arXiv:2307.01315v2 [math.ST] 27 Mar 2024

A log-linear model for non-stationary time series of counts

ANNE LEUCHT[1,a] (ID) and MICHAEL H. NEUMANN[2,b] (ID)

[1]*Universität Bamberg, Institut für Statistik, Feldkirchenstraße 21, D – 96052 Bamberg, Germany,* [a]*E-mail: anne.leucht@uni-bamberg.de*
[2]*Friedrich-Schiller-Universität Jena, Institut für Mathematik, Ernst-Abbe-Platz 2, D – 07743 Jena, Germany,* [b]*E-mail: michael.neumann@uni-jena.de*

We propose a new model for non-stationary integer-valued time series which is particularly suitable for data with a strong trend. In contrast to popular Poisson-INGARCH models, but in line with classical GARCH models, we propose to pick the conditional distributions from nearly scale invariant families where the mean absolute value and the standard deviation are of the same order of magnitude. As an important prerequisite for applications in statistics, we prove absolute regularity of the count process with exponentially decaying coefficients.

Keywords: Absolute regularity; count process; log-linear model; mixing; nonstationary process

1. Motivation and introduction of the model

We propose a new model for time series of counts which is particularly appropriate for modeling explosive processes. So far, the literature on models for integer-valued time series is dominated by processes where the distribution of the count variables conditioned on the past is taken from the family of Poisson distributions, and where the intensities themselves are random and depend on lagged values of the count and the intensity variables; see e.g. Chapter 4 in Weiss (2018). While most of the results on statistical inference for these models are restricted to stationary time series, most of real life count data exhibit strong seasonal patterns and trends, see Figure 1. These features have to be incorporated into the model as simple detrending and deseasonalization is not feasible due to the discrete structure of the data. With a view towards a possible explosive behavior of processes to be modeled, we believe
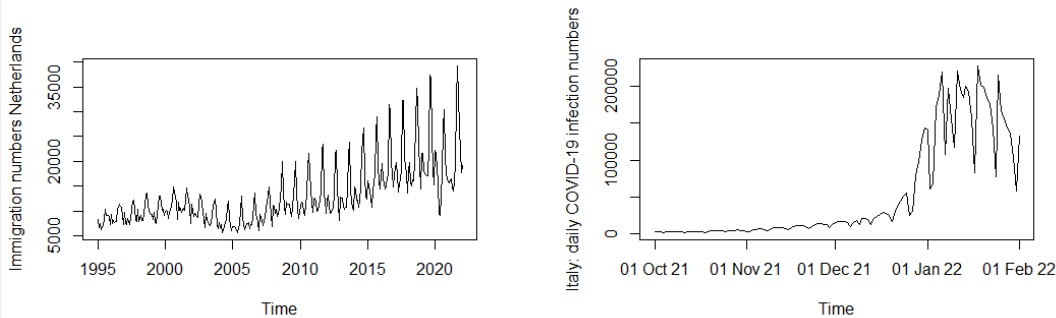

Figure 1. left: Monthly immigration numbers for the Netherlands with increasing trend and strongly increasing seasonality; right: daily COVID-19 infection numbers from Italy with explosive trend.

that it is natural that the expected values of the count variables are of the same order of magnitude as the respective standard deviations. This is in contrast to the family of Poisson distributions where, for $X \sim \text{Poi}(\sigma)$, $EX = \sigma$ but $\sqrt{\text{var}(X)} = \sqrt{\sigma}$, and so $\sqrt{\text{var}(X)}/EX$ tends to zero as $\sigma \to \infty$.

Note that overdispersion in the sense of $\text{var}(X_t) = \eta EX_t$ for some $\eta > 1$ can be incorporated in (log-)linear Poisson regression models by including exogenous regressors, see e.g. Cameron and Trivedi (1986) and Xu *et al.* (2012). This approach is frequently used in various fields; e.g. in demography, health, and biology (as in Renshaw and Haberman (2003), Powell *et al.* (2015), and Aagard, Lyons, and Thogmartin (2018)) and can be adapted to non-linear Poisson autoregressions: If

$$X_t \mid \sigma(\sigma_0, X_0, \ldots, \sigma_{t-1}, X_{t-1}) \sim \text{Poi}(\sigma_t Z_t) \qquad \text{with} \quad \sigma_t = f(\sigma_{t-1}, X_{t-1})$$

where $(Z_t)_t$ is a sequence of i.i.d. random variables with mean μ_z and variance $\sigma_z^2 \in (0, \infty)$ such that Z_t is stochastically independent of $\sigma(\sigma_0, X_0, \ldots, \sigma_{t-1}, X_{t-1})$, then it follows from

$$E[X_t \mid \sigma(\sigma_0, X_0, \ldots, \sigma_{t-1}, X_{t-1})] = f(\sigma_{t-1}, X_{t-1})\mu_z \quad \text{and}$$

$$\text{var}(X_t \mid \sigma(\sigma_0, X_0, \ldots, \sigma_{t-1}, X_{t-1})) = f^2(\sigma_{t-1}, X_{t-1})\sigma_z^2 + f(\sigma_{t-1}, X_{t-1})\mu_z$$

that the degree of conditional dispersion can be changed adapting mean and variance of the exogenous regressors accordingly. Indeed, several well-known count time series models are special cases of these so-called mixed Poisson models. For instance, if the Z_t's are binomial, then the conditional distribution of X_t is a zero-inflated Poisson distribution and Z_t's being Gamma distributed result in a negative binomial distribution, see e.g. Kremer *et al.* (2021) and Doukhan, Leucht, and Neumann (2022).

Here, we propose an alternative approach to assure that conditional expectation and variance are of the same order of magnitude. We pick the conditional distributions from a family of (nearly) scale-invariant distributions. Candidates for such distributions can be found by discretizing scale-invariant continuous distributions. Let Y be a non-negative random variable with a probability density p. Then we can define related integer-valued random variables X_σ by setting

$$X_\sigma = \lfloor \sigma Y \rfloor,$$

i.e. $X_\sigma = k$ if and only if $\sigma Y \in [k, k+1)$. In what follows we denote the distribution of X_σ by P_σ and the corresponding probability mass function by p_σ. For example, if Y is exponentially distributed with rate parameter 1, then σY is exponentially distributed with rate parameter σ and $E[\sigma Y] = \sqrt{\text{var}(\sigma Y)} = \sigma$. We obtain for the corresponding integer-valued random variable X_σ that

$$P_\sigma(\{k\}) = e^{-k/\sigma} - e^{-(k+1)/\sigma} = (1-p)^k p \qquad \forall k \in \mathbb{N}_0,$$

where $p = 1 - e^{-1/\sigma}$. In this case, X_σ has a geometric distribution with success parameter p. Furthermore, $EX_\sigma = (1-p)/p = e^{-1/\sigma}/(1 - e^{-1/\sigma})$ and $\sqrt{\text{var}(X_\sigma)} = \sqrt{(1-p)/p^2} = e^{-1/(2\sigma)}/(1 - e^{-1/\sigma})$ are of the same order of magnitude as $\sigma \to \infty$. (This version of a geometric distribution has support $\mathbb{N}_0$ and describes the number of failures before the first success of independent Bernoulli trials with success parameter p.) Another example can be generated from the family of normal distributions. If $Y \sim N(0, v^2)$, then $|Y|$ has a so-called half-normal distribution, and corresponding integer-valued random variables can be generated by setting $X_\sigma = \lfloor \sigma Y \rfloor$. In both cases, the families $(P_\sigma)_{\sigma>0}$ satisfy the conditions imposed in this paper.

We impose a GARCH-type structure for the count process, i.e.

$$X_t \mid \mathcal{F}_{t-1} \stackrel{d}{=} \lfloor \sigma_t Y \rfloor \sim P_{\sigma_t}, \quad t \in \mathbb{N}, \tag{1.1a}$$

where $\mathcal{F}_s = \sigma(\sigma_0, X_0, Z_0, \ldots, \sigma_s, X_s, Z_s)$ and σ_t is a function of σ_{t-1}, X_{t-1}, and an exogenous covariate Z_{t-1} which may describe e.g. seasonal effects or the effect of a changing environment. To be specific, we will assume that $\sigma_t = f(\sigma_{t-1}, X_{t-1}) \cdot Z_{t-1}$, which can be equivalently rewritten as

$$\ln(\sigma_t) = \ln(f(\sigma_{t-1}, X_{t-1})) + C_{t-1}, \tag{1.1b}$$

where $C_t = \ln(Z_t)$. Note that the initial variable $\sigma_0 \geq 0$ can be chosen arbitrarily as long as some moment conditions are satisfied, see Theorem 2.1 for details.

To work with such processes it is necessary to have some probabilistic properties at our disposal. For classical GARCH processes, mixing properties have been known for a long time; see e.g. Boussama (1998) for linear and Carrasco and Chen (2002) and Francq and Zakoïan (2006) for nonlinear variants. These properties are typically stated for the bivariate process consisting of the observable and the state variables. In sharp contrast, for integer-valued GARCH, the state process $(\sigma_t)_{t \in \mathbb{N}_0}$ is not mixing in general; see Remark 3 in Neumann (2011) for a counterexample. Moreover, mixing properties of classical GARCH processes can be deduced under weak moment conditions on the innovation distribution (see Francq and Zakoïan (2006)), while for Poisson-INGARCH time series all conditional moments given σ_t exist. In contrast, here we require only $E[\ln^+(Y)] < \infty$ with $\ln^+(Y) = \max\{\ln(Y), 0\}$ which means that only logarithmic conditional moments of the X_t may exist.

In this paper we search for conditions that allow us to prove absolute regularity (β-mixing) of the count process $(X_t)_{t \in \mathbb{N}_0}$ ($\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, \ldots\}$). This will be done by a coupling approach described in greater detail in Section 2. To illustrate the usefulness of this result we discuss an application in statistics in Section 3. The proof of the main result is contained in Section 4 and the proof of asymptotic normality of a least squares estimator of a trend parameter is postponed to Section 5. Validity of a bootstrap method to quantify uncertainty of this estimator is proved in Section 6. Proofs of a few technical results are collected in a final Section 7.

2. Absolute regularity of the count process

Let $(\Omega, \mathcal{A}, P)$ be a probability space and $\mathcal{A}_1$, $\mathcal{A}_2$ be two sub-σ-algebras of $\mathcal{A}$. Then the coefficient of absolute regularity is defined as

$$\beta(\mathcal{A}_1, \mathcal{A}_2) = E\big[\sup\{|P(B \mid \mathcal{A}_1) - P(B)| : B \in \mathcal{A}_2\}\big].$$

For a process $\mathbf{X} = (X_t)_{t \in \mathbb{N}_0}$ on $(\Omega, \mathcal{F}, P)$, the coefficients of absolute regularity at the point k are defined as

$$\beta^X(k, n) = \beta\big(\sigma(X_0, X_1, \ldots, X_k), \sigma(X_{k+n}, X_{k+n+1}, \ldots)\big)$$

and the (global) coefficients of absolute regularity as

$$\beta^X(n) = \sup\{\beta^X(k, n) : k \in \mathbb{N}_0\};$$

see e.g. Doukhan (1994). The intended approach to prove absolute regularity is inspired by the fact that one can construct, on a suitable probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ two versions of the process $(X_t)_{t \in \mathbb{N}_0}$, $(\widetilde{X}_t)_{t \in \mathbb{N}_0}$ and $(\widetilde{X}'_t)_{t \in \mathbb{N}_0}$, such that $(\widetilde{X}_0, \ldots, \widetilde{X}_k)$ and $(\widetilde{X}'_0, \ldots, \widetilde{X}'_k)$ are independent and

$$\beta^X(k, n) = \widetilde{P}\left(\widetilde{X}_{k+n+r} \neq \widetilde{X}'_{k+n+r} \text{ for some } r \geq 0\right).$$

Since such an optimal coupling seems to be out of reach in our context we confine ourselves to construct a "reasonably good" coupling. Actually, if $(\widetilde{X}_t)_{t \in \mathbb{N}_0}$ and $(\widetilde{X}'_t)_{t \in \mathbb{N}_0}$ defined on a common probability

space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ are any two versions of $(X_t)_{t\in\mathbb{N}_0}$ such that $(\widetilde{X}_0, \ldots, \widetilde{X}_k)$ and $(\widetilde{X}'_0, \ldots, \widetilde{X}'_k)$ are independent, then

$$
\begin{aligned}
\beta^X(k,n) \leq \widetilde{E}\Big[\sup_{C\in\sigma(C)} \big\{ \big| \widetilde{P}\big((\widetilde{X}_{k+n}, \widetilde{X}_{k+n+1}, \ldots) \in C \mid \widetilde{X}_0, \ldots, \widetilde{X}_k \big) \\
- \widetilde{P}\big((\widetilde{X}'_{k+n}, \widetilde{X}'_{k+n+1}, \ldots) \in C \mid \widetilde{X}'_0, \ldots, \widetilde{X}'_k \big) \big| \big\} \Big] \\
\leq \widetilde{P}\big(\widetilde{X}_{k+n+r} \neq \widetilde{X}'_{k+n+r} \quad \text{for some } r \in \mathbb{N}_0 \big) \\
= \widetilde{P}\big(\widetilde{X}_{k+n} \neq \widetilde{X}'_{k+n} \big) \\
+ \sum_{r=1}^{\infty} \widetilde{P}\big(\widetilde{X}_{k+n+r} \neq \widetilde{X}'_{k+n+r}, \widetilde{X}_{k+n+r-1} = \widetilde{X}'_{k+n+r-1}, \ldots, \widetilde{X}_{k+n} = \widetilde{X}'_{k+n} \big). \quad (2.1)
\end{aligned}
$$

(In the second line of this display, $\sigma(C)$ denotes the σ-algebra generated by the cylinder sets.) We construct on a suitable probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ two versions $\big((\widetilde{\sigma}_t, \widetilde{X}_t, \widetilde{C}_t)\big)_{t\in\mathbb{N}_0}$ and $\big((\widetilde{\sigma}'_t, \widetilde{X}'_t, \widetilde{C}'_t)\big)_{t\in\mathbb{N}_0}$ of the process $\big((\sigma_t, X_t, C_t)\big)_{t\in\mathbb{N}_0}$ such that

$$
\widetilde{P}\big(\widetilde{X}_t = \widetilde{X}'_t \mid \widetilde{\sigma}_t, \widetilde{\sigma}'_t \big) = \sum_{k=0}^{\infty} P_{\widetilde{\sigma}_t}(\{k\}) \wedge P_{\widetilde{\sigma}'_t}(\{k\}) = 1 - d_{TV}\big(P_{\widetilde{\sigma}_t}, P_{\widetilde{\sigma}'_t} \big), \quad (2.2)
$$

where $d_{T_V}(P, Q) = (1/2)\sum_{k=0}^{\infty} |P(k) - Q(k)|$ denotes the total variation distance of two count measures P and Q. Furthermore, if the density p of Y is nonincreasing on $[0, \infty)$ or differentiable everywhere on $(0, \infty)$ and $\int_0^{\infty} x|p'(x)|\, dx < \infty$, then Lemma 7.1 shows that

$$
d_{TV}\big(P_{\sigma}, P_{\sigma'} \big) = O\big(|\ln(\sigma) - \ln(\sigma')| \big).
$$

Hence, it is important to have the evolution of the process $\big(|\ln(\widetilde{\sigma}_t) - \ln(\widetilde{\sigma}'_t)| \big)_{t\in\mathbb{N}_0}$ under control. Since $E\big| \ln(\lfloor \sigma Y \rfloor + 1) - \ln(\lfloor \sigma' Y \rfloor + 1) \big| \leq \gamma\, |\ln(\sigma) - \ln(\sigma')|$ for

$$
\gamma := \int_0^{\infty} \sup\{p(y) \colon y \geq x\}\, dx
$$

(see Lemma 7.3 below) we suppose that the link function f in (1.1b) satisfies the following contractive condition.

$$
\begin{aligned}
\big| \ln\big(f(\sigma, x) \big) - \ln\big(f(\sigma', x') \big) \big| \\
\leq a\, |\ln(\sigma) - \ln(\sigma')| + b\, |\ln(x+1) - \ln(x'+1)| \qquad \forall \sigma, \sigma' > 0, x, x' \in \mathbb{N}_0, \quad (2.3a)
\end{aligned}
$$

where $a, b \geq 0$ and $a + b\gamma < 1$. Note that the latter condition implies validity of the classical contraction condition $a + b < 1$ since it holds $\gamma \geq 1$ by definition. Furthermore, we assume that C_t is independent of $\mathcal{F}_{t-1}$ and X_t, and that

$$
M := \sup_t E|C_t - EC_t| < \infty. \quad (2.3b)
$$

Note that this includes the specification proposed in Fokianos and Tjøstheim (2011),

$$
\ln(\sigma_t) = a\ln(\sigma_{t-1}) + b\ln(X_{t-1}+1) + d \qquad \forall t \in \mathbb{N},
$$

where the count variable X_t given the past had a Poisson distribution with random intensity σ_t, and d was a constant. In that case, for appropriately chosen values of the parameters a and b, there exists a stationary process with such a dynamics. Lemmas 4.1, 4.2, and 7.1 allow us to prove the following main result of this contribution.

Theorem 2.1. *Suppose that* (1.1a)*,* (1.1b)*,* (2.3a)*, and* (2.3b) *are fulfilled, and let* $E|\ln(\sigma_0)| + E\ln^+(Y) < \infty$*. We assume that the density* $p\colon [0,\infty) \to [0,\infty)$ *of* Y *is*

   (i) *monotonously non-increasing*
       or
   (ii) *everywhere differentiable.*

In the former case we set $\gamma = \Gamma = 1$ *and in the latter* $\gamma = \int_0^\infty \sup\{p(y)\colon y \geq x\}\, dx$ *and* $\Gamma = \big(1 + \int_0^\infty x\,|p'(x)|\, dx\big)/2$*. Then the count process* $(X_t)_{t\in\mathbb{N}_0}$ *is absolutely regular (β-mixing), and the corresponding coefficients satisfy*

$$\beta^X(n) \leq (a+b\gamma)^n \, \frac{\Gamma}{1-a} \left\{ 2E|\ln(\sigma_0)| + \frac{2b\,(\|p\|_\infty + E\ln^+(Y)) + 2M}{1-a-b} \right\}.$$

Theorem 2.1 can serve as a basis for various statistical applications. For instance, confidence sets and statistical tests can be developed relying on Rio's (1995) CLT for triangular arrays of nonstationary random variables; see Section 3 for details.

Remark 1. The monotonicity assumption on p is satisfied for instance if Y is exponentially distributed or if it is half-normal. If Y has a chi-square distribution with at least $k = 3$ degrees of freedom, then the corresponding density p_k is unimodal with a mode at $k-2$. However it is differentiable everywhere on $(0,\infty)$ and

$$\int_0^\infty x\,|p_k'(x)|\, dx = \int_0^{k-2} x\, p_k'(x)\, dx \;-\; \int_{k-2}^\infty x\, p_k'(x)\, dx$$

$$= 2(k-2)p_k(k-2) - \int_0^{k-2} p_k(x)\, dx + \int_{k-2}^\infty p_k(x)\, dx < \infty.$$

Hence, our conditions on the distribution of Y are satisfied. Another example is given by the family of Cauchy distributions. Such a distribution has a density $p(x) = \pi\gamma/((x-\mu)^2 + \gamma^2)$ on $\mathbb{R}$, where $\mu \in \mathbb{R}$ is the location and $\gamma > 0$ the scale parameter. If Z follows such a distribution we would call the distribution of $Y = |Z|$ to be half-Cauchy. Such a distribution does not have finite moments of order greater than or equal to one, however, $E\ln^+(Y)$ is finite. Y has a density $p(x) = \pi\gamma/((x-\mu)^2 + \gamma^2) + \pi\gamma/((x+\mu)^2 + \gamma^2)$ which is not non-increasing on $[0,\infty)$ if $|\mu|$ is large, however, it is decreasing on $[|\mu|,\infty)$. Again, our conditions are satisfied by this family of distributions.

Remark 2. The proof of Theorem 2.1 relies heavily on the discretization $\lfloor\sigma Y\rfloor$ to bound (2.1) and (2.2), respectively. Still, these two relations build the basis for proving absolute regularity of other count time series such as Poisson INGARCH processes and variants thereof, see e.g. Doukhan and Neumann (2019), Neumann (2021), or Doukhan, Leucht, and Neumann (2022).

3. An application in statistics

Suppose that we observe the random variables $X_1, \ldots, X_n$ which follow a log-linear model similar to that proposed in Fokianos and Tjøstheim (2011). However, we include an increasing intercept term

which causes a strong trend. To be specific, we assume that

$$\ln(\sigma_t) = a \ln(\sigma_{t-1}) + b \ln(X_{t-1} + 1) + c \ln(t), \tag{3.1}$$

where $a \geq 0$, $b, c > 0$, and $a + b\gamma < 1$. We also assume that

$$X_t \mid \mathcal{F}_{t-1} \stackrel{d}{=} \lfloor \sigma_t Y \rfloor.$$

Although consistent estimation of the parameters a, b and c is a highly relevant issue, it is left for future research as it is mathematically demanding. On the one hand, the σ_t's cannot be observed. Hence, one has to come up with suitable estimators for the volatility process as a preliminary step if $a \neq 0$. Moreover, we expect the convergence rate of least-squares parameter estimators to be comparatively poor in view of the multicollinarity of $(\ln(t))_{t=2,\ldots,n}$ and $(\ln(X_{t-1} + 1))_{t=2,\ldots,n}$. However, under mild regularity conditions we obtain that

$$E \ln(X_t + 1) = \frac{c}{1 - a - b} \ln(t) + O(1); \tag{3.2}$$

see (5.4) and (5.5) in Section 5 below. Hence, the parameter $\theta := c/(1 - a - b)$ characterizes the trend in the data. To estimate it, we may fit the regression model

$$\ln(X_t + 1) = \theta \ln(t) + \varepsilon_t, \qquad t = 1, \ldots, n.$$

The ordinary least squares estimator is given as

$$\widehat{\theta}_n = \frac{\sum_{t=1}^{n} \ln(t) \ln(X_t + 1)}{\sum_{t=1}^{n} \left(\ln(t) \right)^2}$$

and it is centered about the best projection $\bar{\theta} = \bar{\theta}(n)$ which is given as

$$\bar{\theta} = \arg \min_{\theta} \sum_{t=1}^{n} E \left| \ln(X_t + 1) - \theta \ln(t) \right|^2 = \frac{\sum_{t=1}^{n} \ln(t) E \ln(X_t + 1)}{\sum_{t=1}^{n} \left(\ln(t) \right)^2}.$$

Using the mixing property derived in Theorem 2.1 in conjunction with Rio's (1995) central limit theorem for nonstationary and mixing random variables we can prove the following central limit theorem. In particular, it shows a non-standard rate of convergence of our estimator.

Proposition 3.1. *Suppose that the conditions of Theorem 2.1 are satisfied and*

$$\ln(\sigma_t) = a \ln(\sigma_{t-1}) + b \ln(X_{t-1} + 1) + c \ln(t),$$

where $a \geq 0$, $b, c > 0$, and $a + b\gamma < 1$. Moreover, assume that $\sigma_0 \geq 1$ and $\| \ln(\sigma_0) \|_{2+\delta} + \| \ln^+(Y) \|_{2+\delta} < \infty$ for some $\delta > 0$. Then,

$$\sqrt{n} \ln(n) \left(\widehat{\theta}_n - \bar{\theta} \right) \stackrel{d}{\longrightarrow} \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \mathrm{var}\left(\ln(Y) \right) (1 - a)^2 / (1 - (a + b))^2$.

Remark 3. It is known that $\int_0^1 |\ln(y)|^k \, dy = k! \; \forall k \in \mathbb{N}$. Therefore, $\int_0^1 |\ln(y)|^{2+\delta} p(y) \, dy < \infty$ if $\|p\|_\infty < \infty$, which implies that $E \left[|\ln(Y)|^{2+\delta} \right] \leq \|p\|_\infty \int_0^1 |\ln(y)|^{2+\delta} \, dy + E \left[|\ln^+(Y)|^{2+\delta} \right] < \infty$. This means in particular that the variance of $\ln(Y)$ is finite under the above conditions.

To illustrate the dynamics of these models, we revisit the COVID-19 data from Italy. We estimate $\widehat{\theta}_n = 2.48$ and visualize the curve $t \mapsto t^{\widehat{\theta}_n}$, see Figure 2. Moreover, we investigate the finite sample
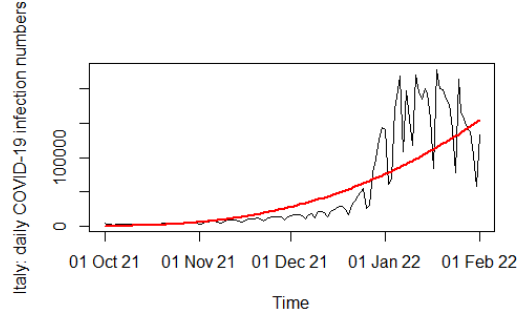

Figure 2. COVID-19 infection numbers with estimated trend curve.

behaviour of our estimator in synthetic samples of sizes $n = 200$ and $n = 500$ with parameters chosen as $a = b = 0.1$ and $c = 2$ which give $\theta = 2.5$, see Figure 3. The distribution of Y is chosen to be either half-normal (left) or exponential (right), both with expectation 1. The box plots are based on 1000 iterations and the red line marks the target value $\bar{\theta}$. We obtained $\bar{\theta}$ by simulation with 20000 Monte Carlo loops as it depends on the unknown distribution of the X_t's as well as the sample size. The performance of the proposed estimator is very convincing in the sense that the median is very close to the target value also for smaller sample size and that the variability becomes smaller as sample sizes increase.
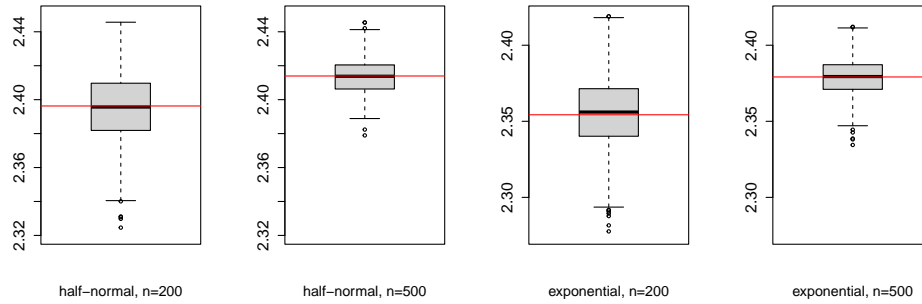

Figure 3. Boxplots for sample sizes $n = 200$ and $n = 500$: left: $Y \sim$ half-normal with $EY = 1$, right: $Y \sim Exp(1)$.

It is possible to construct an asymptotic confidence interval for the parameter $\bar{\theta}$. To this end, we have to approximate the distribution of

$$T_n := \sqrt{n} \ln(n) \left(\widehat{\theta}_n - \bar{\theta} \right) = \sum_{t=1}^{n} w_t \left[\ln(X_t + 1) - E \ln(X_t + 1) \right],$$

where $w_t = \sqrt{n}\,\ln(n)\,\ln(t)/\sum_{s=1}^{n}(\ln(s))^2$. This can be conveniently achieved using the so-called dependent wild bootstrap which was introduced by Shao (2010) for smooth functions of the mean. In our context, we approximate the distribution of T_n by the conditional distribution of

$$T_n^* := \sum_{t=1}^{n} w_t \left[\ln(X_t + 1) - \widehat{m}_n(t) \right] W_t^*,$$

where $\widehat{m}_n(t)$ is an appropriate estimator of $m(t) = E\ln(X_t + 1)$, and $(W_t^*)_{t=1,\ldots,n}$ is a triangular array of auxiliary random variables $W_t^* = W_{t,n}^*$ which are independent of the original sample $X_1, \ldots, X_n$, satisfy $EW_t^* = 0$ and $\mathrm{cov}(W_s^*, W_t^*) = \rho(|s - t|/l_n)$, where $\rho(u) \to_{u\to 0} 1$ and $\sum_{r=1}^{n-1} |\rho(|r|/l_n)| = O(l_n)$. The parameter l_n plays a similar role as the block length in blockwise bootstrap methods and it has to be chosen such that $l_n \to_{n\to\infty} \infty$ and $l_n = o(n)$. To simplify matters we suppose that $W_1^*, \ldots, W_n^*$ are jointly Gaussian. It turns out that the function m is monotonously increasing (see Lemma 6.1 below) and that $m(t) = O(\ln(t))$. Hence $m(t)$ can be well estimated by a nearest neighbor estimator

$$\widehat{m}_n(t) = \sum_{s\,:\,|s-t|\leq N_n} \ln(X_s + 1)/\#\{s \in \{1, \ldots, n\}:\ |s - t| \leq N_n\},$$

where $N_n \to \infty$ as $n \to \infty$. It follows from (5.7) and (5.8) that

$$\mathrm{var}\left(\widehat{m}_n(t)\right) = O\left(1/N_n\right). \tag{3.3}$$

Furthermore, since $\left|E\widehat{m}_n(t) - m(t)\right| \leq E\ln(X_{(t+N_n)\wedge n} + 1) - E\ln(X_{(t-N_n)\vee 1} + 1)$ we obtain that

$$\sum_{t=1}^{n} \left|E\widehat{m}_n(t) - m(t)\right|^2 = O\left(N_n\,(\ln(n))^2\right). \tag{3.4}$$

Remark 4. A simple way to construct the random variables $W_1^*, \ldots, W_n^*$ is to take first an Ornstein-Uhlenbeck process $(U_t)_{t\geq 0}$, i.e. a Gaussian process with continuous sample paths, $EU_t = 0$ and $\mathrm{cov}(U_s, U_t) = \exp(-|s - t|)$, and to define $W_t^* = U_{t/l_n}$, $t = 1, \ldots, n$. Then the practical implementation becomes easy since a discrete sample of an Ornstein-Uhlenbeck process forms an AR(1) process, i.e.

$$W_t^* = e^{-1/l_n} W_{t-1}^* + \sqrt{1 - e^{-2/l_n}}\ \varepsilon_t^*,$$

where $W_1^*, \varepsilon_2^*, \ldots, \varepsilon_n^*$ are independent standard normal variables.

Proposition 3.2. *Suppose that the conditions of Proposition 3.1 are satisfied. Furthermore, we assume that $\sigma_0 = 1$, $\|\ln^+(Y)\|_{4+\delta} < \infty$ for some $\delta > 0$, and that $l_n \xrightarrow[n\to\infty]{} \infty$, $l_n/n \xrightarrow[n\to\infty]{} 0$, and $l_n/N_n +$ $l_n N_n((\ln(n))^2/n \xrightarrow[n\to\infty]{} 0$. Then*

(i) $\sup_{x\in\mathbb{R}} \left|P\left(T_n \leq x\right) - P\left(T_n^* \leq x \mid X_1, \ldots, X_n\right)\right| \xrightarrow{P} 0.$

(ii) *Let $u_{1-\alpha/2}^*$ be the $(1 - \alpha/2)$-quantile of $P^{T_n^* | X_1,\ldots,X_n}$. Then*

$$P\left(\bar{\theta} \in \left[\widehat{\theta}_n - \frac{u_{1-\alpha/2}^*}{\sqrt{n}\,\ln(n)}, \widehat{\theta}_n + \frac{u_{1-\alpha/2}^*}{\sqrt{n}\,\ln(n)}\right]\right) \xrightarrow[n\to\infty]{} 1 - \alpha.$$

Table 1. Rejection frequencies

$n = 500$		half-normal		exponential	
l_n	N_n	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
20	55	0.900	0.953	0.860	0.919
20	65	0.940	0.969	0.904	0.952
25	60	0.912	0.953	0.887	0.933
25	70	0.944	0.976	0.920	0.956
30	65	0.922	0.959	0.895	0.940
30	75	0.951	0.982	0.926	0.961

We illustrate the finite sample performance of the proposed bootstrap method by some simulations regarding the coverage of the bootstrap-based $(1-\alpha)$-confidence intervals for $\alpha = 0.05$ and $\alpha = 0.1$. The data are generated in the same way as in the previous simulations. We use 1000 Monte Carlo loops, each consisting of 1000 bootstrap replications while varying the window size N_n of the nearest-neighbor estimator and the tuning parameter l_n of the dependent wild bootstrap, see Table 3. We observe that the performance of our bootstrap method is convincingly stable with respect to the choice of l_n as long as N_n is not chosen too small. Conversely, choosing N_n too large results in a higher variance of the bootstrap estimator which then gives wider confidence intervals with high coverage.

4. Proof of Theorem 2.1

The proof of our main result is based on the following two lemmas.

Lemma 4.1. *Suppose that* (1.1a), (1.1b), (2.3a), *and* (2.3b) *are fulfilled, that the density p of Y is bounded, and let $E|\ln(\sigma_0)| + E\ln^+(Y) < \infty$. Let $\big((\widetilde{\sigma}_t, \widetilde{X}_t, \widetilde{C}_t)\big)_{t \in \mathbb{N}_0}$ and $\big((\widetilde{\sigma}'_t, \widetilde{X}'_t, \widetilde{C}'_t)\big)_{t \in \mathbb{N}_0}$ be independent versions of $\big((\sigma_t, X_t, C_t)\big)_{t \in \mathbb{N}_0}$ which are defined on a suitable probability space $\big(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P}\big)$. Then*

$$\sup_{k \in \mathbb{N}_0} \widetilde{E}\big|\ln(\widetilde{\sigma}_k) - \ln(\widetilde{\sigma}'_k)\big| \leq 2\, E|\ln(\sigma_0)| + \frac{2b\,(\|p\|_\infty + E\ln^+(Y)) + 2M}{1 - a - b}.$$

Proof. It follows from (1.1b) and (2.3a) that

$$\big|\ln(\widetilde{\sigma}_t) - \ln(\widetilde{\sigma}'_t)\big|$$
$$\leq a\,\big|\ln(\widetilde{\sigma}_{t-1}) - \ln(\widetilde{\sigma}'_{t-1})\big| + b\,\big|\ln(\widetilde{X}_{t-1} + 1) - \ln(\widetilde{X}'_{t-1} + 1)\big| + \big|\widetilde{C}_{t-1} - \widetilde{C}'_{t-1}\big|.$$

Note we obtain from Lemma 7.2 that $\widetilde{E}\big(|\ln(\widetilde{X}_{t-1} + 1) - \ln(\widetilde{\sigma}_{t-1} + 1)|\big|\widetilde{\sigma}_{t-1}, \widetilde{\sigma}'_{t-1}\big) = \widetilde{E}\big(|\ln(\widetilde{X}_{t-1} + 1) - \ln(\widetilde{\sigma}_{t-1} + 1)|\big|\widetilde{\sigma}_{t-1}\big) \leq \|p\|_\infty + E\ln^+(Y)$ and, analogously, $\widetilde{E}\big(|\ln(\widetilde{X}'_{t-1} + 1) - \ln(\widetilde{\sigma}'_{t-1} + 1)|\big|\widetilde{\sigma}_{t-1}, \widetilde{\sigma}'_{t-1}\big) \leq \|p\|_\infty + E\ln^+(Y)$. Therefore we obtain, for $t \geq 1$,

$$\widetilde{E}\Big(\big|\ln(\widetilde{\sigma}_t) - \ln(\widetilde{\sigma}'_t)\big|\,\Big|\,\widetilde{\sigma}_{t-1}, \widetilde{\sigma}'_{t-1}\Big)$$

$$\leq a \left| \ln(\widetilde{\sigma}_{t-1}) - \ln(\widetilde{\sigma}'_{t-1}) \right| + b \left| \ln(\widetilde{\sigma}_{t-1} + 1) - \ln(\widetilde{\sigma}'_{t-1} + 1) \right|$$

$$+ b \, \widetilde{E}\left(\left| \ln(\widetilde{X}_{t-1} + 1) - \ln(\widetilde{\sigma}_{t-1} + 1) \right| \Big| \widetilde{\sigma}_{t-1}, \widetilde{\sigma}'_{t-1} \right)$$

$$+ b \, \widetilde{E}\left(\left| \ln(\widetilde{X}'_{t-1} + 1) - \ln(\widetilde{\sigma}'_{t-1} + 1) \right| \Big| \widetilde{\sigma}_{t-1}, \widetilde{\sigma}'_{t-1} \right)$$

$$+ \widetilde{E}\left(\left| \widetilde{C}_{t-1} - EC_{t-1} \right| \Big| \widetilde{\sigma}_{t-1}, \widetilde{\sigma}'_{t-1} \right) + \widetilde{E}\left(\left| \widetilde{C}'_{t-1} - EC_{t-1} \right| \Big| \widetilde{\sigma}_{t-1}, \widetilde{\sigma}'_{t-1} \right)$$

$$\leq (a + b) \left| \ln(\widetilde{\sigma}_{t-1}) - \ln(\widetilde{\sigma}'_{t-1}) \right| + \bar{M},$$

where $\bar{M} = 2b \left(\|p\|_\infty + E \ln^+(Y) \right) + 2M$. Taking expectation on both sides of this inequality we obtain that

$$\widetilde{E}\left| \ln(\widetilde{\sigma}_t) - \ln(\widetilde{\sigma}'_t) \right|$$

$$\leq (a + b) \, \widetilde{E}\left| \ln(\widetilde{\sigma}_{t-1}) - \ln(\widetilde{\sigma}'_{t-1}) \right| + \bar{M}$$

$$\leq (a + b) \left\{ (a + b) \, \widetilde{E}\left| \ln(\widetilde{\sigma}_{t-2}) - \ln(\widetilde{\sigma}'_{t-2}) \right| + \bar{M} \right\} + \bar{M}$$

$$\leq \ldots \leq (a + b)^t \, \widetilde{E}\left| \ln(\widetilde{\sigma}_0) - \ln(\widetilde{\sigma}'_0) \right| + \bar{M} \left\{ 1 + (a + b) + \cdots + (a + b)^{t-1} \right\}$$

$$\leq 2 E |\ln(\sigma_0)| + \frac{\bar{M}}{1 - a - b}. \qquad \qquad \square$$

Lemma 4.2. *Suppose that* (1.1a), (1.1b), (2.3a), *and* (2.3b) *are fulfilled, and let* $E|\ln(\sigma_0)| < \infty$. *Furthermore, we assume that* $p \colon [0, \infty) \to [0, \infty)$ *is continuous and* $\gamma := \int_0^\infty \sup\{p(y) \colon y \geq x\} \, dx < \infty$. *Then there exist versions* $\left((\widetilde{\sigma}_t, \widetilde{X}_t, \widetilde{C}_t) \right)_{t \in \mathbb{N}_0}$ *and* $\left((\widetilde{\sigma}'_t, \widetilde{X}'_t, \widetilde{C}'_t) \right)_{t \in \mathbb{N}_0}$ *of the process* $\left((\sigma_t, X_t, C_t) \right)_{t \in \mathbb{N}_0}$ *such that, for all* $k, n \geq 0$,

(i) $\widetilde{E}\left| \ln(\widetilde{\sigma}_{k+n}) - \ln(\widetilde{\sigma}'_{k+n}) \right| \leq (a + b\gamma)^n \, \widetilde{E}\left| \ln(\widetilde{\sigma}_k) - \ln(\widetilde{\sigma}'_k) \right|,$

(ii) $\widetilde{E}\left[\left| \ln(\widetilde{\sigma}_{k+n+r}) - \ln(\widetilde{\sigma}'_{k+n+r}) \right| \mathbb{1}\left(\widetilde{X}_{k+n} = \widetilde{X}'_{k+n}, \ldots, \widetilde{X}_{k+n+r-1} = \widetilde{X}'_{k+n+r-1} \right) \right]$

$$\leq a^r \, (a + b\gamma)^n \, \widetilde{E}\left| \ln(\widetilde{\sigma}_k) - \ln(\widetilde{\sigma}'_k) \right| \qquad \forall r \geq 1,$$

(iii) $\widetilde{P}\left(\widetilde{X}_{k+n+r} \neq \widetilde{X}'_{k+n+r}, \widetilde{X}_{k+n+r-1} = \widetilde{X}'_{k+n+r-1}, \ldots, \widetilde{X}_{k+n} = \widetilde{X}'_{k+n} \right)$

$$\leq \Gamma \, \widetilde{E}\left[\left| \ln(\widetilde{\sigma}_{k+n+r}) - \ln(\widetilde{\sigma}'_{k+n+r}) \right| \mathbb{1}\left(\widetilde{X}_{k+n} = \widetilde{X}'_{k+n}, \ldots, \widetilde{X}_{k+n+r-1} = \widetilde{X}'_{k+n+r-1} \right) \right] \; \forall r \geq 0,$$

where $\Gamma = 1$ *if the density* p *is non-increasing on* $[0, \infty)$ *and* $\Gamma = \left(1 + \int_0^\infty x |p'(x)| \, dx \right)/2$ *if* p *is differentiable everywhere on* $(0, \infty)$ *with* $\int_0^\infty x |p'(x)| \, dx < \infty$.

Proof. Let $t \geq k$. For given $\widetilde{\sigma}_t$ and $\widetilde{\sigma}'_t$, we apply a *maximal coupling* of the respective random variables $\widetilde{X}_t$ and $\widetilde{X}'_t$, i.e. $\widetilde{X}_t$ and $\widetilde{X}'_t$ are defined such that

a) $\widetilde{P}\left(\widetilde{X}_t = k \mid \widetilde{\sigma}_t, \widetilde{\sigma}'_t \right) = P_{\widetilde{\sigma}_t}(\{k\})$ and $\widetilde{P}\left(\widetilde{X}'_t = k \mid \widetilde{\sigma}_t, \widetilde{\sigma}'_t \right) = P_{\widetilde{\sigma}'_t}(\{k\})$ $\forall k \in \mathbb{N}_0$,

b) $\widetilde{P}\left(\widetilde{X}_t = \widetilde{X}'_t \mid \widetilde{\sigma}_t, \widetilde{\sigma}'_t \right) = 1 - d_{TV}\left(P_{\widetilde{\sigma}_t}, P_{\widetilde{\sigma}'_t} \right) = \sum_{k=0}^\infty P_{\widetilde{\sigma}_t}(\{k\}) \wedge P_{\widetilde{\sigma}'_t}(\{k\})$.

(Note that our definition of the total variation norm differs from that in Lindvall (1992) by the factor 2.) Furthermore, and in contrast to the construction used for the proof of Theorem 5.2 in Lindvall (1992, Chapter I), we couple $\widetilde{X}_t$ and $\widetilde{X}'_t$ such that $\widetilde{X}_t \geq \widetilde{X}'_t$ holds with conditional probability 1 if

$\widetilde{\sigma}_t \geq \widetilde{\sigma}_t'$ and, vice versa, $\widetilde{X}_t \leq \widetilde{X}_t'$ holds with conditional probability 1 if $\widetilde{\sigma}_t \leq \widetilde{\sigma}_t'$. (This is possible since $\lfloor \sigma Y \rfloor$ is stochastically greater than $\lfloor \sigma' Y \rfloor$ if $\sigma \geq \sigma'$ and $P^{\lfloor \sigma Y \rfloor}((-\infty, t]) \leq P^{\lfloor \sigma' Y \rfloor}((-\infty, t]) \ \forall t$ implies that $(P^{\lfloor \sigma Y \rfloor} - P^{\lfloor \sigma Y \rfloor} \wedge P^{\lfloor \sigma' Y \rfloor})((-\infty, t]) \leq (P^{\lfloor \sigma' Y \rfloor} - P^{\lfloor \sigma Y \rfloor} \wedge P^{\lfloor \sigma' Y \rfloor})((-\infty, t]) \ \forall t$.) And finally, we choose the exogenous variables such that $\widetilde{C}_t = \widetilde{C}_t'$.

(i)　Let $k < t \leq k + n$. Since $\widetilde{X}_t - \widetilde{X}_t'$ has the same sign as $\widetilde{\sigma}_t - \widetilde{\sigma}_t'$, $\ln(\widetilde{X}_t + 1) - \ln(\widetilde{X}_t' + 1)$ is either non-negative or non-positive with conditional probability 1, and we obtain by Lemma 7.3

$$\widetilde{E}\Big(\big|\ln(\widetilde{X}_t + 1) - \ln(\widetilde{X}_t' + 1)\big| \,\Big|\, \widetilde{\sigma}_t, \widetilde{\sigma}_t'\Big) \leq \gamma \big|\ln(\widetilde{\sigma}_t) - \ln(\widetilde{\sigma}_t')\big|.$$

This yields that

$$\widetilde{E}\Big(\big|\ln(\widetilde{\sigma}_t) - \ln(\widetilde{\sigma}_t')\big| \,\Big|\, \widetilde{\sigma}_{t-1}, \widetilde{\sigma}_{t-1}'\Big) \leq (a + b\gamma) \big|\ln(\widetilde{\sigma}_{t-1}) - \ln(\widetilde{\sigma}_{t-1}')\big|.$$

Taking expectation on both sides of this inequality, and using the resulting inequality n times we obtain

$$\widetilde{E}\big|\ln(\widetilde{\sigma}_{k+n}) - \ln(\widetilde{\sigma}_{k+n}')\big| \leq (a + b\gamma)^n \, \widetilde{E}\big|\ln(\widetilde{\sigma}_k) - \ln(\widetilde{\sigma}_k')\big|.$$

(ii)　Let now $t = k + n + r > k + n$. Then we obtain immediately

$$\big|\ln(\widetilde{\sigma}_t) - \ln(\widetilde{\sigma}_t')\big| \mathbb{1}\big(\widetilde{X}_{k+n} = \widetilde{X}_{k+n}', \dots, \widetilde{X}_{t-1} = \widetilde{X}_{t-1}'\big)$$

$$\leq a \big|\ln(\widetilde{\sigma}_{t-1}) - \ln(\widetilde{\sigma}_{t-1}')\big| \mathbb{1}\big(\widetilde{X}_{k+n} = \widetilde{X}_{k+n}', \dots, \widetilde{X}_{t-2} = \widetilde{X}_{t-2}'\big)$$

$$\leq \cdots \leq a^r \big|\ln(\widetilde{\sigma}_{k+n}) - \ln(\widetilde{\sigma}_{k+n}')\big|,$$

which yields (ii). Statement (iii) follows from Lemma 7.1 since b) gives

$$\widetilde{P}\big(\widetilde{X}_{k+n+r} \neq \widetilde{X}_{k+n+r}', \widetilde{X}_{k+n+r-1} = \widetilde{X}_{k+n+r-1}', \dots, \widetilde{X}_{k+n} = \widetilde{X}_{k+n}'\big)$$

$$= E\big(d_{TV}(\widetilde{P}_{\widetilde{\sigma}_{k+n+r}}, \widetilde{P}_{\widetilde{\sigma}_{k+n+r}'}) \, \mathbb{1}\big(\widetilde{X}_{k+n} = \widetilde{X}_{k+n}', \dots, \widetilde{X}_{k+n+r-1} = \widetilde{X}_{k+n+r-1}'\big)\big). \qquad \square$$

Now we are in a position to prove our main result.

Proof of Theorem 2.1. Let $\big((\widetilde{\sigma}_t, \widetilde{X}_t, \widetilde{C}_t)\big)_{t \in \mathbb{N}_0}$ and $\big((\widetilde{\sigma}_t', \widetilde{X}_t', \widetilde{C}_t')\big)_{t \in \mathbb{N}_0}$ be two versions of the process $\big((\sigma_t, X_t, C_t)\big)_{t \in \mathbb{N}_0}$, where $(\widetilde{\sigma}_0, \widetilde{X}_0, \widetilde{C}_0, \dots, \widetilde{\sigma}_{k-1}, \widetilde{X}_{k-1}, \widetilde{C}_{k-1}, \widetilde{\sigma}_k, \widetilde{X}_k)$ and $(\widetilde{\sigma}_0', \widetilde{X}_0', \widetilde{C}_0', \dots, \widetilde{\sigma}_{k-1}', \widetilde{X}_{k-1}', \widetilde{C}_{k-1}', \widetilde{\sigma}_k', \widetilde{X}_k')$ are independent, and where $\widetilde{C}_k, \widetilde{\sigma}_{k+1}, \widetilde{X}_{k+1}, \widetilde{\sigma}_{k+2}, \widetilde{X}_{k+2}, \dots$ are coupled with their respective counterparts $\widetilde{C}_k', \widetilde{\sigma}_{k+1}', \widetilde{X}_{k+1}', \widetilde{\sigma}_{k+2}', \widetilde{X}_{k+2}', \dots$ as described in the proof of Lemma 4.2. Then we obtain from (2.1) and Lemmas 4.1 and 4.2 that

$$\beta^X(k, n)$$

$$\leq \widetilde{P}\Big(\widetilde{X}_{k+n} \neq \widetilde{X}_{k+n}'\Big)$$

$$+ \sum_{r=1}^{\infty} \widetilde{P}\Big(\widetilde{X}_{k+n+r} \neq \widetilde{X}_{k+n+r}', \widetilde{X}_{k+n+r-1} = \widetilde{X}_{k+n+r-1}', \dots, \widetilde{X}_{k+n} = \widetilde{X}_{k+n}'\Big)$$

$$\leq \Gamma \sum_{r=0}^{\infty} a^r (a + b\gamma)^n \left\{ 2 E\big|\ln(\sigma_0)\big| + \frac{2b \, (\|p\|_\infty + E \ln^+(Y)) + 2M}{1 - a - b} \right\}. \qquad \square$$

5. Proof of Proposition 3.1

The following lemma is a consequence of a CLT in Rio (1995) and it serves as a basis for the proof of Proposition 3.1.

Lemma 5.1. *Suppose that $(X_{t,n})_{t=1,\dots,n}$, $n \in \mathbb{N}$, is a triangular array of centered random variables with $\sup_{t \leq n}(\|X_{t,n}\|_{2+\delta}) = O(n^{-1/2})$ for some $\delta > 0$, where $\|X_{t,n}\|_q := (E|X_{t,n}|^q)^{1/q}$. Further assume that the array is absolutely regular with exponentially decaying mixing coefficients $\beta_{(n)}(k) \leq \rho^k$ for some $\rho < 1$. If additionally $V_{n,n} := \mathrm{var}(\sum_{t=1}^{n} X_{t,n})$ satisfies $\liminf_{n\to\infty} V_{n,n} > 0$, then*

$$\sum_{t=1}^{n} \frac{X_{t,n}}{V_{n,n}^{1/2}} \xrightarrow{d} Z \sim \mathcal{N}(0,1).$$

Proof of Lemma 5.1. We apply Corollary 1 in Rio (1995). To this end, we validate conditions (a) and (b) therein. Assumption (a) reads

$$\limsup_{n\to\infty} \max_{i \leq n} \frac{\mathrm{var}(\sum_{t=1}^{i} X_{t,n})}{V_{n,n}} < \infty.$$

In view of our assumption $\liminf_{n\to\infty} V_{n,n} > 0$, it suffices to show that

$$\limsup_{n\in\mathbb{N}} \sum_{s,t=1}^{n} \left| \mathrm{cov}(X_{s,n}, X_{t,n}) \right| \leq C. \tag{5.1}$$

Recall that the β-mixing coefficients serve as an upper bound for the α-mixing coefficients. By the covariance inequality for α-mixing random variables (see e.g. Doukhan (1994, Thm. 3, Sect. 1.2.2)) we obtain

$$\left| \mathrm{cov}(X_{s,n}, X_{t,n}) \right| \leq 8\,\alpha_{(n)}^{X}(|s-t|)^{(2+\delta)/\delta}\,\|X_{s,n}\|_{2+\delta}\,\|X_{t,n}\|_{2+\delta}$$

which then gives (5.1) under our moment conditions by the exponential decay of the strong mixing coefficients.

It remains to check (b) in Corollary 1 in Rio (1995), that is

$$\sum_{t=1}^{n} \int_0^1 \alpha_{(n)}^{-1}\left(\frac{x}{2}\right) Q_{t,n}^2(x) \min\left\{1, \alpha_{(n)}^{-1}\left(\frac{x}{2}\right) Q_{t,n}(x)\right\} dx \xrightarrow[n\to\infty]{} 0. \tag{5.2}$$

Here, $Q_{t,n}$ denotes the inverse of the survival function $z \mapsto P(|X_{t,n}| > z V_{n,n}^{1/2})$ and $\alpha_{(n)}^{-1}$ denotes the inverse of $z \mapsto \alpha(\lfloor z \rfloor)$. First, note that $\alpha_{(n)}^{-1}(x/2) \leq C_1 - C_2 \ln(x)$ for some $C_1, C_2 \in (0,\infty)$ in view of the exponential decay of the mixing coefficients. Second, we obtain by Markov's inequality

$$P(|X_{t,n}| > z V_{n,n}^{1/2}) \leq \frac{E|X_{t,n}|^{2+\delta}}{z^{2+\delta} V_{n,n}^{(2+\delta)/2}}$$

which leads to

$$Q_{t,n}(x) \leq x^{-1/(2+\delta)} \left(\frac{\|X_{t,n}\|_{2+\delta}}{V_{n,n}^{1/2}} \right).$$

From this, and using the rough estimate $\min\left\{1, \alpha_{(n)}^{-1}\left(\frac{x}{2}\right) Q_{t,n}(x)\right\} \leq \alpha_{(n)}^{-\delta/2}\left(\frac{x}{2}\right) Q_{t,n}^{\delta/2}(x)$, we can deduce that the r.h.s. of (5.2) can be bounded from above by

$$\widetilde{C}\int_0^1 |\ln(x)|^{1+\delta/2} x^{-(2+\delta/2)/(2+\delta)}\, dx \cdot \sum_{t=1}^n \left(\frac{\|X_{t,n}\|_{2+\delta}}{V_{n,n}^{1/2}}\right)^{2+\delta/2}$$

which tends to zero as $\sup_{t\leq n}(\|X_{t,n}\|_{2+\delta}) = O(n^{-1/2})$. $\qquad\square$

Before we turn to the proof of Proposition 3.1 we derive a few useful approximations. Since $\sigma_0 \geq 1$ and $\ln(\sigma_t) \geq a\ln(\sigma_{t-1}) + c\ln(t)$, we see inductively that under the conditions of Proposition 3.1

$$\sigma_t \geq t^c \qquad \forall t \geq 1. \tag{5.3}$$

In what follows we shall get rid of the cumbersome terms $\ln(X_t + 1)$. Note that the vector $(X_0, \ldots, X_n)$ has the same distribution as $(\lfloor \sigma_0 Y_0\rfloor, \ldots, \lfloor \sigma_n Y_n\rfloor)$, where Y_t has the same distribution as Y and is independent of $\sigma_0, \ldots, \sigma_t, Y_0, \ldots, Y_{t-1}$. To simplify the representation we suppose that $X_t = \lfloor \sigma_t Y_t\rfloor$.
Since

$$\ln(\sigma_t Y_t) \leq \ln(X_t + 1) \leq \ln(\sigma_t Y_t + 1) \leq \ln(\sigma_t(Y_t + 1/t^c)),$$

we obtain

$$\ln(X_t + 1) - \ln(\sigma_t Y_t) \leq \ln(\sigma_t(Y_t + 1/t^c)) - \ln(\sigma_t Y_t)$$

and hence, for $\rho \leq 2+\delta$,

$$E\left|\ln(X_t + 1) - \ln(\sigma_t Y_t)\right|^\rho \leq E\left|\ln(\sigma_t(Y_t + 1/t^c)) - \ln(\sigma_t Y_t)\right|^\rho$$
$$= E\left|\ln((Y_t + 1/t^c)) - \ln(Y_t)\right|^\rho \underset{t\to\infty}{\longrightarrow} 0. \tag{5.4}$$

The above representation allows us to represent $\ln(\sigma_t)$ in a suitable form:

$$\ln(\sigma_t) = (a+b)\ln(\sigma_{t-1}) + b\ln(Y_{t-1}) + b\left[\ln(X_{t-1}+1) - \ln(\sigma_{t-1}Y_{t-1})\right] + c\ln(t)$$

$$= (a+b)\left\{(a+b)\ln(\sigma_{t-2}) + b\ln(Y_{t-2}) + b\left[\ln(X_{t-2}+1) - \ln(\sigma_{t-2}Y_{t-2})\right] + c\ln(t-1)\right\}$$

$$+ b\ln(Y_{t-1}) + b\left[\ln(X_{t-1}+1) - \ln(\sigma_{t-1}Y_{t-1})\right] + c\ln(t)$$

$$= \ldots = (a+b)^t \ln(\sigma_0)$$

$$+ b\sum_{s=1}^t (a+b)^{s-1}\ln(Y_{t-s})$$

$$+ b\sum_{s=1}^t (a+b)^{s-1}\left[\ln(X_{t-s}+1) - \ln(\sigma_{t-s}Y_{t-s})\right]$$

$$+ c\left\{\sum_{s=0}^{t-1}(a+b)^s \ln(t-s) - \frac{1}{1-a-b}\ln(t)\right\}$$

$$+ \frac{c}{1-a-b}\ln(t). \tag{5.5}$$

The first four terms on the right-hand side are stochastically bounded, and the fifth one is dominating. To see boundedness of the fourth term, note that

$$\left| \sum_{s=0}^{t-1} (a+b)^s \ln(t-s) - \frac{1}{1-a-b} \ln(t) \right|$$

$$\leq \left| \sum_{s=0}^{t-1} (a+b)^s \left[\ln(t-s) - \ln(t) \right] \right| + \left| \frac{1-(a+b)^t}{1-a-b} \ln(t) - \frac{1}{1-a-b} \ln(t) \right|$$

$$\leq \sum_{s=0}^{t-1} (a+b)^s \, s \quad + \quad \frac{(a+b)^t}{1-a-b} \ln(t).$$

Now we are in a position to approximate the covariance structure of $\ln(X_t + 1)$. Since it follows from (5.4) that $\left\| \ln(X_{t-s}+1) - \ln(\sigma_{t-s} Y_{t-s}) \right\|_2 \xrightarrow[t\to\infty]{} 0$, we obtain by Minkowski's inequality

$$\left\| b \sum_{s=1}^{t} (a+b)^{s-1} \left[\ln(X_{t-s}+1) - \ln(\sigma_{t-s} Y_{t-s}) \right] \right\|_2$$

$$\leq b \sum_{s=1}^{t} (a+b)^{s-1} \left\| \ln(X_{t-s}+1) - \ln(\sigma_{t-s} Y_{t-s}) \right\|_2 \xrightarrow[t\to\infty]{} 0.$$

This means that the limits of the covariances arise from the terms in the fourth from last row in (5.5):

$$\operatorname{cov}\left(\ln(\sigma_t), \ln(\sigma_{t-u}) \right) \xrightarrow[t\to\infty]{} \operatorname{var}(\ln(Y)) \, b^2 \, \frac{(a+b)^u}{1 - (a+b)^2}. \tag{5.6}$$

This implies

$$\operatorname{var}\left(\ln(X_t + 1) \right) \xrightarrow[t\to\infty]{} \lim_{t\to\infty} \operatorname{var}\left(\ln(\sigma_t) + \ln(Y_t) \right)$$

$$= \lim_{t\to\infty} \operatorname{var}\left(\ln(\sigma_t) \right) \quad + \quad \operatorname{var}\left(\ln(Y) \right)$$

$$= \operatorname{var}\left(\ln(Y) \right) \left\{ \frac{b^2}{1 - (a+b)^2} + 1 \right\} \tag{5.7}$$

and, for $u \geq 1$,

$$\operatorname{cov}\left(\ln(X_t + 1), \ln(X_{t-u} + 1) \right) \xrightarrow[t\to\infty]{} \lim_{t\to\infty} \operatorname{cov}\left(\ln(\sigma_t), \ln(\sigma_{t-u}) \right) + \lim_{t\to\infty} \operatorname{cov}\left(\ln(\sigma_t), \ln(Y_{t-u}) \right)$$

$$= \operatorname{var}\left(\ln(Y) \right) \left\{ \frac{b^2 (a+b)^u}{1 - (a+b)^2} + b (a+b)^{u-1} \right\}. \tag{5.8}$$

In the subsequent proof, (5.7) and (5.8) allow us to identify $\lim_{n\to\infty} V_{n,n}$.

Proof of Proposition 3.1. Let $J_{t,n} = \frac{\ln(t)}{\sqrt{n}\,\ln(n)} \left(\ln(X_t + 1) - E\ln(X_t + 1) \right)$. By Lemma 7.4

$$\sum_{t=1}^{n} \ln^2(t) / (n \, \ln^2(n)) \xrightarrow[n\to\infty]{} 1.$$

Therefore, it remains to verify

$$\sum_{t=1}^{n} J_{t,n} \xrightarrow{d} Z \sim \mathcal{N}(0, \sigma^2).$$

To this end, we apply Lemma 5.1 with $X_{t,n} := J_{t,n}$. Note that we obtain

$$\sup_{t \leq n} \|J_{t,n}\|_{2+\delta} \leq O(n^{-1/2}) \sup_{t \leq n,\, n \in \mathbb{N}} \| \ln(X_t + 1) - E \ln(X_t + 1) \|_{2+\delta}$$

$$\leq O(n^{-1/2}) \sup_{t \leq n,\, n \in \mathbb{N}} \Big[\| \ln(X_t + 1) - \ln(\sigma_t Y_t) \|_{2+\delta}$$

$$+ \| \ln(\sigma_t Y_t) - E \ln(\sigma_t Y_t) \|_{2+\delta} + \| E [\ln(\sigma_t Y_t) - \ln(X_t + 1)] \|_{2+\delta} \Big].$$

Uniform boundedness of the first and the last summand can be deduced from (5.4). Uniform boundedness of the middle term follows from Remark 3 together with (5.5) since the latter gives

$$\| \ln(\sigma_t Y_t) - E \ln(\sigma_t Y_t) \|_{2+\delta}$$

$$\leq \| \ln(\sigma_t) - E \ln(\sigma_t) \|_{2+\delta} + \| \ln(Y_t) - E \ln(Y_t) \|_{2+\delta}$$

$$\leq (a + b)^t \| \ln(\sigma_0) - E \ln(\sigma_0) \|_{2+\delta}$$

$$+ \sum_{s=1}^{t} (a + b)^{s-1} b \big\| [\ln(X_{t-s} + 1) - \ln(\sigma_{t-s})] - E [\ln(X_{t-s} + 1) - \ln(\sigma_{t-s})] \big\|_{2+\delta}$$

$$+ \| \ln(Y) - E \ln(Y) \|_{2+\delta}. \tag{5.9}$$

It remains to derive $\liminf_{n \to \infty} V_{n,n}$. It follows from (5.4) and (5.9) that the moments of order $2 + \delta$ of the random variables $\ln(X_t + 1)$ are bounded. Hence, we obtain from the mixing property that $\sup_t |\text{cov}(\ln(X_t + 1), \ln(X_{t+h} + 1))| \leq C \rho^{|h|}$ for some $\rho \in (0, 1)$. Furthermore, Lebesgue's theorem, Cauchy's limit theorem, Lemma 7.4, (5.7) and (5.8) give

$$\lim_{n \to \infty} V_{n,n} = \sum_{h \in \mathbb{Z}} \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\ln(t + h) \ln(t)}{\ln^2(n)} \text{cov}(\ln(X_t + 1), \ln(X_{t+h} + 1)) \mathbb{1}_{1 \leq t+h \leq n}$$

$$= \text{var}\big(\ln(Y) \big) \Big\{ 1 + \sum_{h \in \mathbb{Z}} \frac{b^2 (a + b)^{|h|}}{1 - (a + b)^2} + 2 \sum_{h \in \mathbb{N}} b (a + b)^{h-1} \Big\}$$

$$= \text{var}\big(\ln(Y) \big) \frac{(1 - a)^2}{(1 - (a + b))^2} > 0. \qquad \square$$

6. Proof of Proposition 3.2

Proof of Proposition 3.2.

(i)   It follows from Proposition 3.1 that

$$\sup_{x \in \mathbb{R}} \big| P(T_n \leq x) - \Phi(x/\sigma) \big| \xrightarrow[n \to \infty]{} 0. \tag{6.1}$$

Since $\sigma_n^2 := \text{var}(T_n) \xrightarrow[n \to \infty]{} \sigma^2$ and since T_n^* conditioned on $X_1, \ldots, X_n$ is by construction normally distributed, it only remains to show, for $\sigma_n^{*2} := \text{var}(T_n^* \mid X_1, \ldots, X_n)$, that

$$\left| \sigma_n^{*2} - \sigma_n^2 \right| \xrightarrow{P} 0. \tag{6.2}$$

Then (i) follows from (6.1) and (6.2).

Let $Z_t = w_t \left[\ln(X_t + 1) - E \ln(X_t + 1) \right]$ and $Z_t' = w_t \left[E \ln(X_t + 1) - \widehat{m}_n(t) \right]$. Then $T_n = \sum_{t=1}^n Z_t$ and $T_n^* = \sum_{t=1}^n (Z_t + Z_t') W_t^*$. Note that it follows from (3.3), (3.4) and $w_t = O(n^{-1/2})$ that $\sum_{t=1}^n E\left[(Z_t')^2 \right] = O\left(n^{-1} \left(n/N_n + N_n (\ln(n))^2 \right) \right)$. Since $\sum_{r=-(n-1)}^{n-1} |\rho(r/l_n)| = O(l_n)$, we obtain that

$$E\left[\left(\sum_{t=1}^n Z_t' W_t^* \right)^2 \right] = E\left[E\left(\left(\sum_{t=1}^n Z_t' W_t^* \right)^2 \mid X_1, \ldots, X_n \right) \right]$$

$$= E\left[\sum_{s,t=1}^n Z_s' Z_t' \rho(|s-t|/l_n) \right]$$

$$= \sum_{r=-(n-1)}^{n-1} |\rho(r/l_n)| \sum_{\max\{1,1-r\} \le s \le \min\{n,n-r\}} \left| E\left[Z_s' Z_{s+r}' \right] \right|$$

$$= O\left(\frac{l_n}{n} \left(\frac{n}{N_n} + N_n (\ln(n))^2 \right) \right) = o(1).$$

Next we compare $\text{var}(T_n)$ and $\text{var}(\sum_t Z_t W_t^* \mid X_1, \ldots, X_n)$. We obtain

$$E\left[T_n^2 \right] - E\left(\left(\sum_t Z_t W_t^* \right)^2 \mid X_1, \ldots, X_n \right)$$

$$= \sum_{s,t} E\left[Z_s Z_t \right] - \sum_{s,t} Z_s Z_t \rho(|s-t|/l_n)$$

$$= \sum_{s,t} E\left[Z_s Z_t (1 - \rho(|s-t|/l_n)) \right]$$

$$- \sum_{s,t} (Z_s Z_t - E[Z_s Z_t]) \rho(|s-t|/l_n). \tag{6.3}$$

The last but one term converges to zero by majorized convergence. We conclude from (5.4) and (5.5) that

$$\sup_t \left\{ E | \ln(X_t + 1) - E \ln(X_t + 1) |^{4+\delta} \right\} < \infty.$$

This and the exponential decay of the mixing coefficients $\beta^X(r)$ imply, for $1 \le s \le t \le u \le v \le n$ and $r = \max\{t - s, u - t, v - u\}$, that

$$\left| \text{cum}(Z_s, Z_t, Z_u, Z_v) \right| = O\left(\rho^r n^{-2} \right)$$

for some $\rho < 1$, where

$\text{cum}(Z_s, Z_t, Z_u, Z_v) = E[Z_s Z_t Z_u Z_v] - E[Z_s Z_t]E[Z_u Z_v] - E[Z_s Z_u]E[Z_t Z_v] - E[Z_s Z_v]E[Z_t Z_u]$

denotes the joint cumulant of Z_s, Z_t, Z_u and Z_v. Hence,

$$\sum_{s,t,u,v=1}^{n} \left|\mathrm{cum}(Z_s, Z_t, Z_u, Z_v)\right| \leq 4! \sum_{1 \leq s \leq t \leq u \leq v \leq n} \left|\mathrm{cum}(Z_s, Z_t, Z_u, Z_v)\right| = O(1/n).$$

Furthermore, it follows from (5.7) and (5.8) that $\mathrm{cov}(Z_s, Z_t) = O(n^{-1}(a+b)^{|s-t|})$. Therefore we obtain for the second term on the right-hand side of (6.3) that

$$E\left[\left(\sum_{s,t}(Z_s Z_t - E[Z_s Z_t])\rho(|s-t|/l_n)\right)^2\right]$$

$$= \sum_{s,t,u,v=1}^{n} \rho(|s-t|/l_n)\,\rho(|u-v|/l_n)\,\mathrm{cum}(Z_s, Z_t, Z_u, Z_v)$$

$$+ 2 \sum_{s,t,u,v=1}^{n} \rho(|s-t|/l_n)\,\mathrm{cov}(Z_t, Z_u)\,\rho(|u-v|/l_n)\,\mathrm{cov}(Z_v, Z_s)$$

$$= O\left(\frac{1}{n}\right) + O\left(\frac{l_n}{n}\right).$$

(ii)   Follows from (i). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Lemma 6.1. *Suppose that (1.1a) and (3.1) are fulfilled and that $\sigma_0 = 1$. Then the function $t \mapsto E\ln(X_t + 1)$ is monotonously increasing.*

Proof of Lemma 6.1. We construct two versions $(\widetilde{\sigma}_t)_{t \in \mathbb{N}_0}$ and $(\widetilde{\sigma}'_t)_{t \in \mathbb{N}_0}$ of the intensity process such that $\widetilde{\sigma}_0 = \widetilde{\sigma}'_0 = 1$ and

$$\widetilde{\sigma}_t \leq \widetilde{\sigma}'_{t+1} \qquad \forall t \in \mathbb{N} \tag{6.4}$$

holds with probability 1. This will be achieved by feeding the first process with innovations $Y_1, Y_2, \ldots$ while the second one is fed with $Y_0, Y_1, Y_2, \ldots$, where $(Y_t)_{t \in \mathbb{N}_0}$ is a sequence of independent random variables such that $Y_t \stackrel{d}{=} Y$, i.e. we define for $t \geq 1$

$$\ln(\widetilde{\sigma}_t) = a\ln(\widetilde{\sigma}_{t-1}) + b\ln(\lfloor\widetilde{\sigma}_{t-1}Y_t\rfloor + 1) + c\ln(t)$$

$$\text{and}$$

$$\ln(\widetilde{\sigma}'_t) = a\ln(\widetilde{\sigma}'_{t-1}) + b\ln(\lfloor\widetilde{\sigma}'_{t-1}Y_{t-1}\rfloor + 1) + c\ln(t).$$

Then

$$\ln(\widetilde{\sigma}'_1) = a\ln(\widetilde{\sigma}'_0) + b\ln(\lfloor\widetilde{\sigma}'_0 Y_0\rfloor + 1) + c\ln(1)$$

$$\geq 0 = \ln(\widetilde{\sigma}_0).$$

Suppose now that, for $t \geq 1$, $\ln(\widetilde{\sigma}'_t) \geq \ln(\widetilde{\sigma}_{t-1})$. Then

$$\ln(\widetilde{\sigma}'_{t+1}) = a\ln(\widetilde{\sigma}'_t) + b\ln(\lfloor\widetilde{\sigma}'_t Y_t\rfloor + 1) + c\ln(t+1)$$

$$> a\ln(\widetilde{\sigma}_{t-1}) + b\ln(\lfloor\widetilde{\sigma}_{t-1}Y_t\rfloor + 1) + c\ln(t)$$

$$= \ln(\widetilde{\sigma}_t),$$

which proves that (6.4) holds for all $t \in \mathbb{N}$ and with probability 1. This implies that

$$E \ln(X_t + 1) = E \ln(\lfloor \widetilde{\sigma}_t Y_t \rfloor + 1) < E \ln(\lfloor \widetilde{\sigma}'_{t+1} Y_t \rfloor + 1) = E \ln(X_{t+1} + 1). \qquad \square$$

7. A few auxiliary results

In this section we collect a few technical results which contribute to the proof of our main result.

Lemma 7.1. *Suppose that Y is a non-negative random variable with a density p.*

(i) *If p is monotonously non-increasing on $[0, \infty)$, then*

$$d_{TV}\big(P^{\lfloor \sigma Y \rfloor}, P^{\lfloor \sigma' Y \rfloor}\big) \le \big|\ln(\sigma) - \ln(\sigma')\big| \qquad \forall \sigma, \sigma' > 0.$$

(ii) *If p is everywhere differentiable on $(0, \infty)$ and $\int_0^\infty x|p'(x)|\,dx < \infty$, then*

$$d_{TV}\big(P^{\lfloor \sigma Y \rfloor}, P^{\lfloor \sigma' Y \rfloor}\big) \le \big|\ln(\sigma) - \ln(\sigma')\big| \left\{ 1 + \int_0^\infty x|p'(x)|\,dx \right\}/2. \qquad \forall \sigma, \sigma' > 0.$$

Proof. Note that the total variation distance $d_{TV}(P_\sigma, P_{\sigma'}) = (1/2) \sum_{k=0}^\infty |P_\sigma(\{k\}) - P_{\sigma'}(\{k\})|$ between P_σ and $P_{\sigma'}$ can be bounded from above by

$$
\begin{aligned}
d_{TV}(P_\sigma, P_{\sigma'}) &= 1 - \sum_{k=0}^\infty P_\sigma(\{k\}) \wedge P_{\sigma'}(\{k\}) \\
&\le 1 - \int_0^\infty \frac{1}{\sigma} p(x/\sigma) \wedge \frac{1}{\sigma'} p(x/\sigma')\,dx \\
&= \frac{1}{2} \int_0^\infty \Big| \frac{1}{\sigma} p\Big(\frac{x}{\sigma}\Big) - \frac{1}{\sigma'} p\Big(\frac{x}{\sigma'}\Big) \Big|\,dx,
\end{aligned}
\tag{7.1}
$$

which is just the total variation distance between the distributions of σY and $\sigma' Y$.

Let $\sigma < \sigma'$. If p is non-increasing on $[0, \infty)$, then

$$
\begin{aligned}
1 - \int_0^\infty \frac{1}{\sigma} p(x/\sigma) \wedge \frac{1}{\sigma'} p(x/\sigma')\,dx &\le 1 - \int_0^\infty \frac{1}{\sigma'} p\Big(\frac{x}{\sigma}\Big)\,dx \\
&= 1 - \frac{\sigma}{\sigma'} \le \int_\sigma^{\sigma'} \frac{1}{x}\,dx = \ln(\sigma') - \ln(\sigma).
\end{aligned}
$$

If p is everywhere differentiable on $(0, \infty)$ and $\int_0^\infty x|p'(x)|\,dx < \infty$, then

$$
\begin{aligned}
\frac{1}{2} \int_0^\infty \Big| \frac{1}{\sigma} p\Big(\frac{x}{\sigma}\Big) - \frac{1}{\sigma'} p\Big(\frac{x}{\sigma'}\Big) \Big|\,dx &\le \frac{1}{2} \int_0^\infty \int_\sigma^{\sigma'} \Big| \frac{\partial}{\partial u}\Big(\frac{1}{u} p\Big(\frac{x}{u}\Big)\Big) \Big|\,du\,dx \\
&\le \frac{1}{2} \int_0^\infty \int_\sigma^{\sigma'} \frac{1}{u^2} p\Big(\frac{x}{u}\Big) + \frac{x}{u^3}\Big|p'\Big(\frac{x}{u}\Big)\Big|\,du\,dx \\
&= \frac{1}{2} \int_\sigma^{\sigma'} \frac{1}{u} \left\{ \int_0^\infty \frac{1}{u} p\Big(\frac{x}{u}\Big) + \frac{x}{u^2}\Big|p'\Big(\frac{x}{u}\Big)\Big|\,dx \right\}\,du \\
&= \frac{\ln(\sigma') - \ln(\sigma)}{2} \left\{ 1 + \int_0^\infty x|p'(x)|\,dx \right\},
\end{aligned}
$$

which completes the proof. □

Lemma 7.2. *Suppose that Y is a non-negative random variable with a bounded density p and $E[\ln^+(Y)] < \infty$. Then*

$$E\left|\ln(\lfloor \sigma Y \rfloor + 1) - \ln(\sigma + 1)\right| \leq \|p\|_\infty + E[\ln^+(Y)] \qquad \forall \sigma > 0.$$

Proof. Using

$$\sum_{k=0}^{\lfloor \sigma \rfloor} \left(\ln(\sigma + 1) - \ln(k + 1)\right) = \left(\lfloor \sigma \rfloor + 1\right) \ln(\sigma + 1) - \sum_{k=1}^{\lfloor \sigma \rfloor} \ln(k + 1)$$

$$\leq \left(\lfloor \sigma \rfloor + 1\right) \ln(\sigma + 1) - \int_1^{\lfloor \sigma \rfloor + 1} \ln(y)\, dy$$

$$= \left(\lfloor \sigma \rfloor + 1\right) \ln(\sigma + 1) - \left[x\ln(x) - x\right]_1^{\lfloor \sigma \rfloor + 1}$$

$$= \left(\lfloor \sigma \rfloor + 1\right) \underbrace{\left(\ln(\sigma + 1) - \ln(\lfloor \sigma \rfloor + 1)\right)}_{\leq (\sigma - \lfloor \sigma \rfloor)/(\lfloor \sigma \rfloor + 1)} + \lfloor \sigma \rfloor$$

$$\leq \sigma,$$

we obtain that

$$E\left[|\ln(\lfloor \sigma Y \rfloor + 1) - \ln(\sigma + 1)|\, \mathbb{1}(\lfloor \sigma Y \rfloor \leq \sigma)\right]$$

$$= \sum_{k=0}^{\lfloor \sigma \rfloor} \left(\ln(\sigma + 1) - \ln(k + 1)\right) P\left(\sigma Y \in [k, k+1)\right) \leq \|p\|_\infty.$$

Furthermore, we have that

$$E\left[|\ln(\lfloor \sigma Y \rfloor + 1) - \ln(\sigma + 1)|\, \mathbb{1}(\lfloor \sigma Y \rfloor > \sigma)\right]$$

$$\leq E\left[(\ln(\sigma Y + 1) - \ln(\sigma + 1))\, \mathbb{1}(Y \geq 1)\right]$$

$$\leq E\left[(\ln(\sigma Y) - \ln(\sigma))\, \mathbb{1}(Y \geq 1)\right]$$

$$= E\left[\ln^+(Y)\right],$$

which completes the proof. □

Lemma 7.3. *Suppose that Y is a non-negative random variable with a continuous probability density p satisfying $\gamma := \int_0^\infty \sup\{p(y)\colon y \geq x\}\, dx < \infty$. Then the function $\sigma \mapsto E\ln(\lfloor \sigma Y \rfloor + 1)$ is differentiable and*

$$\frac{d}{d\sigma} E\ln\left(\lfloor \sigma Y \rfloor + 1\right) \leq \frac{\gamma}{\sigma} \qquad \forall \sigma > 0.$$

(If p is monotonously non-increasing, then $\gamma = 1$.)

Proof. First note that

$$E\ln\left(\lfloor \sigma Y \rfloor + 1\right) = \sum_{k=1}^\infty \left(\ln(k+1) - \ln(k)\right) P\left(\lfloor \sigma Y \rfloor \geq k\right) = \sum_{k=1}^\infty \left(\ln(k+1) - \ln(k)\right) P\left(\sigma Y \geq k\right).$$

To prove differentiability we consider corresponding difference quotients. Let $g(x) := \sup\{p(y) \colon y \geq x\}$. For $0 < \epsilon < \sigma$, we have that

$$\frac{P\big((\sigma + \epsilon)Y \geq k\big) - P\big(\sigma Y \geq k\big)}{\epsilon} = \frac{1}{\epsilon} \int_{k/(\sigma+\epsilon)}^{k/\sigma} p(y)\, dy$$

$$\leq \frac{1}{\epsilon} k \Big(\frac{1}{\sigma} - \frac{1}{\sigma + \epsilon}\Big) g\Big(\frac{k}{\sigma + \epsilon}\Big)$$

$$\leq \frac{k}{\sigma^2} g\Big(\frac{k}{2\sigma}\Big) \leq \frac{k}{\sigma} \int_{k-1}^{k} \frac{1}{\sigma} g\Big(\frac{y}{2\sigma}\Big)\, dy =: h_1(k),$$

and, for $-\sigma/2 < \epsilon < 0$, we obtain that

$$\frac{P\big((\sigma + \epsilon)Y \geq k\big) - P\big(\sigma Y \geq k\big)}{\epsilon} = -\frac{1}{\epsilon} \int_{k/\sigma}^{k/(\sigma+\epsilon)} p(y)\, dy$$

$$\leq \frac{1}{-\epsilon} k \Big(\frac{1}{\sigma + \epsilon} - \frac{1}{\sigma}\Big) g\Big(\frac{k}{\sigma}\Big)$$

$$\leq \frac{k}{\sigma^2/2} g\Big(\frac{k}{\sigma}\Big) \leq \frac{k}{\sigma} \int_{k-1}^{k} \frac{1}{\sigma/2} g\Big(\frac{y}{\sigma}\Big)\, dy =: h_2(k).$$

In both cases, the corresponding quantities $h_1(k)$ and $h_2(k)$ do not depend on the value of ϵ, and we obtain that

$$\sum_{k=1}^{\infty} \underbrace{\big(\ln(k+1) - \ln(k)\big)}_{\leq 1/k} h_i(k) \leq \frac{2}{\sigma} \int_0^{\infty} g(y)\, dy < \infty,$$

which allows us to invoke Lebesgue's theorem on dominated convergence. Since p is continuous, we have

$$\lim_{\epsilon \to 0} \frac{P\big((\sigma + \epsilon)Y \geq k\big) - P\big(\sigma Y \geq k\big)}{\epsilon} = \frac{k}{\sigma^2} p\Big(\frac{k}{\sigma}\Big),$$

which implies that

$$\frac{d}{d\sigma} E \ln\big(\lfloor \sigma Y \rfloor + 1\big) = \sum_{k=1}^{\infty} k \underbrace{\big(\ln(k+1) - \ln(k)\big)}_{\leq 1} \frac{1}{\sigma^2} p\Big(\frac{k}{\sigma}\Big)$$

$$\leq \frac{1}{\sigma} \sum_{k=1}^{\infty} \int_{k-1}^{k} \frac{1}{\sigma} g(y/\sigma)\, dx \leq \frac{\gamma}{\sigma}. \qquad \square$$

Lemma 7.4. *It holds, for fixed $h \in \mathbb{Z}$,*

$$\sum_{t \colon 1 \leq t, t+h \leq n} \ln(t+h) \ln(t) = n \big(\ln(n)\big)^2 + O\big(n \ln(n)\big).$$

Proof. Let, w.l.o.g., $h \geq 0$. Then

$$\sum_{t \colon 1 \leq t, t+h \leq n} \ln(t+h) \ln(t) = (n-h) \big(\ln(n)\big)^2 - \sum_{t=1}^{n-h} \big(\ln(n)\big)^2 - \ln(t+h) \ln(t).$$

Since $\ln(s+1) - \ln(s) \leq s^{-1}$ for $s \geq 1$ we can estimate the right-hand side by

$$\sum_{t=1}^{n-h} \big(\ln(n)\big)^2 - \ln(t+h)\ln(t)$$

$$\leq \sum_{t=1}^{n-h} \big(\ln(n)\big)^2 - \big(\ln(t)\big)^2$$

$$= \sum_{t=1}^{n-h} \big(\ln(n)\big)^2 - \big(\ln(n-1)\big)^2 + \big(\ln(n-1)\big)^2 - \cdots - \big(\ln(t+1)\big)^2 + \big(\ln(t+1)\big)^2 - \big(\ln(t)\big)^2$$

$$= \sum_{s=1}^{n-1} \min\{s, n-h\} \left[\big(\ln(s+1)\big)^2 - \big(\ln(s)\big)^2 \right]$$

$$\leq \sum_{s=1}^{n-1} \ln(s+1) + \ln(s)$$

$$\leq 2 \sum_{t=1}^{n} \ln(t) = O\big(n\ln(n)\big),$$

which completes the proof. $\qquad\square$

Acknowledgments

References

AAGARD, K., LYONS, J. E., AND THOGMARTIN, W. E. (2018). Accounting for surveyor effort in large-scale monitoring programs. *J. Fish Wildl. Manag.* **9** 459–466.

BOUSSAMA, F. (1998). Ergodicité, mélange et estimation dans les modèles GARCH. PhD thesis, Université Paris 7.

CAMERON, A. C. AND TRIVEDI, P. K. (1986). Econometric models based on count data: comparisons and applications of some estimators and tests. *J. Appl. Econom.* **1** 29–53.

CARRASCO, M. AND CHEN, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econom. Theory* **18** 17–39.

DOUKHAN, P. (1994). *Mixing: Properties and Examples*. Lecture Notes in Statistics **85**, Springer, New York.

DOUKHAN, P. AND NEUMANN, M. H. (2019). Absolute regularity of semi-contractive GARCH-type processes. *J. Appl. Probab.* **56** 91–115.

DOUKHAN, P., LEUCHT, A., AND NEUMANN, M. H. (2022). Mixing properties of non-stationary IN-GARCH(1,1) processes. *Bernoulli* **28** 663–688.

FOKIANOS, K. AND TJØSTHEIM, D. (2011). Log-linear Poisson autoregression. *J. Multivariate Anal.* **102** 563–578.

FRANCQ, C. AND ZAKOÏAN, J.-M. (2006). Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econom. Theory* **22** 815–834.

KREMER, C., TORNERI, A., BOESMANS, S. *et al.* (2021). Quantifying superspreading for COVID-19 using Poisson mixture distributions. *Sci. Rep.* **11** 14107.

LINDVALL, T. (1992). *Lectures on the Coupling Method*. Wiley, New York.

NEUMANN, M. H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* **17** 1268–1284.

NEUMANN, M. H. (2021). Bootstrap for integer-valued GARCH(p,q) processes. *Stat. Neerl.* **75** 343–363.

POWELL, H., KRALL, J. R., WANG, Y., BELL, M. L., AND PENG, R. D. (2015). Ambient coarse particulate matter and hospital admissions in the medicare cohort air pollution study, 1999-2010. *EHP* 123:11

RENSHAW, A. E. AND HABERMAN, S. (2003). Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales. *J. Roy. Statist. Soc. Ser. C* **52** 119–137.

RIO, E. (1995). About the Lindeberg method for strongly mixing sequences. *ESAIM Probab. Statist.* **1** 33–61.

SHAO, X. (2010). The dependent wild bootstrap. *J. Amer. Statist. Assoc.* **105** 218–235.

WEISS, C. H. (2010). The INARCH(1) model for overdispersed time series of counts. *Comm. Statist. Simulation Comput.* **39** 1269–1291.

WEISS, C. H. (2018). *An Introduction to Discrete-Valued Time Series*. Wiley, Hoboken, NJ.

XU, H.-Y., XIE, M., GOE, T. N., AND FU, X. (2012). A model for integer-valued time series with conditional overdispersion. *Comput. Statist. Data Anal.* **56** 4229–4242.