arXiv:2601.06317v1 [math.ST] 9 Jan 2026

积分模型中截距参数的估计

Galton-Watson 过程中的截距参数估计

Yang Lu*

2026年1月13日

摘要

我们研究积分 Galton-Watson 过程中的截距参数估计问题。该过程是许多计数值时间序列模型的基本构件。在这一单位根情形下，普通最小二乘估计量不一致，而现有的加权最小二乘（WLS）估计量仅在过程为暂态时才一致；这一条件取决于未知的截距参数。我们提出一种基于 $1/t$ 的新权函数构造的替代 WLS 估计量，并证明无论过程是暂态还是零常返，该估计量都一致，其收敛速度为 $\sqrt{\ln n}$。

关键词：过程的收敛性、非平稳性、整数值时间序列、紧性。

数学主题分类代码：62M10

经济文献分类代码：
C22。
致谢：感谢加拿大自然科学与工程研究委员会提供资助，项目编号为 RGPIN-2021-04144。

数据可用性声明：未生成任何数据。

利益冲突声明：作者声明不存在利益冲突。

1 引言

近年来，人们对非平稳整数值过程的研究兴趣日益浓厚（Barczy 等，2011，2014；Michel，2020；Pei 和 Lu，2025）。许多这类模型可以表示为高阶

*加拿大蒙特利尔康考迪亚大学数学与统计系。电子邮箱：yang.lu@concordia.ca

Galton-Watson 过程或含移民分枝过程的扩展:

$$X_0 = 0, \qquad X_t = \sum_{i=1}^{X_{t-1}} Z_{i,t} + \epsilon_t, \forall t > 0, \tag{1}$$

其中，$Z_{i,t}, i, t$ 随时间变化，且相互独立同分布，均值为 m，方差为 $\sigma^2 > 0$； ϵ_t, t 随时间变化，且相互独立同分布，均值为 $\mu > 0$，方差为 $b > 0$，并且与序列 $(Z_{i,t})$ 相互独立。该模型可得:

$$\mathbb{E}[X_t \mid X_{t-1}] = mX_{t-1} + \mu,$$

若过程 (X_t) 满足 $m = 1$，则称其具有单位根，或称其为积分过程。这类积分过程包括 INARCH 模型（其中$Z_{i,t}$ 和 ϵ_t 均服从泊松分布）、Gouri´eroux 和 Lu（2019年）提出的 NBAR模型（其中 $Z_{i,t}$ 服从几何分布， ϵ_t 服从负二项分布，且二者具有相同的概率参数），以及许多其他模型 [，例如参见 Lu（2021）及其中的参考文献]。然而，在单位根条件 $m = 1$下，INAR(1)模型不在此列。事实上，当 $m = 1$时，我们几乎必然有 $Z_{i,t} = 1$ 。因此，其方差 σ^2 等于零，这是不允许的。在这种情况下，INAR(1)模型可写为： $X_t = X_{t-1} + \epsilon_t$。于是，$\mu = \mathbb{E}[\epsilon_t]$ 可通过 $\frac{X_n}{n}$以标准参数收敛速度 $\sqrt{n}$直接估计。

这方面文献中的一个重要空白是，迄今为止，对于许多这类非平稳模型，普通最小二乘法（OLS）估计量并不具有一致性。具体而言，至少从 Winnicki（1986年）起，人们就已知道，基准马尔可夫模型（1）的 (m, μ) 的 OLS 估计量定义如下:

$$(\hat{m}, \hat{\mu})' := \arg\min_{(m,\mu)'} \sum_{t=1}^{n} (X_t - mX_{t-1} - \mu)^2, \tag{2}$$

该估计量使得 $\hat{}\mu$ 不一致（其大样本性质见下文定理 3）。为解决这一不一致问题，Wei and Winnicki（1990）考虑了以下加权最小二乘估计量:

$$(\hat{\hat{m}}, \hat{\hat{\mu}})' := \arg\min_{(m,\mu)} \sum_{t=1}^{n} \frac{(X_t - mX_{t-1} - \mu)^2}{1 + X_{t-1}}. \tag{3}$$

他们只能证明 [参见其定理2.5及（Wei，1991，定理2.1）]$\hat{}\mu$ 的一致性（收敛速度为 $\sqrt{\ln n}$），但需要对参数施加以下附加条件:

$$\tau_0 := \frac{2\mu_0}{\sigma_0^2} > 1, \tag{4}$$

wh这里， μ_0 （分别为 σ_0）表示 μ （分别为 σ）的真实参数值。该不等式可以

被解释为（Wei，1991）马尔可夫链（X_t）的暂留条件，而相反的不等式 $\tau_0 \leq 1$ 则是该链零常返的条件。他们指出（见其注释 2.6）：“当 $m_0 = 1$ 且 $\tau_0 \leq 1$时，尚不清楚 $\hat{}\mu$ 的极限分布是什么。”主要困难似乎在于量化诸如$\sum_{t=1}^{n} \frac{1}{1+X_{t-1}}$ 和 $\sum_{t=1}^{n} \frac{X_t - X_{t-1} - \mu_0}{1+X_{t-1}}$（如果存在）等量的发散速度。例如，Wei and Winnicki（1989）指出，如果 $\tau_0 = 1$，那么 $\sum_{t=1}^{n} \frac{1}{1+X_{t-1}}$ 的收敛速度“似乎更难确定；目前只给出了某些上界和下界”。[1] 总之，迄今为止，加权最小二乘（3）的渐近理论尚未完全明了；即使它收敛，其收敛速度也可能取决于未知参数值 τ_0，这给统计推断带来不便。

本文针对模型（1）中的 μ，提出一种替代的加权最小二乘估计量，使用权重 $\frac{1}{t}$ 而非 $\frac{1}{1+X_{t-1}}$。其动机在于，给定 X_{t-1} 时，X_t 的条件方差为 $\sigma_0^2 X_{t-1} + b_0$；在单位根假设下，可以证明，随着 t 趋于无穷大，$X_t \sim \mu_0 t$，因此选择权重 $\frac{1}{t}$。与 $\sum_{t=1}^{n} \frac{1}{1+X_{t-1}}$ 相比，调和级数 $\sum_{t=1}^{n} \frac{1}{t}$ 更易于处理，因为随着 n 趋于无穷大，$\sum_{t=1}^{n} \frac{1}{t} = \ln n + \gamma_E + o(\frac{1}{n})$，其中 $\gamma_E \approx 0.577$ 为欧拉常数。我们证明，对于模型（1），由此得到的新加权最小二乘估计量是一致的，无需额外的暂态条件（4）。我们还讨论了这一结果对统计推断的影响。

本文其余部分安排如下。第 2 节提出新的估计量，并推导其在积分 Galton-Watson 过程下的渐近性质。第 3 节讨论这些新结果的意义。第 4 节总结全文。技术性证明汇集于附录中。

2 估计量

我们考虑 $1/t$ 加权的加权最小二乘估计量：

$$(\tilde{m}, \tilde{\mu})' := \arg \min_{(m,\mu)'} \sum_{t=1}^{n} \frac{1}{t} (X_t - mX_{t-1} - \mu)^2. \tag{5}$$

[1] 在 $\tau_0 < 1$ 的情况下，Wei and Winnicki（1989年，定理 2.18）证明，$n^{-1+\tau_0} \sum_{t=1}^{n} \frac{1}{1+X_{t-1}}$ 依分布收敛于 $cW_{1-\tau_0}$，其中 c 是常数，而随机变量 W_q 服从参数为 q 的米塔格-莱弗勒分布。然而，Wei and Winnicki（1989年）并未给出常数 c 的值。此外，$\sum_{t=1}^{n} \frac{X_t - X_{t-1} - \mu_0}{1+X_{t-1}}$ 的渐近行为也尚不清楚。

将 $\tilde{m}$ 和 $\tilde{\mu}$ 分别以其真实值 1 和 μ_0 为中心化处理后，得到：

$$\begin{bmatrix} \tilde{m}-1 \\ \tilde{\mu}-\mu_0 \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{n} \frac{1}{t} X_{t-1}^2 & \sum_{t=1}^{n} \frac{1}{t} X_{t-1} \\ \sum_{t=1}^{n} \frac{1}{t} X_{t-1} & \sum_{t=1}^{n} \frac{1}{t} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{n} X_{t-1} W_t \frac{1}{t} \\ \sum_{t=1}^{n} W_t \frac{1}{t} \end{bmatrix} \tag{6}$$

其中

$$W_t = X_t - X_{t-1} - \mu_0$$

是一个鞅差序列。我们还回顾积分 Galton-Watson 过程的以下著名标度性质 [见（Wei and Winnicki，1989年，定理2.1）]：随着 n趋于无穷，经过重新缩放和时间变换的过程（$\tilde{X}_{n,\ s}$）$=$（$\frac{X_{\lfloor ns \rfloor}}{n}$，$s>0$）在非负右连续且有左极限函数空间 $[0,\infty[$, 上的斯科罗霍德空间$\mathcal{D}^+([0,\infty[)$ 中弱收敛于 Cox-Ingersoll-Ross（CIR）扩散过程：

$$\mathrm{d}Y_s = \mu_0 \mathrm{d}s + \sigma_0 \sqrt{Y_s} \mathrm{d}B_s, \forall s > 0, \qquad Y_0 = 0, \tag{7}$$

其中（B_s）是标准布朗运动。对于我们的估计问题，通过引入连续时间鞅（$M_{n,\ s}$）$_s$，可以轻松地加强这一结果：

$$M_{n,s} = \frac{1}{n} \sum_{t=1}^{\lfloor ns \rfloor} W_t = \frac{X_{\lfloor ns \rfloor}}{n} - \frac{\lfloor ns \rfloor}{n} \mu_0, \qquad \forall s > 0.$$

于是，根据连续映射定理（CMT），我们得到该过程的联合弱收敛：

$$\left(\frac{X_{\lfloor ns \rfloor}}{n}, M_{n,s} \right) \Rightarrow \left(Y_s, M_s = Y_s - \mu_0 s = \sigma_0 \int_0^s \sqrt{Y_s} \mathrm{d}B_s \right). \tag{8}$$

随后，WLS 估计量的分析基于以下两个定理。

定理 1。 当 n 趋于无穷大时，

- $a)$ $\frac{1}{n} \sum_{t=1}^{n} \frac{1}{t} X_{t-1} \Rightarrow \int_0^1 \frac{Y_s}{s} \mathrm{d}s$，其中 $\Rightarrow$ 表示弱收敛
- $b)$ $\frac{1}{n^2} \sum_{t=1}^{n} \frac{1}{t} X_{t-1}^2 \Rightarrow \int_0^1 \frac{Y_s^2}{s} \mathrm{d}s$,
- $c)$ $\frac{1}{n} \sum_{t=1}^{n} \frac{1}{t} W_t X_{t-1} \Rightarrow \sigma_0 \int_0^1 \frac{Y_s^{3/2}}{s} \mathrm{d}B_s$
- $d) \frac{1}{\sqrt{\ln n}} \sum_{t=1}^{n} \frac{1}{t} W_t \Rightarrow Z$，其中 Z服从正态分布 $N(0,\ \sigma_0^2 \mu_0)$，且与过程(Y_r)相互独立
- $e)$ 在 $a) - d)$ 中的弱收敛也构成联合弱收敛

我们注意到，a) b) c)中的弱极限都存在，因为 $\frac{Y_s}{s}$在 $[0,1]$上连续。事实上，当s 接近于零时，有 $Y_s = \mu_0 s + \sigma_0 \int_0^s \sqrt{Y_u}\mathrm{d}u$，因此当 s 递减至 0 时，$\frac{Y_s}{s} \to \mu_0$ 几乎必然成立。

定理 1 的证明相当繁琐，涉及若干随机过程极限定理。下面我们给出一些直观解释，并将严格证明放在附录中。对于 a)，我们有：

$$\frac{1}{n}\sum_{t=1}^{n}\frac{1}{t}X_{t-1} = \int_{1/n}^{1+1/n}\frac{X_{\lfloor ns\rfloor}}{\lfloor ns\rfloor}\mathrm{d}s.$$

接下来要证明的是，后者收敛到：$\int_0^1 \frac{Y_s}{s}\mathrm{d}s$. 困难在于，由于零点处的奇异性，$(\frac{X_{\lfloor ns\rfloor}}{n}) \Rightarrow Y_s$ 并不能直接推出在整个定义域 $(0, \infty)$ 上有 $(\frac{X_{\lfloor ns\rfloor}}{\lfloor ns\rfloor}) \Rightarrow Y_s/s$。换言之，无法直接应用连续映射定理。相反，我们必须控制 $\frac{X_{\lfloor ns\rfloor}}{\lfloor ns\rfloor}$ 在零点附近的行为。

b)可以用与证明 a)相同的方法来证明，只需注意：

$$\frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{t}X_{t-1}^2 = \int_{1/n}^{1+1/n}\frac{X_{\lfloor ns\rfloor}}{\lfloor ns\rfloor}\frac{X_{\lfloor ns\rfloor}}{n}\mathrm{d}s.$$

至于 c)，我们有：

$$\frac{1}{n}\sum_{t=1}^{n}\frac{1}{t}W_t X_t = \int_{1/n}^{1+1/n}\frac{X_{\lfloor ns\rfloor}}{\lfloor ns\rfloor}dM_{n,s}. \tag{9}$$

后者是关于鞅 $(M_{n,\ s})$ 的随机积分。思路是，过程 $\frac{X_{\lfloor ns\rfloor}}{\lfloor ns\rfloor}$ 弱收敛到 (Y_s/s)，而 $(M_{n,\ s})$ 弱收敛到 (M_s)，并且这些收敛实际上是联合收敛。然而，这个简单的思路需要进行两项技术处理。首先，与 a)类似，我们需要处理 0 处的奇异性。其次，(9) 是一个随机积分；与确定性积分不同，仅有积分器过程和被积过程的联合弱收敛，并不足以推出随机积分的弱收敛。相反，还需要验证关于鞅 $(M_{n,\ s})$ 的某种一致性条件，称为一致紧性（UT）。

对于 d)，我们写作：

$$\sum_{t=1}^{n} W_t\frac{1}{t} = \sum_{t=1}^{n}\int_{\frac{t}{n}}^{\frac{t+1}{n}}\frac{1}{\lfloor ns\rfloor}\mathrm{d}M_{n,s} \approx \sigma_0\int_{\frac{1}{n}}^{1+\frac{1}{n}}\frac{\sqrt{Y_s}}{s}\mathrm{d}B_s.$$

This 近似是不严谨的，因为上下界和鞅（$M_{n,\ s}$）都取决于

n。假设这一点成立，则随机积分 $\sigma_0 \int_a^1 \frac{\sqrt{Y_s}}{s} \mathrm{d}B_s$ 是以 a 为指标的局部鞅，其二次变差为 $\sigma_0^2 \int_a^1 \frac{Y_s}{s^2} \mathrm{d}s$。同样容易验证，当 a 递减至 0 时，$\frac{\sigma_0^2}{\ln a} \int_a^1 \frac{Y_s}{s^2} \mathrm{d}s$ 收敛于 $\sigma_0^2 \mu_0$，因为当 r 递减至零时，$Y_s/s \to \mu_0$ 几乎必然成立。因此，根据随机积分的中心极限定理（Kutoyants, 2013，定理 1.19），我们得到：

$$\frac{\sigma_0}{\sqrt{\ln n}} \int_{1/n}^1 \frac{\sqrt{Y_s}}{s} \mathrm{d}B_s \Rightarrow N(0, \sigma_0^2 \mu_0) \tag{10}$$

随着 n 趋于无穷大。因此，我们还可以得到：

$$\frac{1}{\sqrt{\ln n}} \sum_{t=1}^n W_t \frac{1}{t} \Rightarrow N(0, \sigma_0^2 \mu_0).$$

此外，随着 n 趋于无穷大，$\frac{1}{\sqrt{\ln n}} \int_{1/n}^1 \frac{\sqrt{Y_s}}{s} \mathrm{d}B_s$ 越来越取决于 (B_s) 在零附近的行为。因此，方程 (9) 和 (10) 中的联合弱极限变量相互独立。

最后，弱收敛 a) 到 d) 蕴含在 e) 中联合弱收敛，这一事实是连续映射定理的直接推论。参见 Wei and Winnicki (1989) 的注 2.4。

下面的定理是定理 1 和连续映射定理的直接推论。

定理 2。 若存在一个正指数 δ，使得 $\mathbb{E}[Z_{i,t}^{2+\delta}]$ 和 $\mathbb{E}[\epsilon_t^{2+\delta}]$ 均有限，则 $1/t$–加权最小二乘估计量 $(\tilde{m}, \tilde{\mu})$ 满足：

$$\begin{bmatrix} n(\tilde{m}-1) \\ \sqrt{\ln n}(\tilde{\mu}-\mu_0) \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{\int_0^1 \sigma_0 \frac{Y_s^{3/2}}{s} \mathrm{d}B_s /}{\int_0^1 \frac{Y_s^2}{s} \mathrm{d}s} \\ Z \end{bmatrix}, \tag{11}$$

随着 n 趋于无穷大，其中 Z 服从正态分布 $N(0, \sigma_0^2 \mu_0)$，且与扩散过程 (Y_s) 相互独立。特别地，我们有：$\sqrt{\ln n}(\tilde{\mu} - \mu_0) \Rightarrow N(0, \sigma_0^2 \mu_0)$。

我们的 $1/t$-加权最小二乘估计量的收敛速度（$\sqrt{\ln n}$）与 Wei and Winnicki (1990) 的 $\frac{1}{1+X_{t-1}}$–加权最小二乘估计量相同，但后者仅在 $\tau_0 > 1$ 时有效。在后一种情形下，（Wei and Winnicki，1990，定理2.5）和（Wei，1991，定理 1.2）表明，他们的加权最小二乘估计量的渐近方差为（$\mu_0 - \frac{\sigma_0^2}{2}$）$\sigma_0^2$，小于我们的加权最小二乘估计量的渐近方差（$\mu_0 \sigma_0^2$）。这意味着，如果确定 $\tau_0 > 1$，那么 Wei and Winnicki (1990) 的加权最小二乘估计量仍然更有效。然而，我们始终可以使用 $\tilde{\mu}$ 作为 μ 的初步估计量。随后，我们可以使用（Winnicki,

1991，定理 3.5）所建议的单位根情形下的相合估计量ˆ，并通过ˆ$\tau = 2\tilde{\mu}/\hat{}\sigma^2$估计 τ，然后检验其是否确实大于 1。如果结果表明情况确实如此，那么我们可以使用 Wei and Winnicki (1990) 的原始估计量作为改进后的估计量。另一方面，如果ˆτ 不大于 1，则只有我们的估计量 $\tilde{\mu}$ 有效。

3 讨论

3.1 与 OLS 估计量的比较

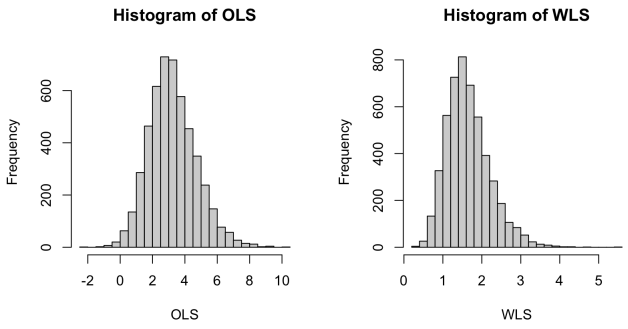
作为比较，对于积分 Galton-Watson 过程，OLS 估计量具有以下性质：

定理 3. *OLS* 估计量 ˆμ 是不一致的，具体来说：

$$\begin{bmatrix} n(\hat{m}-1) \\ \hat{\mu}-\mu_0 \end{bmatrix} \Rightarrow \begin{bmatrix} \int_0^1 Y_s^2 \mathrm{d}s & \int_0^1 Y_s \mathrm{d}s \\ \int_0^1 Y_s \mathrm{d}s & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_0 \int_0^1 Y_s^{3/2} \mathrm{d}B_s \\ \sigma_0 \int_0^1 \sqrt{Y_s} \mathrm{d}B_s \end{bmatrix}.$$

3.2 模拟实验

作为示例，我们考虑这样的模型：给定 X_{t-1}时， X_t 服从参数为 $X_{t-1}+\mu_0$ 的条件泊松分布（因此 $\sigma^2 = 1$）。我们设定 $\mu_0 = 2$，并在样本量为 $n = 100$ 的 $B = 5000$ 个模拟数据集上计算 OLS 和 $1/t$-加权最小二乘估计量。图 1 绘出了这两个估计量的样本直方图。



Fig图 1：OLS 估计量 ˆμ （左图）的样本分布与 e
$1/t$-加权最小二乘估计量（右图）的比较

我们可以看到，与 OLS 估计量相比，加权最小二乘估计量的样本分布范围显著更窄。我们再报告一些汇总统计量。OLS

估计量的样本均值为 3.23（真实值为 2），样本方差为 2.12；而加权最小二乘估计量的样本均值为 1.65，样本方差为 0.31。换言之，尽管 $\sqrt{\ln n}$ 的收敛速度较慢，加权最小二乘估计量仍比 OLS 估计量准确得多。若将样本量 n 分别增加到 $n = 500$ 和 $n = 1000$，则 OLS 估计量的样本均值分别变为 3.40 和 3.42，而加权最小二乘估计量的样本均值分别变为 1.79 和 1.85。因此，随着 n 增加，加权最小二乘估计量的有限样本偏差会减小，而 OLS 估计量的有限样本偏差则不会。这证实了 $1/t$-加权最小二乘估计量是一致的，而 OLS 估计量则不是。

我们还在表1中报告了不同真实μ_0 值和不同样本量下加权最小二乘估计量的样本均值。

μ_0	0.4	0.5	0.8	2	3	10
$n = 100$	0.45	0.53	0.77	1.65	2.37	7.34
$n = 500$	0.45	0.53	0.79	1.79	2.55	8.10

表1：不同 μ_0 和 n下加权最小二乘估计量的样本均值。

以 μ_0标准化后的有限样本偏差在 μ_0较小时似乎较小。此外，正如预期，随着 n 增大，相对偏差通常会减小，而且 μ_0较大时，改善更为显著。

最后，如果 Galton-Watson 模型已完全参数化（例如 INARCH 情形），也可以进行有限样本偏差校正：首先估计 μ，然后像表1中那样通过模拟计算样本偏差。

3.3 一致推断?

Galton-Watson 过程缺乏统一推断这一点已有充分记载（Heyde and Seneta，1974；Wei and Winnicki，1990）。我们也能在平稳情形（$m < 1$）下使用本文提出的新加权最小二乘估计量吗？我们有如下定理：

定理4。 如果 $m < 1$，则有：

- $a)$ $\frac{1}{\ln n} \sum_{t=1}^{n} \frac{1}{t} X_{t-1} \to \lim_{t\to\infty} \mathbb{E}[X_t] = \frac{\mu_0}{1-m_0}$，几乎必然，
- $b)$ $\frac{1}{\ln n} \sum_{t=1}^{n} \frac{1}{t} X_{t-1}^2 \to \lim_{t\to\infty} \mathbb{E}[X_t^2]$，几乎必然，
- $c) \sum_{t=1}^{n} \frac{1}{t} W_t X_t$ 和 $\sum_{t=1}^{n} \frac{1}{t} W_t$ 都依分布收敛于某些随机变量 Z_1、 Z_2。特别地，它们也联合弱收敛于$(Z_1$、 $Z_2)$。

and

$$\begin{bmatrix} (\ln n)(\tilde{m}-m_0) \\ (\ln n)(\tilde{\mu}-\mu_0) \end{bmatrix} \Rightarrow \begin{bmatrix} \lim_{t\to\infty}\mathbb{E}[X_t^2] & \lim_{t\to\infty}\mathbb{E}[X_t] \\ \lim_{t\to\infty}\mathbb{E}[X_t] & 1 \end{bmatrix}^{-1} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix},$$

其中$\lim_{t\to\infty}\mathbb{E}[X_t^2]$是$(X_t)$的平稳分布的二阶矩，且其表达式为：

$$\mathbb{E}[X_t^2] = \left(\frac{\mu_0}{1-m_0}\right)^2 + \frac{\sigma_0^2 \frac{\mu_0}{1-m_0}+b_0}{1-m_0^2}.$$

这个定理显而易见，因此省略证明。

因此，新的加权最小二乘估计量在平稳情形和积分情形下均具有一致性，但与其他估计量（如 OLS 估计量）相比，其在平稳情形下的收敛速度并非最优[2] 在 $m_0 < 1$ 的情况下，OLS 估计量以标准参数收敛速度收敛$\sqrt{n}$。因此，如果研究者得出过程（X_t）是平稳的结论，仍然优先选择 OLS。

3.4 平稳性与单位根检验

第3.3节引出了一个非常自然的问题：在实践中，如何根据数据判断过程（X_t）是平稳的还是积分的？可以采用以下程序，该程序最早由 Kwiatkowski等人（1992年）（KPSS）提出：

i) 进行平稳性检验，例如 KPSS 检验，其原假设为 $m_0 < 1$，备择假设为 $m_0 = 1$。例如，Michel（2019年）曾建议将此检验用于计数值过程。

ii) 使用 m的单位根检验补充平稳性检验，并利用定理 2 或定理 3 的极限理论检验 $\hat{m}$ 或 $\tilde{m}$。在这两种情况下，极限分布都涉及参数μ 和 σ^2，这些参数必须进行估计。新的 $1/t-$加权最小二乘可用于估计 μ，而 σ^2 可根据（Winnicki，1991年，定理 3.5）进行估计。如果单位根检验未被拒绝，但平稳性检验被拒绝，那么该过程很可能是非平稳的。

iii) 如果单位根检验被拒绝，但平稳性检验未被拒绝，那么该过程很可能是平稳的。在这种情况下，应使用 OLS 估计量来估计 m 和 μ，因为在平稳情形下，它的收敛速度远快于加权最小二乘估计量。关于该平稳情形下 σ^2 和 b 的估计量及其性质，另请参见（Winnicki，1991年，命题3.1）。

[2]或者，如果条件分布已完全指定，则使用最大似然（ML）估计量。

4 结论

本文提出了一种新的积分 Galton-Watson 过程截距参数 μ 估计量，无论 $\tau_0 = 2\mu_0/\sigma_0^2$ 是否大于 1，该估计量都具有一致性。本文提出的权函数 $1/t$在相关文献中尚属新颖，而且可能非常有用，因为许多更复杂的模型中也存在类似的不一致性问题。例如，Barreto-Souza 和 Chan（2024）研究了一种 INARCH 模型，其中参数 m_n具有局部单位根性质（即 $m_n = 1 + \frac{\gamma}{n}$），该模型通过允许 γ 可能非零，扩展了严格积分 INARCH 模型。他们的分析假设 μ 先验固定，这可能是为了避免不一致性问题。Barczy 等人（2014）研究了 INAR(2) 模型：

$$X_{-1} = X_0 = 0, \qquad X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + \epsilon_t, \forall t, \tag{12}$$

其中，概率参数 α_1, α_2 均为正，计数序列 (ϵ_t) 在 t 上独立同分布，均值为 μ。(Barczy 等，2014，定理2.1) 表明，在单位根假设$\alpha_1 + \alpha_2 = 1$, 下， μ 的 OLS 估计量不相合。他们的模型可视为（1）的二阶扩展。Michel（2020）推导了非马尔可夫 INGARCH(1,1) 过程趋向 CIR 扩散的标度性质，但未能进一步推导出 μ 的相合估计量等。该模型可视为（1）的无穷阶扩展。这些模型的分析留待未来研究。

参考文献

Barczy, M., Isp´any, M., and Pap, G.(2011). 非稳定 INAR(p) 过程的渐近行为。《随机过程及其应用》, 121(3):583–608.

Barczy, M., Isp´any, M., 和 Pap, G. (2014). 二阶不稳定整数值自回归模型的条件最小二乘估计量的渐近行为。《斯堪的纳维亚统计学杂志》, 41(4):866–892.

Barreto-Souza, W.和Chan, N. H. (2024). 近不稳定整数值 ARCH 过程与单位根检验。《斯堪的纳维亚统计学杂志》, 51(1):402–424.

Gouri´eroux, C. 和 Lu, Y.（2019年）。具有随机强度的负二项自回归过程。《时间序列分析杂志》， 40(2)：225–247。

Heyde, C. C. 和 Seneta, E.（1974）。关于“乘法过程中增长率与移入率的估计理论”的札记。《应用概率杂志》，11(3)：572–577。Jacod, J. 和 Shiryaev, A.（2013）。随机过程极限定理，第288卷。施普林格科学与商业媒体。Jakubowski, A.、M´emin, J. 和 Pag`es, G.（1989）。Skorokhod 空间 1 中随机积分序列的依分布收敛。《概率论及其相关领域》，81(1)：111–137。Kurtz, T. G. 和 Protter, P.（1991）。随机积分和随机微分方程的弱极限定理。《概率年刊》，第1035–1070页。Kutoyants, Y. A.（2013）。遍历扩散过程的统计推断。施普林格科学与商业媒体。Kwiatkowski, D.、Phillips, P. C.、Schmidt, P. 和 Shin, Y.（1992）。检验平稳性原假设与单位根备择假设：我们有多大把握认为经济时间序列具有单位根？《计量经济学杂志》，54(1-3)：159–178。Lu, Y.（2021）。基于稀疏化的计数过程的预测分布。《斯堪的纳维亚统计学杂志》，48(1)：42–67。Michel, J.（2019）。非线性时间序列分析论文集。博士论文，俄亥俄州立大学经济学系。Michel, J.（2020）。非平稳整数值 GARCH (1, 1) 过程的极限分布。《时间序列分析杂志》，41(2)：351–356。Pei, J. 和 Lu, Y.（2025年）。利用非平稳计数时间序列模型预测自然灾害频率。《统计学论文集》，66(72)。Wei, C.（1991年）。含移民临界分枝过程的收敛速度。《中华统计学志》，1(1)：175–184。Wei, C. 和 Winnicki, J.（1989年）。含移民分枝过程的一些渐近结果。《随机过程及其应用》，31(2)：261–282。Wei, C. 和 Winnicki, J.（1990年）。含移民分枝过程的均值估计。《统计年刊》，18(4)：1757–1773。

Winnicki, J.（1986年）。含移民分枝过程的统一估计理论。博士论文，马里兰大学帕克分校。

Winnicki, J.（1991年）。含移民分枝过程中的方差估计。《概率论及其相关领域》，88(1)：77–106。

定理 1 的证明

对于 a）。我们指出：

- 由于 $\mathbb{E}[X_t] = t\mu_0$，我们得到，对于任意 n 以及任意给定的 $\eta > 0$:

$$\mathbb{E}[\frac{1}{n}\sum_{t=1}^{\lfloor n\eta \rfloor}\frac{1}{t}X_{t-1}] \leq \frac{\lfloor n\eta \rfloor + 1}{n}\mu_0 \leq 2\eta\mu_0$$

因此，随着 η 递减至零，$\frac{1}{n}\sum_{t=1}^{\lfloor n\eta \rfloor}\frac{1}{t}X_{t-1}$ 关于 n 一致依概率收敛至零

- 根据连续映射定理，对于固定的 $\eta > 0$，我们有 $\frac{1}{n}\sum_{t=\lfloor n\eta \rfloor}^{n}\frac{1}{t}X_{t-1} \Rightarrow \int_{\eta}^{1}\frac{Y_s}{s}\mathrm{d}s$

- 最后，随着 η 趋于 0，$\int_{0}^{\eta}\frac{Y_s}{s}\mathrm{d}s \to 0$ 几乎必然成立。

因此我们得出结论：

$$\frac{1}{n}\sum_{t=1}^{n}\frac{1}{t}X_{t-1} \Rightarrow \int_{0}^{1}\frac{Y_s}{s}\mathrm{d}s$$

当 n 趋于无穷时。因此 a)成立。

对于 c)，我们只需证明：对任意给定的 $\epsilon > 0$，都有

$$\int_{\epsilon}^{1}\frac{X_{\lfloor ns \rfloor}}{\lfloor ns \rfloor}dM_{n,s} \Rightarrow \int_{\epsilon}^{1}\frac{Y_s}{s}\mathrm{d}M_s \tag{13}$$

根据（Jakubowski 等人，1989年，定理 2.6）[另见 Kurtz 和 Protter（1991年）以及（Jacod 和 Shiryaev，2013年，第 VI 章第 6 节）]，，我们需要检验鞅序列（$M_{n,\ s},\ \epsilon < s < 1$）满足 UT 条件。由于（$W_t$）是（$M_{n,\ s}$）鞅序列，根据（Jakubowski 等人，1989年，命题 3.2），只需检验：

$$\sup_n \mathbb{E}[\max_{1\leq t\leq n}\frac{|W_t|}{n}] < \infty.$$

我们注意到 $\max_{1\leq t\leq n}\frac{|W_t|}{n} \leq \max_{1\leq t\leq n}\frac{|X_t|}{n} + \frac{\mu_0}{n}$。因此，只需证明 $\sup_n \mathbb{E}[\max_{1\leq t\leq n}\frac{|X_t|}{n}] < \infty$. 由于（$X_t$）是次鞅，根据 Doob 不等式，我们有：

$$\mathbb{E}[(\max_{1\leq t\leq n}|X_t|)^q] \leq (\frac{q}{q-1})^q\mathbb{E}[X_n^q]$$

对于任意指数 $q > 1$。但根据过程 $(\frac{X_{\lfloor ns \rfloor}}{n})$ 弱收敛到 (Y_s)，我们有：$\mathbb{E}[\frac{X_n^q}{n}] = O(n^q)$，其中 q 介于 1 和 $1+\delta$ 之间。因此，随着 $\mathbb{E}[\max_{1\leq t\leq n}\frac{|W_t|}{n}]$、$n$ 变化，序列 $\mathbb{E}[\max$ 有界。因此，满足 UT 条件，且 (13) 成立。

对于 d)，我们将鞅中心极限定理应用于三角阵 $W_{t,\ n} = \frac{1}{\sqrt{\ln n}} W_t \frac{1}{t}$。我们验证：

- i) $V_n := \sum_{t=1}^{n} \mathbb{V}ar[W_{t,n} \mid X_{t-1}, X_{t-2}, \cdots] \to \mu_0$ 随着 n 趋于无穷大而依概率收敛。
- ii)（林德伯格条件）：$\sum_{t=1}^{n} \mathbb{E}[W_{t,n}^2 \mathbb{1}_{X_{t,n} > \epsilon\sqrt{V_n}} \mid X_{t-1}, X_{t-2}, \cdots] \to 0$ 对于任意给定的 $\epsilon > 0$，依概率成立。

事实上，若这两个条件均满足，则 $\frac{1}{\mu_0} \sum_{t=1}^{n} W_{t,\ n}$ 弱收敛于一个正态随机变量。

对于 i)，我们有：

$$\begin{aligned}
\sum_{t=1}^{n} \mathbb{V}ar[X_{t,n} \mid X_{t-1}, X_{t-2}, \cdots] &= \frac{1}{\ln n} \sum_{t=1}^{n} \frac{\sigma_0^2 X_{t-1} + b_0}{t^2} \\
&= \frac{1}{\ln n} \sum_{t=1}^{n} \frac{\sigma_0^2 \mathbb{E}[X_{t-1}] + \sigma_0^2 (X_{t-1} - \mathbb{E}[X_{t-1}]) + b_0}{t^2} \\
&= \frac{1}{\ln n} b_0 \sum_{t=1}^{n} \frac{1}{t^2} + \frac{1}{\ln n} \mu_0 \sum_{t=1}^{n} \frac{1}{t} + \frac{1}{\ln n} \sigma_0^2 \sum_{t=1}^{n} \frac{X_{t-1} - (t-1)\mu_0}{t^2},
\end{aligned}$$

其中用到了 $\mathbb{E}[X_{t-1}] = \mu_0(t-1)$。随着 n 趋于无穷大，第一项和第二项分别收敛于 0 和 μ_0。对于第二项，我们可以证明它依概率收敛于零，因此也依概率收敛。因此 $\sum_{t=1}^{n} \mathbb{V}ar[W_{t,n} \mid X_{t-1}, X_{t-2}, \cdots] \to b_0$ 依概率收敛。

对于 ii)，由马尔可夫不等式可得，

$$\begin{aligned}
\sum_{t=1}^{n} \mathbb{E}[|W_{t,n}|^{2+\delta} \mathbb{1}_{W_{t,n} > \epsilon\sqrt{V_n}} \mid X_{t-1}, X_{t-2}, \cdots] &= \sum_{t=1}^{n} \mathbb{E}[|W_{t,n}|^{2+\delta} \mathbb{1}_{|W_{t,n}|^{\delta} > \epsilon^{\delta} {V_n}^{\delta/2}} \mid X_{t-1}, X_{t-2}, \cdots] \\
&\leq \frac{1}{\ln n} \sum_{t=1}^{n} \frac{1}{t^2} \frac{\mathbb{E}[|W_t|^{2+\delta}]}{\epsilon^{\delta} (\ln n)^{\delta/2} V_n^{\delta/2} t^{\delta}}.
\end{aligned}$$

因此，只需证明 $\sum_{t=1}^{n} \mathbb{E}[|W_t|^{2+\delta}]/t^{2+\delta}$ 有界，即可使上述不等式右侧（RHS）在 n 趋于无穷大时趋于零。为此，我们来证明：

$$\mathbb{E}[|W_t|^{2+\delta}] = O(t^{1+\delta/2}). \tag{14}$$

我们记作：

$$W_t = \underbrace{\sum_{i=1}^{X_{t-1}} (Z_{i,t} - 1)}_{:=A_t} + \underbrace{\epsilon_t - \mu_0}_{:=B_t}.$$

By 根据詹森不等式，我们有： $|W_t|^{2+\delta} \le 2^{1+\delta}(A_t^{2+\delta} + B_t^{2+\delta})$。通过取期望，我们得到：

$$\mathbb{E}[|W_t|^{2+\delta}] \le 2^{1+\delta}\mathbb{E}[A_t^{2+\delta}] + 2^{1+\delta}\mathbb{E}[B_t^{2+\delta}].$$

右侧的第二项是一个常数。对于第一项，我们使用罗森塔尔不等式，推得存在一个正常数 C_1，使得：

$$\mathbb{E}[|A_t|^{2+\delta}\,|X_{t-1}] \;\le\; C_1\left[X_{t-1}\,\mathbb{E}\,|Z-1|^{2+\delta} \;+\; (X_{t-1}\,\mathrm{Var}(Z))^{(2+\delta)/2}\right].$$

因此，取边际期望，并利用以下事实：对于任意 $x>0,\ \ x^{(2+\delta)/2}+1 \ge x$，我们得到：

$$\mathbb{E}[|A_t|^{2+\delta}] \le C_2\mathbb{E}[X_{t-1}^{(2+\delta)/2}] + C_3,$$

其中 C_2、C_3 是正常数。然后，根据 $X_{\lfloor ns\rfloor}$ 的弱收敛，我们可以证明：$\mathbb{E}[X_{t-1}^{1+\delta/2}] = O(t^{1+\delta/2})$。因此，（14）成立。结论得证。

B 定理 2 的证明

我们将 F_n 和 d_n 定义为：

$$F_n = \begin{bmatrix} \sum_{t=1}^n \frac{1}{t}X_{t-1}^2 & \sum_{t=1}^n \frac{1}{t}X_{t-1} \\ \sum_{t=1}^n \frac{1}{t}X_{t-1} & \sum_{t=1}^n \frac{1}{t} \end{bmatrix}, \qquad d_n = \begin{bmatrix} \sum_{t=1}^n X_{t-1}W_t\frac{1}{t} \\ \sum_{t=1}^n W_t\frac{1}{t} \end{bmatrix},$$

然后对它们进行重新缩放，并通过以下方式定义 F_n 和 d_n：

$$\tilde{F}_n = \begin{bmatrix} n^{-1} & 0 \\ 0 & 1/\sqrt{\ln n} \end{bmatrix} F_n \begin{bmatrix} n^{-1} & 0 \\ 0 & 1/\sqrt{\ln n} \end{bmatrix} = \begin{bmatrix} n^{-2}\sum_{t=1}^n X_{k-1}^2\frac{1}{t} & \frac{1}{n\sqrt{\ln n}}\sum_{t=1}^n X_{t-1}\frac{1}{t} \\ \frac{1}{n\sqrt{\ln n}}\sum_{t=1}^n X_{t-1}\frac{1}{t} & \frac{1}{\ln n}\sum_{t=1}^n \frac{1}{t} \end{bmatrix},$$

and

$$\tilde{d}_n = \begin{bmatrix} n^{-1} & 0 \\ 0 & 1/\sqrt{\ln n} \end{bmatrix} d_n = \begin{bmatrix} n^{-1}\sum_{t=1}^n X_{t-1}W_t\frac{1}{t} \\ \frac{1}{\sqrt{\ln n}}\sum_{t=1}^n W_t\frac{1}{t} \end{bmatrix}.$$

于是根据定理 1，$(F_n,\ d_n)$ 的联合分布弱收敛于 $(F,\ d)$ 的分布，其形式如下：

$$\tilde{F} = \begin{bmatrix} \int_0^1 \frac{Y_s^2}{s}\mathrm{d}s & 0 \\ 0 & 1 \end{bmatrix}, \qquad \tilde{d} = \begin{bmatrix} \int_0^1 \sigma_0\frac{Y_s^{3/2}}{s}\mathrm{d}B_s \\ Z \end{bmatrix}.$$

因此可得:

$$F_n^{-1}d_n = \begin{bmatrix} n^{-1} & 0 \\ 0 & 1/\sqrt{\ln n} \end{bmatrix} \tilde{F}_n^{-1} \begin{bmatrix} n^{-1} & 0 \\ 0 & 1/\sqrt{\ln n} \end{bmatrix} \begin{bmatrix} n & 0 \\ 0 & \sqrt{\ln n} \end{bmatrix} \tilde{d}_n,$$

or

$$\begin{bmatrix} n & 0 \\ 0 & \sqrt{\ln n} \end{bmatrix} F_n^{-1}d_n = \tilde{F}_n^{-1}\tilde{d}_n \Rightarrow \tilde{F}^{-1}\tilde{d}$$

根据连续映射定理。特别地，$\sqrt{\ln n}(\tilde{\mu} - \mu)$ 收敛于正态分布。

C 定理 3 的证明

OLS 估计量满足:

$$\begin{bmatrix} n(\hat{m}-1) \\ \hat{\mu} - \mu_0 \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^n X_{t-1}^2 & \sum_{t=1}^n X_{t-1} \\ \sum_{t=1}^n X_{t-1} & n \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^n X_{t-1}W_t \\ \sum_{t=1}^n W_t \end{bmatrix}. \tag{15}$$

接下来只需指出以下弱收敛:

- $\frac{1}{n^2}\sum_{t=1}^n X_{t-1} \Rightarrow \int_0^1 Y_s \mathrm{d}s$
- $\frac{1}{n^3}\sum_{t=1}^n X_{t-1}^2 \Rightarrow \int_0^1 Y_s^2 \mathrm{d}s,$
- $\frac{1}{n^2}\sum_{t=1}^n W_t X_t \Rightarrow \sigma_0 \int_0^1 Y_s^{3/2} \mathrm{d}B_s$
- $\frac{1}{n}\sum_{t=1}^n W_t \Rightarrow \sigma_0 \int_0^1 \sqrt{Y_s}\mathrm{d}B_s$。

这里，前两个收敛分别是定理 1 中性质 a) 和 b) 的对应结果，并且是连续映射定理的直接推论。接下来的两个收敛与定理 1 中的 c) 类似。最后，根据连续映射定理，上述弱收敛构成联合收敛。结论得证。