\to\infty}\mathbb{E}[\tau(t)]=K$$

成立。因此，

$$\lim_{t\to\infty}\frac{\mathbb{E}[\tau(t)]}{t}=0.$$

证明。 (4) 中的每一项 $p_i e^{-\lambda p_i t}$ 都随着 t 的增加而单调递减，因此其总和也单调递减。由于每一项的极限均为 0，总和的极限也为 0。此外，由 $0\leq q_i(t)\leq 1$，(3) 可推出 $\mathbb{E}[\tau(t)]\leq K$。此外，由于 $t\to\infty$， $q_i(t)\to 1$，因此 $\mathbb{E}[\tau(t)]\to K$。最后，由于 $\mathbb{E}[\tau(t)]$有界，故 $\mathbb{E}[\tau(t)]/t\to 0$。

□

(4) 的每一项都对应“未观察到的代码比例”。随着时间推移，已观测代码列表逐渐增长，因此未观察到的代码减少，并且

内部时间的推进速度随之减慢。由于到达过程服从泊松过程，这种减速可以明确写成指数形式。由于内部时间 $\tau(t)$ 是“观测到的代码类型数量”，因此不会超过 K。当时间足够长时，每种代码至少都会被观测到一次，因此均值趋近于 K 并达到饱和。因此，在较长的时间尺度上，无论物理时间推进多少，内部时间几乎都不会推进，内部时间也就相对物理时间变得更短。

B. 均匀分布 $p_i = 1/K$：闭式

在均匀情形 $p_i = 1/K$ 下，所有 $q_i(t)$ 都相等：

$$q(t) = 1 - e^{-\lambda t/K}.$$

因此，

$$\mathbb{E}[\tau(t)] = K\big(1 - e^{-\lambda t/K}\big), \tag{5}$$

$$\frac{d}{dt}\mathbb{E}[\tau(t)] = \lambda e^{-\lambda t/K}, \tag{6}$$

平均推进速度呈指数下降。

a. 时间尺度 方程（5）表明，自然时间尺度为 K/λ。K 越大，达到饱和所需的物理时间越长；λ 越大，饱和出现得越早。在后文讨论物理时间与内部时间之间的对应关系时，这个 K/λ 将作为参照。

IV. 加权内部时间：与编码一致的推广

内部时间 $\tau(t)$ 是对“观测到的代码种类数量”的最简定义。另一方面，即使是同样“观测到一个新代码”，观测到稀有代码所携带的信息量也更大。为纳入这一点，我们为代码赋予权重（描述长度），并将观测代码列表中包含的总权重作为另一种内部时间。为每个代码 i赋予正权重（描述长度） $\ell_i > 0$ ，并定义

$$T(t) := \sum_{i \in S(t)} \ell_i \tag{7}$$

作为**加权内部时间**。 $T(t)$ 对应于“已观测代码列表中所包含代码的总描述长度”；而 $\tau(t)$ 只统计“种类数量”， $T(t)$ 的推进方式则取决于“观测到了哪些代码”。

A. 泊松到达下的均值与减速

在泊松到达的情况下，由于 $q_i(t) = \mathbb{P}(i \in S(t)) = 1 - e^{-\lambda p_i t}$，

$$\mathbb{E}[T(t)] = \sum_{i=1}^{K} \ell_i\, q_i(t) = \sum_{i=1}^{K} \ell_i \left(1 - e^{-\lambda p_i t}\right), \tag{8}$$

而平均推进速度为

$$\frac{d}{dt}\mathbb{E}[T(t)] = \lambda \sum_{i=1}^{K} \ell_i p_i e^{-\lambda p_i t}. \tag{9}$$

因此，随着 t 增加，每一项都单调递减，总和也随之递减。因此，加权内部时间的均值也会减速。对于有限方法 K, $\mathbb{E}[T(t)]$„ 例如 $t \to \infty$,

$$\lim_{t\to\infty} \mathbb{E}[T(t)] = \sum_{i=1}^{K} \ell_i.$$

也就是说，对于有限数量的代码， $T(t)$也会趋于一个上限而饱和。从这个意义上说， $\tau(t)$的饱和是由于“种类数量的上限”所致，而 $T(t)$的饱和则是由于“总描述长度的上限”所致。

B. 编码一致性条件（熵编码要求）

作为一种与编码分布 p_i一致地选择权重 ℓ_i 的方法，假设

$$-\log p_i = \beta \ell_i + \text{const}, \qquad \beta > 0, \tag{10}$$

这是一项“越罕见的编码，其描述长度越大”的要求，并且与信息论中的编码一致（自信息 $-\log p_i$ 与描述长度之间的对应关系） [3, 12]。

忽略常数项（10）意味着 ℓ_i 与 $-\log p_i$成正比。在这种情况下， $T(t)$可以解释为“已观测代码列表所保留的自信息总量”的自然代表量。

a. 与均匀分布的一致性 对于均匀分布 $p_i = 1/K$，$-\log p_i = \log K$与 i无关，为常数。因此，（10）要求所有 ℓ_i 都相等。于是， $T(t)$成为 $\tau(t)$的常数倍，而加权内部时间是“以种类数量表示的内部时间”的一种推广。

b. 对非均匀分布的含义 对于非均匀分布，更罕见的编码具有更大的 ℓ_i，因此，如果在较晚阶段观测到一个稀有代码， $T(t)$ 就会大幅增加。稀有代码更难被观测到，因此它们的贡献会较晚显现。因此， $T(t)$ 往往呈现出这样的形式：“在早期阶段推进相对较快，但后期阶段的增长会持续很长时间。”这一特性在后文关于齐夫定律输入的章节中会更加直观地体现出来。

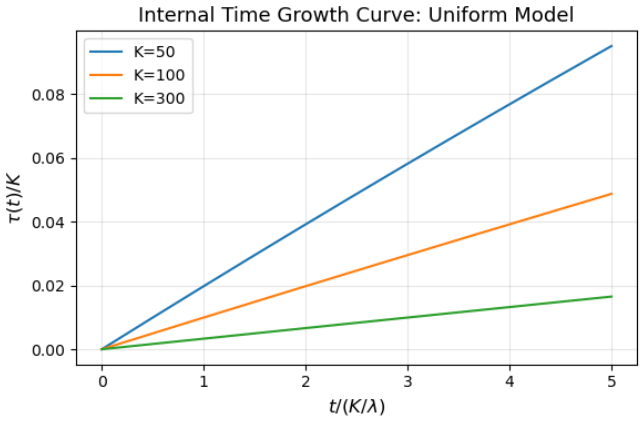


图 1. 均匀情形下，平均内部时间 $\mathbb{E}[\tau(t)]$ 与物理时间 t 之间的对应关系。它在早期阶段迅速增加，随后逐渐趋于饱和，因此内部时间的推进显著减缓。

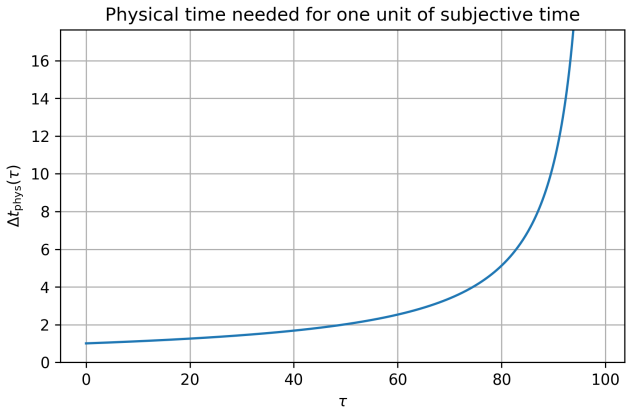


图 2. 均匀情形下获得“一个单位内部时间”所需的物理时间 $\Delta t_{\text{phys}}(\tau)$（式（12））。在后期阶段，该时间急剧增加，表明内部时间的推进显著减慢。

V. 物理时间与内部时间之间的对应关系

我们首先给出均匀分布情况下的闭式，然后描述非均匀分布所造成的失真。

A. 均匀参考系统： t 与 τ之间的非线性对应关系

在均匀情形下，内部时间的均值由式（5）给出。在早期阶段， $\mathbb{E}[\tau(t)] \approx \lambda t$ 几乎线性增加，但随着时间推移，增幅逐渐减缓， $\mathbb{E}[\tau(t)] \to K$ 最终趋于饱和。

a. “相对更短”是什么意思？由于$\mathbb{E}[\tau(t)]$ 有界，当 $t \to \infty$ 时， $\mathbb{E}[\tau(t)]/t \to 0$ 成立。这意味着：“在较长的时间尺度上，即使物理时间继续前进，内部时间也几乎不再推进。”从这个意义上说，内部时间相对于物理时间变得更短。

B. 逆映射：达到某一内部时间水平所需的物理时间

对式（5）求逆，达到内部时间水平 τ 所需的物理时间为

$$t(\tau) = \frac{K}{\lambda} \log\left(\frac{K}{K-\tau}\right) \tag{11}$$

（当 $\tau \to K$ 时发散）。对应于一个单位内部时间的物理时间增量为

$$\Delta t_{\text{phys}}(\tau) = t(\tau+1) - t(\tau) = \frac{K}{\lambda} \log\left(\frac{K-\tau}{K-\tau-1}\right), \tag{12}$$

其随 $\tau \to K$ 急剧增加。

a. 重新表述 方程（12）具体表明，在后期阶段，“使内部时间推进相同幅度所需的物理时间”会增加。这种“所需物理时间的增加”，是对内部时间相对物理时间变短这一主张最直接的形式化表述。

C. 非均匀分布： $\mathbb{E}[\tau（t）]$的失真

对于一般分布 p，式（3）成立。因此，初始斜率为

$$\frac{d}{dt}\mathbb{E}[\tau(t)]\Big|_{t=0} = \lambda \sum_{i=1}^{K} p_i = \lambda,$$

因此，在早期阶段，它的增长近似为 λt ，与分布无关。另一方面，后期形状则强烈依赖于 p。

由于高概率编码很早就会被观测到，而低概率编码直到较晚时才会被观测到，因此，与均匀情形相比，可能会出现一种失真：“早期上升很快，但后期长尾绵延。”这种失真源于：每种编码的观测概率 $q_i(t) = 1 - e^{-\lambda p_i t}$ 会根据 $\lambda p_i t$ 的大小以不同方式上升。

D. 重尾（齐夫定律）输入：延迟饱和与长尾

如果信息输入的统计具有重尾，内部时间饱和的方式就会发生显著变化。作为重尾的一个典型例子，考虑齐夫定律（一种幂律秩频分布） [11, 13]。这里，对于代码的秩 i （$i = 1$ 出现频率最高；

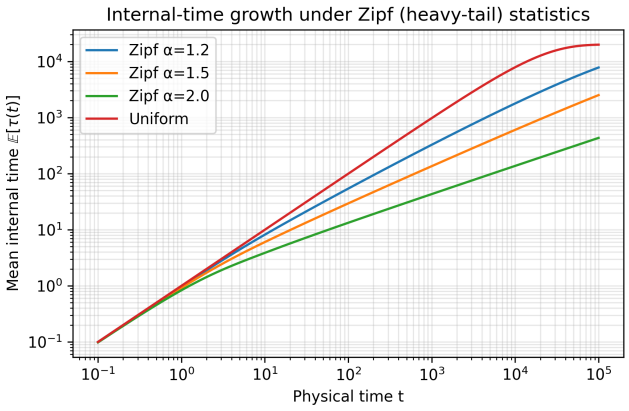


图 3. 齐夫定律输入中的平均内部时间 $\mathbb{E}[\tau(t)]$（式（13））（式（14））。由于重尾，趋向饱和的后期尾部往往会变长。

较大的 i 表示出现频率较低），假设采用截断齐夫分布

$$p_i = \frac{i^{-\alpha}}{\sum_{j=1}^{K} j^{-\alpha}}, \qquad i = 1, \ldots, K,\ \alpha > 0 \tag{13}$$

（关于截断分布估计的实用说明，参见 [2]）。

在泊松到达下，编码 i 截至时间 t 至少被观测到一次的概率为 $q_i(t) = 1 - e^{-\lambda p_i t}$，因此平均内部时间为

$$\mathbb{E}[\tau(t)] = \sum_{i=1}^{K} \left(1 - e^{-\lambda p_i t}\right) \tag{14}$$

（式（3）的一个具体例子）。在齐夫定律下，存在许多具有较小 p_i 的稀有代码，因此它们往往会在很长一段时间内仍未被观测到。因此，尽管 $\mathbb{E}[\tau(t)]$在早期阶段相对快速地增加，但在后期阶段，“最后几种”很难填满，趋向饱和的尾部会变长。

a. 中间时间窗口内的增加（定性标度） 为了直观理解式（14）中间时间窗口内的增加，可根据 $\lambda p_i t$ 是否超过 1 来划分各项贡献。对于满足$\lambda p_i t \gtrsim 1$的秩不超过 i 的编码，有 $q_i(t) \approx 1$；而对于更罕见的编码，有$q_i(t) \approx \lambda p_i t$。在齐夫定律 $p_i \propto i^{-\alpha}$ 下，边界秩$i_c(t)$往往近似满足

$$\lambda p_{i_c} t \approx 1 \tag{15}$$

那么，
$i_c(t)$趋于按 t的某个幂成比例增加。

因此，在饱和之前的中间时间窗口内，$\mathbb{E}[\tau(t)]$ 在 t 中可能呈现出慢于线性的增长（幂律式增长）。特别是，随着 α 增加，概率会集中在少数几个排名靠前的编码上，因此早期增长会很快，但观测到稀有代码会变得更加困难，后期的尾部也往往会变长。

内部时间 $\tau(t)$仅保留“观测代码的种类数量”，因此，即使 $\tau(t)$相同，“观测到了哪些代码”通常也无法确定。为了衡量尚存的信息，我们使用条件熵 $H(S(t) \mid \tau(t))$。引入这一条件熵，是为了以辅助方式直观呈现“仅保留内部时间时会丢失什么”。

E. 定义与恒等式

设截至时间 t 的代码序列为 $X^{N(t)} =(X_1、\ldots\ldots、X_{N(t)})$。为比较不同的约简，定义以下两个条件熵:

$$L_A(t) := H\left(X^{N(t)} \mid S(t)\right), \qquad L_B(t) := H\left(X^{N(t)} \mid \tau(t)\right),$$

其中，$H(\cdot)$是香农熵（自然对数）[3, 12]。引入差值

$$U(t) := L_B(t) - L_A(t).$$

那么，根据链式法则，有如下结论。

定理 2 (恒等式). 在任意概率模型下，

$$U(t) = H(S(t) \mid \tau(t)). \tag{16}$$

证明。 $\tau(t)$ 是 $S(t)$ 的函数，而 $S(t)$ 是 $X^{N(t)}$ 的函数。根据链式法则，

$$H\left(X^{N(t)}, S(t) \mid \tau(t)\right) = H(S(t) \mid \tau(t)) + H\left(X^{N(t)} \mid S(t), \tau(t)\right).$$

另一方面，由于 $S(t)$ 是 $X^{N(t)}$ 的函数，由 $H(S(t) \mid X^{N(t)}, \tau(t)) = 0$ 可得

$$H\left(X^{N(t)} \mid \tau(t)\right) = H\left(X^{N(t)}, S(t) \mid \tau(t)\right).$$

此外，由于 $\tau(t)$ 是 $S(t)$ 的函数，$H(X^{N(t)} \mid S(t))$ 是 t 的函数，因此，

$$H(X^{N(t)} \mid \tau(t)) = H(S(t) \mid \tau(t)) + H(X^{N(t)} \mid S(t)),$$

两边同时减去 $H(X^{N(t)} \mid S(t))$，即可得到结论。 □

a. 恒等式的含义 $U(t)$衡量的是“保留已观测代码列表 $S(t)$” 与 “ 仅保留内部时间 $\tau(t)$” 之间的差异。恒等式（16）表明，这种差异恰好等于“仅知道内部时间时，已观测代码列表中仍存在的不确定性”。因此，$H(S(t) \mid \tau(t))$是一个自然的量，用来描述简化为内部时间时丢失的信息。

F. 均匀分布：闭式

考虑到达过程为泊松过程且代码均匀分布的情形 $p_i = 1/K$。根据泊松分裂，

$$N_i(t) \sim \text{Poisson}\left(\frac{\lambda t}{K}\right) \text{ independently} \tag{17}$$

$$\mathbb{P}(i \in S(t)) =: q(t) = 1 - e^{-\lambda t/K}. \tag{18}$$

因此，

$$\tau(t) = \sum_{i=1}^{K} \mathbf{1}\{N_i(t) \geq 1\} \sim \text{Binomial}\big(K, q(t)\big).$$

定理3（均匀分布下的条件等概率性）。在均匀分布下，以 $\tau(t) = m$ 为条件时，所有“从 K 个代码中选取 m 个代码所得到的候选项”都具有相同的可能性。因此，

$$H(S(t) \mid \tau(t) = m) = \log \binom{K}{m}, \tag{19}$$

and

$$H(S(t) \mid \tau(t)) = \mathbb{E}\left[\log \binom{K}{\tau(t)}\right] \tag{20}$$

成立。

引理1（对称性）。将 (20) 视为关于 $q \in [0,1]$ 的函数，其中 $U(q)$，我们有

$$U(q) = U(1-q).$$

定理4（单峰性与最大化时间）。在均匀分布下，$H(S(t) \mid \tau(t))$ 关于 t 是单峰的，且唯一的最大化点为

$$t^* = \frac{K}{\lambda} \log 2 \tag{21}$$

t^* 是辅助量 $H(S \mid \tau)$达到最大值时的时间，表示这样一个中间阶段：“仅知道内部时间时，已观测代码列表的候选项数量最多。”主要结果（减速与饱和）是一种长时间尺度结构，其中包含 $t \to \infty$时的行为，而 t^* 则量化了另一种视角（对约简的识别不确定性）。

G. 均匀分布：渐近求值（斯特林 + 集中性）

在均匀情形下，

$$\tau \sim \quad K\ q \quad q = q\ t \ = 1 - \text{Bin}(,), \ ( ) \text{。}$$

$e^{-\lambda t/K}$ $m = Kx$

令。根据二项式系数的斯特林近似（参见标准概率论教材，[6]），

$$\log \binom{K}{Kx} = K\,h(x) - \frac{1}{2}\log\big(2\pi K x(1-x)\big) + o(1), \tag{22}$$

$$h(x) := -x \log x - (1-x)\log(1-x). \tag{23}$$

此外，由于 τ 集中在均值 $\mu = Kq$和方差 $\sigma^2 = Kq$附近（$1-q$），对凹函数 $f(m) =\log \binom{K}{m}$ 应用二阶展开（增量法），可得

$$H(S(t) \mid \tau(t)) = \mathbb{E}[f(\tau)] = f(\mu) + \frac{1}{2} f''(\mu)\sigma^2 + o(1). \tag{24}$$

Using $f''(\mu) \approx -\{(Kq)^{-1} + (K(1-q))^{-1}\} = -1/(Kq(1-q))$,

$$H(S(t) \mid \tau(t)) = K\,h(q) - \frac{1}{2}\log\big(2\pi K q(1-q)\big) - \frac{1}{2} + o(1), \tag{25}$$

$(K \to \infty,\ q \in (\varepsilon, 1-\varepsilon))$.

I 在更晚阶段，对于有限的$\to \infty$， K和 $q \to 1$会集中；又因为 $\log \binom{K}{\tau} \to 0$，我们得到$H(S(t) \mid \tau(t)) \to 0$。

VI. 非均匀分布：上界、等号成立条件、差距表示与近似

本节将说明当编码分布为非均匀分布时，$H(S(t) \mid \tau(t))$如何变化。这里同样需要强调，本节讨论的是辅助量的性质，而非主要结果（减速与饱和）。主要结果本身已在§III中证明，而非均匀性在其中表现为 $\mathbb{E}[\tau(t)]$ 的扭曲（见§V）。

A. 普适上界（与模型无关 ent)

对于任意 m，在 $\tau(t) = m$ 条件下，$S(t)$ 可取的候选项数量为 $\binom{K}{m}$。因此，

$$H(S(t) \mid \tau(t) = m) \leq \log \binom{K}{m}, \tag{26}$$

and

$$H(S(t) \mid \tau(t)) \leq \mathbb{E}\left[\log \binom{K}{\tau(t)}\right] \tag{27}$$

成立。只有当“在 $\tau(t) = m$ 条件下，各候选项出现的概率相同”时，等式才可能成立。

B. 泊松到达下的等式条件

假设到达过程服从泊松分布，且代码序列是独立同分布的。那么

$$N_i(t) \sim \text{Poisson}(\lambda p_i t) \text{ independently}, \tag{28}$$

$$q_i(t) := \mathbb{P}(i \in S(t)) = 1 - e^{-\lambda p_i t}. \tag{29}$$

因此， $I_i(t) = \mathbf{1}\{i \in S(t)\}$ 是相互独立的伯努利变量（q_i）。对于任意 $A \subseteq \{1, \dots, K\}$,

$$\mathbb{P}(S(t) = A) = \prod_{i \in A} q_i \prod_{j \notin A} (1 - q_j),$$

因此，在给定 $\tau(t) = m$ 的条件下,

$$\mathbb{P}(S(t) = A \mid \tau(t) = m) \;\propto\; \prod_{i \in A} \frac{q_i}{1 - q_i}.$$

定理5 （等式条件：均匀分布是唯一的达成者（泊松到达））。 在泊松到达下，对于任意 $m \in \{1, \dots, K-1\}$,

$$H(S(t) \mid \tau(t) = m) = \log \binom{K}{m}$$

当且仅当 $q_1(t) = \cdots = q_K(t)$ 时成立。此外，根据 $q_i(t) = 1 - e^{-\lambda p_i t}$ 的单调性，这等价于

$$p_1 = \cdots = p_K = 1/K.$$

当“候选项尽可能均匀地分布”时，上界（27）可以达到。在泊松到达下，只有当所有 $q_i(t)$ 都相等时才会出现这种情况，这等价于 p_i 的均匀性。因此，对于非均匀分布，仅知道内部时间时，候选项会带有偏差，且条件熵小于均匀情形下的条件熵。

C. 与上界的差距：条件KL表示

将与上界的差定义为

$$G(t) := \mathbb{E}\left[\log \binom{K}{\tau(t)}\right] - H(S(t) \mid \tau(t)).$$

令 U_m 为“从 K 中选择 m 所得到的候选者”上的等概率分布，并令 $P_m(A) := \mathbb{P}(S(t) = A \mid \tau(t) = m)$。于是

$$\log \binom{K}{m} - H(S \mid \tau = m) = D_{\mathrm{KL}}(P_m \| U_m),$$

因此

$$G(t) = \mathbb{E}\big[D_{\mathrm{KL}}(P_{\tau(t)} \| U_{\tau(t)})\big] \geq 0. \tag{30}$$

差距 $G(t)$ 衡量的是“候选项偏离等概率状态的程度”，取其平均值。对于非均匀分布， $G(t) > 0$ ，因为容易观测到的代码往往会进入已观测代码列表，从而使结果偏向容易输入的组合。

D. 分解恒等式及其近似（泊松二项分布）

在泊松到达的情况下，由于 $I_i \sim$伯努利$(q_i(t))$ 相互独立,

$$H(S(t)) = H(I_1, \dots, I_K) = \sum_{i=1}^{K} h(q_i(t)).$$

根据链式法则 $H(S) = H(\tau) + H(S \mid \tau)$ 和（16），我们有:

定理6 （分解恒等式）。 在泊松到达下,

$$H(S(t) \mid \tau(t)) = \sum_{i=1}^{K} h(q_i(t)) - H(\tau(t)), \qquad q_i(t) = 1 - e^{-\lambda p_i t}. \tag{31}$$

这里 $\tau(t) = \sum I_i$ 服从泊松二项分布（计算与近似: [5, 8]）。

$$\mu(t) = \mathbb{E}[\tau(t)] = \sum q_i(t), \tag{32}$$

$$\sigma^2(t) = \mathrm{Var}(\tau(t)) = \sum q_i(t)(1 - q_i(t)). \tag{33}$$

在 $\sigma^2(t)$ 足够大的区域,

$$H(\tau(t)) \approx \frac{1}{2} \log(2\pi e\, \sigma^2(t))$$

可用作 [3] 的近似。因此,

$$H(S(t) \mid \tau(t)) \approx \sum_{i=1}^{K} h(q_i(t)) - \frac{1}{2} \log(2\pi e\, \sigma^2(t)). \tag{34}$$

在 λt 较小的早期阶段，由于 $q_i(t) \approx \lambda p_i t$,

$$\sum_{i=1}^{K} h(q_i(t)) \approx \lambda t \Big[\log \frac{1}{\lambda t} + 1 + H(p)\Big], \qquad \sigma^2(t) \approx \lambda t$$

可得，且 $H(p)$ 出现在初始增长率中。

E. 最大化时间

a. 均匀分布（精确）在均匀情形下，$H(S(t) \mid \tau(t))$ 是单峰的，最大化时间由式（21）给出。

b. 非均匀分布（通常没有闭式） 对于非均匀分布，最大化时间通常没有闭式。可以直接对（31）进行最大化，或根据近似式（34）的一阶条件对其进行近似确定。

VII. 讨论

本节将从直观和方程两个方面梳理主要结果“内部时间的减慢与饱和”源于哪些假设，以及它受输入统计的哪些特征控制。我们还将描述重尾（齐夫定律）输入所产生的长尾的含义，以及加权内部时间推广的定位。对于作为辅助量引入的条件熵，我们将阐明它并非本文的核心，并说明在哪个范围内它可作为辅助指标发挥作用。

A. 重申本文核心：减速与饱和是结论

本文的核心结论是：响应系统所计量的内部时间 $\tau(t)$（观测代码的种类数量）受输入统计（到达率 λ 和编码分布 p）调控；平均推进速度随时间单调递减，并在有限 K下达到饱和。这一结论可直接从泊松到达且编码序列独立同分布的标准设定中得出：平均内部时间由式（3）给出，其时间导数（平均推进速度）具有式（4）的形式。也就是说，推进速度是 $p_i e^{-\lambda p_i t}$之和；由于每一项都随 t单调递减，平均推进速度必然减慢；此外，由于$q_i(t) \to 1$，$\mathbb{E}[\tau(t)] \to K$ 并达到饱和。

从这个意义上说，“内部时间变得相对短于物理时间”并非一种比喻，而是可以表述为严格的结论 $\lim_{t\to\infty} \mathbb{E}[\tau(t)]/t = 0$，它由 $\mathbb{E}[\tau(t)]$的有界性推得。这一结论适用于一般分布 p，而不同分布会改变“它如何减慢以及尾部如何保持”。

B. 输入统计决定减速曲线：λ 是时间尺度，p是形状

根据（4），到达率 λ主要决定“时间尺度”。例如，若用 $u = \lambda t$对时间进行无量纲化，减速的基本形式可由 u 描述。另一方面，编码分布 p 决定减速的形状。由于高概率编码很可能较早被观测到，而低概率编码往往会长时间未被观测到，因此，$\mathbb{E}[\tau(t)]$ 中“早期上升”“中期增长”和“后期尾部”的时间分配会随 p而改变。本模型的一个优点是，将内部时间限定为“种类数量”这一最简记录后，输入统计便能直接显现。

此外，在均匀情形下，物理时间与内部时间之间的对应关系可以写成闭式。逆映射（11）表明，达到内部时间水平 τ 所需的物理时间随着 $\tau \to K$ 发散，而增量（12）具体说明，“内部时间每前进一个单位所需的物理时间”在后续阶段会增加。“所需物理时间在后续阶段增加”这一表述，以最直观的方式传达了减速与饱和的内涵。

C. 重尾（齐夫定律）输入：长尾并非“例外”，而是统计规律的结果

对于重尾分布，由于稀有编码数量众多，后期趋向饱和的尾部会变长。本文以齐夫定律作为一个典型示例（概述 [11], 原文 [13]）。即使在截断齐夫分布（13）下，平均内部时间仍可由（14）以具体形式表示为一般公式（3），而模型的核心（减速和饱和）并未改变。改变的是趋向饱和的时间分配。

特别是在齐夫定律下，由于存在许多概率很小的编码，“经过相当长的时间后仍有未观测到的编码”这一现象自然会出现。这意味着，即使对于有限的 K，饱和也可能极其缓慢，并可作为一种最简模型，用来解释真实数据（自然语言、行为序列等）中“新事物永远不会完全耗尽”的印象。由于重尾分布的估计和截断处理需要统计谨慎，因此在应用于真实数据时，[2] 中的讨论很有用。

D. 加权内部时间：种类数量模型的自然推广

本文采用 $\tau(t)$（种类数量）作为内部时间的最简定义，但若要纳入“越罕见的编码携带的信息越多”这一标准观点，则加权内部时间 $T(t)$ 是自然的（§IV）。特别是，假设权重 ℓ_i 与分布 p_i 之间满足式（10），则与自信息 $-\log p_i$ 相一致，而 $T(t)$ 可解释为“已观测代码列表保留的信息量”的代表性量。这种一致性基于信息论 [3, 12], 的基本原理，并提供了一种一致的推广，其中包含种类数量模型（τ）。

加权内部时间的均值由（8）给出，其推进速度由（9）给出。与 $\tau(t)$ 类似，它的均值会减速，并在 K 有限时趋于饱和。然而，由于在较晚阶段观测到稀有代码时，其贡献更大，因此对于重尾输入，$T(t)$ 的“后期尾部”可能比 $\tau(t)$ 的后期尾部更加突出。在既希望根据种类数量、也希望根据信息量来定义内部时间的情形下，这一点很重要。

E. 辅助量（条件熵）的定位：化简为内部时间后舍弃了什么

条件熵 $H(S(t) \mid \tau(t))$ 并非支撑本文核心内容（减速与饱和）的假设或结论，而是一个辅助量，用来衡量“仅知道内部时间时，已观测代码列表中还剩下多少不确定性”。恒等式（16）表明，约化差 $U(t)$ 等于这一条件熵，并从熵的链式法则 [3, 12] 的基本结构出发，对约化为内部时间时丢失的信息进行了梳理。

在均匀情形下，候选项变为等概率，从而得到闭式（20），并可推出单峰性和最大化时间（21）。在非均匀情形下，可以整理出普适上界（27）及其等号成立条件，并将与上界的差异表示为 KL散度的均值（30）。本节无意取代“减速与饱和”，而是作为一项辅助结果，量化采用内部时间作为最小记录时，哪些信息会变得无法区分。

F. 局限与扩展：最小模型的适用范围

由于本文有意采用最小模型，因此存在以下局限。第一，内部时间定义为“观测代码的种类数量（或加权和）”，不会保留更详细的历史信息，例如频率、顺序或到达间隔。第二，为便于分析，到达过程被限定为泊松过程。第三，本文不假定其与外部心理指标或行为指标相对应。

另一方面，后续研究方向十分明确。(i) 将到达过程推广至非泊松情形（周期性、自激等），并考察减速的普适性及其失效方式（点过程的一般理论：[4, 9]）。(ii) 稍微扩展内部状态，将频率信息或近期历史纳入其中；设计包含此类信息的内部时间，并探究它与本文所述减速和饱和现象之间的联系。(iii) 考察重尾极限 $K \to \infty$，并整理饱和实际上无法观测到的区域中的标度问题。这些都是值得开展的自然研究课题，可用于考察依赖于输入统计的内部时间减速与饱和这一核心现象在多大程度上具有普遍性。

VIII. 结论

本文针对一种根据信息输入更新其内部状态的响应系统，提出了一个最小模型，将观测到的编码种类数定义为内部时间，并根据信息输入统计，严格推导出内部时间的减慢与饱和。在泊松到达的情况下，内部时间的平均推进速度单调下降；若编码种类数量有限，则该速度会趋于饱和。因此，在较长的时间尺度上，内部时间相较于物理时间会变得更短。在均匀情形下，我们给出了物理时间与内部时间之间的反向关系，并明确表明，到了后期，内部时间会变得更难推进。作为推广，我们还引入了依赖于代码的权重（描述长度），并给出了根据信息量定义内部时间的公式。

作为辅助量，我们引入了条件熵，并量化了仅知道内部时间时仍然存在的不确定性。对于均匀情形，我们整理了闭式表达、单峰性和最大化时间；对于非均匀情形，则整理了上界、等号成立条件、间隔表示和近似计算。这些结果独立于减速与饱和这一主要结论，但为理解约化到内部时间时所舍弃的信息提供了辅助视角。

附录 A：附录 A：均匀分布中的单峰性与最大化点

在均匀情形下，$\tau \sim \mathrm{Bin}(K,\ q)$，且

$$U(q) = H(S \mid \tau) = \mathbb{E}\left[\log \binom{K}{\tau}\right] \tag{A1}$$

$$= \sum_{m=0}^{K} \binom{K}{m} q^m (1-q)^{K-m} \log \binom{K}{m}. \tag{A2}$$

这是离散凹函数 $f(m) = \log \binom{K}{m}$ 的伯恩斯坦多项式。

a. 对称性 由 $\binom{K}{m} = \binom{K}{K-m}$ 及二项权重的对称性可知，(A2) 意味着 $U(q) = U(1-q)$。

b. 严格凹性（概要）序列 $m \mapsto \log \binom{K}{m}$ 在 $m = 0, \ldots\ldots, K$ 上离散严格凹，且当 $1 \le m \le K-1$ 时其二阶差分为负。根据伯恩斯坦算子的标准性质，伯恩斯坦多项式的二阶导数可表示为二阶差分的正权重平均（见 [10]）。因此，对于 $q \in (0,\ 1)$，$U''(q) < 0$ 成立，并且 $U(q)$ 在 $(0,\ 1)$ 上严格凹。因此它是单峰的。

c. 最大化点 根据严格凹性以及关于 $q = 1/2$ 的对称性，最大化点唯一为 $q^* = 1/2$。由于 $q(t) = 1 - e^{-\lambda t/K}$，

$$1 - e^{-\lambda t^*/K} = \frac{1}{2}, \qquad t^* = \frac{K}{\lambda} \log 2.$$

[1]Richard A. Block、Dan Zakay 和 Peter A. Hancock。人类老化与时长判断：元分析综述。 *Psychology and Aging*，13(4):584–596，1998。[2] Aaron Clauset、Cosma Rohilla Shalizi 和 M. E. J. New- man。经验数据中的幂律分布。 *SIAM Review* 51 4 661 703, 2009 10 1137 070710111 ，( ): –。doi: 。/ 。[3] Thomas M. Cover 和 Joy A. Thomas。 *Elements of Information Theory* 2 2006 。Wiley，版次，。[4] D. J. Daley 和 D. Vere-Jones。 *An Introduction to the Theory of Point Processes Volume I: Elementary Theory , and Methods* 2 2003 。Springer，版次，。[5] Werner Ehm。泊松二项分布的二项近似。 *Statistics & Probability Letters*，11 1 7 16，1991 10 1016 0167 7152 91 90170 ( ); – doi: . / - ( ) -V。[6] William Feller。 *An Introduction to Probability Theory and Its Applications Volume I* 2 1957 ，。Wiley，版次，。[7] William J. Friedman。老化与时间速度。 *Acta Psychologica*，134(2):130–141，2010。doi: 10.1016/j.actpsy.2010.02.004. [8] Yili Hong. 关于计算泊松二项分布的分布函数。 *Computational Statistics & Data Analysis* 59 41 51, 2013 , : – 。doi: 10.1016/j.csda.2012.10.006. [9] J. F. C. Kingman. *Poisson Processes*. Clarendon Press（牛津大学出版社），1993 年。[10] G. G. Lorentz. *Bernstein Polynomials*. Chelsea Publishing Company，修订版，1986 年。[11] M. E. J. Newman. 幂律、帕累托分布和齐普夫定律。 *Contemporary Physics*, 46(5):323–351, 2005. doi:10.1080/00107510500052444. [12] Claude E. Shannon. 通信的数学理论。 *Bell System Technical Journal* 27 3 379 423, , 1948 10 1002 1538 7305 1948 01338 ( ): – 。doi: . /j. - . .tb .x.[13] George Kingsley Zipf. *Human Behavior and the Principle of Least Effort* . Addison-Wesley Press，马萨诸塞州剑桥，1949 年。