Statistica Sinica: Supplement

TEST FOR CONDITIONAL VARIANCE OF INTEGER-VALUED TIME SERIES

Yuichi Goto, Kou Fujimori

Waseda University, Shinshu University

Supplementary Material

This is the supplementary material of the article "Test for conditional variance of integer-valued time series". In this supplementary material, we provide all proofs of the main theorems 1-5, additional examples of the goodness-of-fit test, and expressions of higher moments to derive asymptotic variances for concrete examples.

S1 Proofs of main article

This section provides the proofs of Theorems 1-5 in the article "Test for conditional variance of integer-valued time series".

Proof of Theorem 1. From (B2), it follows that

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\tilde{L}_n(\boldsymbol{\theta})-L_n(\boldsymbol{\theta})\right|=\frac{1}{n}\sum_{t=1}^{n}\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\ell(Z_t,\tilde{\lambda}_t(\boldsymbol{\theta}))-\ell(Z_t,\lambda_t(\boldsymbol{\theta}))\right|\to 0$$

as $n\to\infty$. Define an open ball as $B_r(\boldsymbol{\theta}_1):=\{\boldsymbol{\theta}:\|\boldsymbol{\theta}-\boldsymbol{\theta}_1\|_{\ell_1}<1/r\}$. By the ergodic theorem for non-integrable processes (see Francq and Zakoian

(2010, p.181, Problem 7.3)), it can be shown that, for any $\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}$,

$$\begin{aligned}\limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_1)} \tilde{L}_n(\boldsymbol{\theta}) \leq & \limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_1)} L_n(\boldsymbol{\theta}) \\ & + \limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_1)} \left|\tilde{L}_n(\boldsymbol{\theta}) - L_n(\boldsymbol{\theta})\right| \\ \leq & \mathrm{E}\left(\sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_1)} \ell(Z_t,\lambda_t(\boldsymbol{\theta}))\right) \quad \text{a.s.}.\end{aligned}$$

Since $\mathrm{E}\left(\sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_1)} \ell(Z_t,\lambda_t(\boldsymbol{\theta}))\right)$ is a decreasing function with respect to r, by Assumption (B1) and the Beppo-Levi theorem, we observe $\mathrm{E}\left(\sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_1)} \ell(Z_t,\lambda_t(\boldsymbol{\theta}))\right) \to \mathrm{E}\left(\ell(Z_t,\lambda_t(\boldsymbol{\theta}_1))\right)$ as $r \to \infty$. Hence, from Assumptions (B3) and (B4), there exists a neighborhood $B(\boldsymbol{\theta}_1)$ of $\boldsymbol{\theta}_1(\neq \boldsymbol{\theta}_0)$,

$$\begin{aligned}\limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in B(\boldsymbol{\theta}_1)} \tilde{L}_n(\boldsymbol{\theta}) \leq & \mathrm{E}\left(\sup_{\boldsymbol{\theta}\in B(\boldsymbol{\theta}_1)} \ell(Z_t,\lambda_t(\boldsymbol{\theta}_1))\right) \quad \text{a.s.} \\ < & \mathrm{E}\left(\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right) \quad \text{a.s.} \\ = & \lim_{n\to\infty} \tilde{L}_n(\boldsymbol{\theta}_0) \quad \text{a.s.}.\end{aligned}$$

From Assumption (B5), there exists, for any covering set $\{B(\boldsymbol{\theta}_i) : i = 0,\dots,\infty\}$ of $\boldsymbol{\Theta}$ such that an open neighborhood $B(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ and open neighborhoods $B(\boldsymbol{\theta}_i)$ with $\limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in B(\boldsymbol{\theta}_i)} \tilde{L}_n(\boldsymbol{\theta}) < \lim_{n\to\infty} \tilde{L}_n(\boldsymbol{\theta}_0)$ and $\boldsymbol{\theta}_i \in \boldsymbol{\Theta}\backslash B(\boldsymbol{\theta}_0)$ for $i = 1,\dots,\infty$, the finite covering set $\{B(\boldsymbol{\theta}_i) : i = 0,\dots,s\}$. Hence, it holds that, for sufficiently large n,

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \tilde{L}_n(\boldsymbol{\theta}) = \max_{i=0,\dots,s} \sup_{\boldsymbol{\theta}\in B(\boldsymbol{\theta}_i)} \tilde{L}_n(\boldsymbol{\theta})$$

$$= \sup_{\boldsymbol{\theta} \in B(\boldsymbol{\theta}_0)} \tilde{L}_n(\boldsymbol{\theta}) \quad \text{a.s..}$$

Since $\hat{\boldsymbol{\theta}}_n$ belongs to the neighborhood $B(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ almost surely for large n and $B(\boldsymbol{\theta}_0)$ can be taken arbitrary, $\hat{\boldsymbol{\theta}}_n$ converges to $\boldsymbol{\theta}_0$ a.s. as $n \to \infty$. $\square$

Proof of Theorem 2. By Assumption (C7), we know that

$$\begin{aligned}
&\sqrt{n} \left\| \frac{\partial}{\partial \boldsymbol{\theta}} \tilde{L}_n(\boldsymbol{\theta}) - \frac{\partial}{\partial \boldsymbol{\theta}} L_n(\boldsymbol{\theta}) \right\|_{\ell_1} \\
\leq &\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\left\| \ell'(Z_t, \tilde{\lambda}_t(\boldsymbol{\theta})) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\lambda}_t(\boldsymbol{\theta}) - \frac{\partial}{\partial \boldsymbol{\theta}} \lambda_t(\boldsymbol{\theta}) \right) \right\|_{\ell_1} \right. \\
&\left. + \left\| \frac{\partial}{\partial \boldsymbol{\theta}} \lambda_t(\boldsymbol{\theta}) \left(\ell'(Z_t, \tilde{\lambda}_t(\boldsymbol{\theta})) - \ell'(Z_t, \lambda_t(\boldsymbol{\theta})) \right) \right\|_{\ell_1} \right) \\
&\to 0 \quad \text{a.s. as } n \to \infty.
\end{aligned}$$

The definition of $\hat{\boldsymbol{\theta}}_n$, Assumptions (C6) and (C12), and Taylor's expansion yield that

$$\begin{aligned}
0 = &\sqrt{n} \frac{\partial}{\partial \boldsymbol{\theta}} \tilde{L}_n(\hat{\boldsymbol{\theta}}_n) \\
= &\sqrt{n} \frac{\partial}{\partial \boldsymbol{\theta}} L_n(\hat{\boldsymbol{\theta}}_n) + o_p(1) \\
= &\sqrt{n} \frac{\partial}{\partial \boldsymbol{\theta}} L_n(\boldsymbol{\theta}_0) - \sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right) \left(- \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} L_n(\boldsymbol{\theta}^\ddagger) \right) + o_p(1),
\end{aligned} \tag{S1.1}$$

where $\boldsymbol{\theta}_0 \lessgtr \boldsymbol{\theta}^\ddagger \lessgtr \hat{\boldsymbol{\theta}}_n$.

We define $B_r(\boldsymbol{\theta}_0) := \{\boldsymbol{\theta} : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|_{\ell_1} < 1/r\}$, and it holds that $B_r(\boldsymbol{\theta}_0) \subset V(\boldsymbol{\theta}_0)$ for sufficiently large r. By the strong consistency of $\hat{\boldsymbol{\theta}}_n$, we observe that $\boldsymbol{\theta}^\ddagger \in B_r(\boldsymbol{\theta}_0)$ for large n. By Assumption (C8) and the ergodic theorem, we have

$$\begin{aligned}
&\left|\frac{\partial^2}{\partial\theta_i\partial\theta_j}L_n(\boldsymbol{\theta}^{\ddagger}) - \mathrm{E}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right| \\
\leq&\frac{1}{n}\sum_{t=1}^{n}\sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_0)}\left|\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta})) - \mathrm{E}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right| \\
&\to \mathrm{E}\sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_0)}\left|\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta})) - \mathrm{E}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right| \quad \text{a.s. as } n\to\infty.
\end{aligned}$$

From Assumption (C6) and the Beppo-Levi theorem, we can see that

$$\mathrm{E}\sup_{\boldsymbol{\theta}\in B_r(\boldsymbol{\theta}_0)}\left|\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta})) - \mathrm{E}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right| \to 0 \quad \text{as } r\to\infty,$$

which shows that

$$-\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}L_n(\boldsymbol{\theta}^{\ddagger}) \to J \quad \text{a.s. as } n\to\infty. \tag{S1.2}$$

Assumptions (A0), (C9), and (C10), and the martingale central limit theorem give that

$$\begin{aligned}
\sqrt{n}\frac{\partial}{\partial\boldsymbol{\theta}}L_n(\boldsymbol{\theta}_0) =&\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial\boldsymbol{\theta}}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0)) \\
&\Rightarrow N(0,I) \quad \text{as } n\to\infty.
\end{aligned} \tag{S1.3}$$

By Assumption (C11), we can show that J is a non-singular matrix. Consequently, from (S1.1)-(S1.3), we obtain the desired result. □

Proof of Theorem 3. (a) We shall prove that $T_n \Rightarrow N(0,\sigma^2)$ as $n\to\infty$.

The proof will be completed once we show that

(i)

$$\Bigg|\frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\Big\{(Z_t-\tilde\lambda_t(\hat{\boldsymbol\theta}_n))^2-\kappa_0(\tilde\lambda_t(\hat{\boldsymbol\theta}_n))\Big\} - \frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\Big\{(Z_t-\lambda_t(\hat{\boldsymbol\theta}_n))^2-\kappa_0(\lambda_t(\hat{\boldsymbol\theta}_n))\Big\}\Bigg| = o_p(1),$$

(ii)

$$\Bigg|\frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\Big\{(Z_t-\lambda_t(\hat{\boldsymbol\theta}_n))^2-\kappa_0(\lambda_t(\hat{\boldsymbol\theta}_n))\Big\} - \frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\left\{(Z_t-\lambda_t(\boldsymbol\theta_0))^2-\kappa_0(\lambda_t(\boldsymbol\theta_0))\right\}\Bigg| = o_p(1),$$

(iii)

$$\frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\left\{(Z_t-\lambda_t(\boldsymbol\theta_0))^2-\kappa_0(\lambda_t(\boldsymbol\theta_0))\right\} \Rightarrow N\left(0,\mathrm{E}\left((Z_t-\lambda_t(\boldsymbol\theta_0))^2-\kappa_0(\lambda_t(\boldsymbol\theta_0))\right)^2\right) \quad \text{as } n\to\infty, \text{ and}$$

(iv)

$$\hat\sigma^2 = \frac{1}{M_n}\sum_{t=1}^{M_n}\Big\{(Z_t-\tilde\lambda_t(\hat{\boldsymbol\theta}_n))^2-\kappa_0(\tilde\lambda_t(\hat{\boldsymbol\theta}_n))\Big\}^2 \to \sigma^2 = \mathrm{E}\left((Z_t-\lambda_t(\boldsymbol\theta_0))^2-\kappa_0(\lambda_t(\boldsymbol\theta_0))\right)^2 \quad \text{as } n\to\infty.$$

First, we prove (i). By Assumption (A1), it is easy to see that

$$\Bigg|\frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\Big\{(Z_t-\tilde\lambda_t(\hat{\boldsymbol\theta}_n))^2-\tilde\kappa_0(\lambda_t(\hat{\boldsymbol\theta}_n))\Big\} - \frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\Big\{(Z_t-\lambda_t(\hat{\boldsymbol\theta}_n))^2-\kappa_0(\lambda_t(\hat{\boldsymbol\theta}_n))\Big\}\Bigg|$$

$$
\begin{aligned}
\leq & \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left| (Z_t - \tilde{\lambda}_t(\boldsymbol{\theta}))^2 - (Z_t - \lambda_t(\boldsymbol{\theta}))^2 \right| \\
& + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left| \kappa_0(\tilde{\lambda}_t(\boldsymbol{\theta})) - \kappa_0(\lambda_t(\boldsymbol{\theta})) \right| \\
\leq & 2 \left| \frac{V}{\sqrt{n}} \sum_{t=1}^{n} Z_t \rho^t \right| + \left| \frac{V}{\sqrt{n}} \sum_{t=1}^{n} \left\{ V \rho^{2t} + 2\rho^t \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \lambda_t(\boldsymbol{\theta}) \right\} \right| \\
& + \frac{V}{\sqrt{n}} \frac{\rho}{1-\rho},
\end{aligned}
$$

which, by the stationarity of $\{Z_t\}$ and the Markov's inequality, tends to 0 as $n \to \infty$.

From the Taylor's expansion, it holds that

$$
\begin{aligned}
& \left| \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n)) \right\} \right. \\
& \quad \left. - \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right\} \right| \\
= & \sqrt{\frac{M_n}{n}} \left| \sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right) \frac{1}{M_n} \sum_{t=1}^{M_n} \frac{\partial}{\partial \boldsymbol{\theta}} \left\{ (Z_t - \lambda_t(\boldsymbol{\theta}^*))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}^*)) \right\} \right| \\
\leq & \sqrt{\frac{M_n}{n}} \left\| \sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right) \right\|_{\ell^1} \left\| \frac{2}{M_n} \sum_{t=1}^{M_n} (Z_t - \lambda_t(\boldsymbol{\theta}^*)) \frac{\partial}{\partial \boldsymbol{\theta}} \lambda_t(\boldsymbol{\theta}^*) \right\|_{\ell^1} \qquad \text{(S1.4)} \\
& + \sqrt{\frac{M_n}{n}} \left\| \sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right) \right\|_{\ell^1} \left\| \frac{1}{M_n} \sum_{t=1}^{M_n} \frac{\partial}{\partial \boldsymbol{\theta}} \kappa_0(\lambda_t(\boldsymbol{\theta}^*)) \right\|_{\ell^1}, \qquad \text{(S1.5)}
\end{aligned}
$$

where $\boldsymbol{\theta}_0 \lesseqgtr \boldsymbol{\theta}^* \lesseqgtr \hat{\boldsymbol{\theta}}_n$. By Assumption (M_2), the $\sqrt{n}$-consistency of $\hat{\boldsymbol{\theta}}_n$, and the continuous mapping theorem, (S1.4) and (S1.5) converge to 0 as $n \to \infty$. Hence, (ii) is established.

Next, we show (iii). By Assumption (A0) and (M_1), it holds that

$$\begin{aligned}
&\frac{1}{M_n}\sum_{t=1}^{M_n}\mathrm{E}\Bigg(\big\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\big\}^2 \\
&\quad\times \mathbb{I}\Big\{\big|(Z_t-\lambda_t)^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\big|>\sqrt{M_n}\epsilon\Big\}\Bigg) \\
=&\mathrm{E}\Bigg(\big\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\big\}^2 \\
&\quad\times \mathbb{I}\bigg\{\frac{|(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))|^\delta}{(\sqrt{M_n}\epsilon)^\delta}>1\bigg\}\Bigg) \\
\le&\frac{1}{(\sqrt{M_n}\epsilon)^\delta}\mathrm{E}\left|(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\right|^{2+\delta}\to 0 \quad \text{as } n\to\infty,
\end{aligned}$$

and the ergodic theorem yields

$$\begin{aligned}
&\frac{1}{M_n}\sum_{t=1}^{M_n}\mathrm{E}\left(\big\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\big\}^2\,\big|\mathcal{F}_{t-1}\right) \\
&\to \mathrm{E}\left(\big\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\big\}^2\right) \quad \text{as } n\to\infty.
\end{aligned}$$

Since $\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\}$ is a martingale difference sequence, the martingale central limit theorem immediately gives (iii).

Finally, we prove (iv). By the Assumption (A1), a simple algebra gives that

$$\begin{aligned}
&\bigg|\frac{1}{M_n}\sum_{t=1}^{M_n}\Big((Z_t-\tilde\lambda_t(\hat{\boldsymbol{\theta}}_n))^2-\kappa_0(\tilde\lambda_t(\hat{\boldsymbol{\theta}}_n))\Big)^2 \\
&-\frac{1}{M_n}\sum_{t=1}^{M_n}\Big((Z_t-\lambda_t(\hat{\boldsymbol{\theta}}_n))^2-\kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n))\Big)^2\bigg| = o_p(1).
\end{aligned}$$

Next, we see that

$$\left| \frac{1}{M_n} \sum_{t=1}^{M_n} \left((Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n)) \right)^2 - \frac{1}{M_n} \sum_{t=1}^{M_n} \left((Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right)^2 \right| = o_p(1). \tag{S1.6}$$

By the Taylor's expansion, we have

$$\begin{aligned} &\frac{1}{M_n} \sum_{t=1}^{M_n} \left\{ \left((Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n)) \right)^2 - \left((Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right)^2 \right\} \\ \leq & \frac{1}{\sqrt{n}} \left\| \sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right) \right\|_{\ell^1} \left\| \frac{1}{M_n} \sum_{t=1}^{M_n} \frac{\partial}{\partial \boldsymbol{\theta}} \left((Z_t - \lambda_t(\boldsymbol{\theta}^\dagger))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}^\dagger)) \right)^2 \right\|_{\ell^1}, \end{aligned}$$

where $\boldsymbol{\theta}_0 \lesseqqgtr \boldsymbol{\theta}^\dagger \lesseqqgtr \hat{\boldsymbol{\theta}}_n$. From Assumption ($M_2$) and the $\sqrt{n}$-consistency of $\hat{\boldsymbol{\theta}}_n$, (S1.6) is established. By Assumption (B3) and the ergodic theorem, it holds that

$$\begin{aligned} &\frac{1}{M_n} \sum_{t=1}^{M_n} \left((Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right)^2 \\ &\to \mathrm{E} \left((Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right)^2 \quad \text{a.s. as } n \to \infty, \end{aligned}$$

which shows the required result.

(b) As well as proof of (a), we shall prove that $T_n \Rightarrow N(0, \tilde{\sigma}^2)$ as $n \to \infty$.

Define that

$$T_n^0 := \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{ (Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right\}.$$

Under our assumptions, it follows, from the Taylor expansion, that

$$
\begin{aligned}
& T_n - T_n^0 \\
=& \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Bigg[\Big\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - (Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 \Big\} \\
& - \Big\{ \kappa_0(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)) - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \Big\} \Bigg] \\
=& \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Bigg[\Big\{ (Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - (Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 \Big\} \\
& - \Big\{ \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n)) - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \Big\} \Bigg] + o_p(1) \\
=& \frac{1}{n} \sum_{t=1}^{n} \Bigg[\{2(Z_t - \lambda_t(\boldsymbol{\theta}_n^*)) - \kappa_0'(\lambda_t(\boldsymbol{\theta}_n^*))\} \left(-\frac{\partial}{\partial \boldsymbol{\theta}} \lambda_t(\boldsymbol{\theta}_n^*) \right)^{\top} \\
& \times \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \Bigg] + o_p(1) \\
=& \mathrm{E} \left(\{2(Z_t - \lambda_t(\boldsymbol{\theta}_0)) - \kappa_0'(\lambda_t(\boldsymbol{\theta}_0))\} \left(-\frac{\partial}{\partial \boldsymbol{\theta}} \lambda_t(\boldsymbol{\theta}_0) \right)^{\top} \right) \\
& \times J^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial \boldsymbol{\theta}} \ell(Z_t, \lambda_t(\boldsymbol{\theta}_0)) + o_p(1),
\end{aligned}
$$

where $\boldsymbol{\theta}_n^*$ is a point between $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}_n$. Putting that

$$
\begin{aligned}
L :=& \mathrm{E} \left(\{2(Z_t - \lambda_t(\boldsymbol{\theta}_0)) - \kappa_0'(\lambda_t(\boldsymbol{\theta}_0))\} \left(-\frac{\partial}{\partial \boldsymbol{\theta}} \lambda_t(\boldsymbol{\theta}_0) \right) \right) \\
=& \mathrm{E} \left(\kappa_0'(\lambda_t(\boldsymbol{\theta}_0)) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \lambda_t(\boldsymbol{\theta}_0) \right) \right),
\end{aligned}
$$

we have that

$$
T_n - T_n^0 = L^{\top} J^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial \boldsymbol{\theta}} \ell(Z_t, \lambda_t(\boldsymbol{\theta}_0)) + o_p(1).
$$

For the vector-valued sequence $(T_n - T_n^0, T_n^0)^\top$, it holds that

$$\begin{pmatrix} T_n - T_n^0 \\ T_n^0 \end{pmatrix} = \begin{pmatrix} L^\top J^{-1} & O \\ O & 1 \end{pmatrix} \left(\sum_{t=1}^{n} \zeta_t \right) + o_p(1),$$

where

$$\zeta_t = \begin{pmatrix} \frac{1}{\sqrt{n}} \frac{\partial}{\partial \boldsymbol{\theta}} \ell(Z_t, \lambda_t(\boldsymbol{\theta}_0)) \\ \frac{1}{\sqrt{n}} \left\{ (Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right\} \end{pmatrix}.$$

By Assumption 3.1 and the definition of κ_0, we have that $\{\zeta_t\}_{t \in \mathbb{Z}}$ is an $\mathbb{R}^{d+1}$-valued martingale difference. Moreover, we can see that this martingale difference satisfies the Lindeberg's condition from Assumptions 2.1 and 3.1. We obtain the following asymptotic normality under null hypothesis by using the martingale central limit theorem:

$$\sum_{t=1}^{n} \zeta_t \Rightarrow N(\mathbf{0}, C) \quad \text{as } n \to \infty,$$

where

$$C := \begin{pmatrix} I & C_{12} \\ C_{12}^\top & \sigma_2^2 \end{pmatrix}$$

with

$$C_{12} = \mathrm{E}\left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ell(Z_t, \lambda_t(\boldsymbol{\theta}_0)) \right)^\top \left\{ (Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) \right\} \right].$$

Therefore, we have that

$$\begin{pmatrix} T_n - T_n^0 \\ T_n^0 \end{pmatrix} \Rightarrow N\left(\mathbf{0}, \begin{pmatrix} L^\top J^{-1} & O \\ O & 1 \end{pmatrix} C \begin{pmatrix} L^\top J^{-1} & O \\ O & 1 \end{pmatrix}^\top \right).$$

Noting that

$$T_n = (1,1)\begin{pmatrix} T_n - T_n^0 \\ \\ T_n^0 \end{pmatrix},$$

we obtain the conclusion by using continuous mapping theorem.

□

Proof of Theorem 4. Under H_1, we prove that $P(|T_n| > C) \to 1, \quad n \to \infty$. We have that

$$\begin{aligned} \{|T_n| > C\} &= \{T_n - C > 0\} \cup \{T_n + C < 0\} \\ &= \left\{\frac{1}{\sqrt{M_n}}T_n - \frac{1}{\sqrt{M_n}}C > 0\right\} \cup \left\{\frac{1}{\sqrt{M_n}}T_n + \frac{1}{\sqrt{M_n}}C < 0\right\}. \end{aligned}$$

Put

$$A_n := \frac{1}{\sqrt{M_n}}T_n - \frac{1}{\sqrt{M_n}}C, \text{ and } \quad A_n' := \frac{1}{\sqrt{M_n}}T_n + \frac{1}{\sqrt{M_n}}C.$$

Then, we can decompose A_n and A_n' as follows:

$$A_n = (I) + (II) + (III) + (IV), \text{ and } \quad A_n' = (I) + (II) + (III) - (IV),$$

where

$$\begin{aligned} (I) = \quad & \tfrac{1}{M_n}\textstyle\sum_{t=1}^{M_n}\left\{(Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))\right\} \\ & -\tfrac{1}{M_n}\textstyle\sum_{t=1}^{M_n}\left\{(Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n))\right\}, \end{aligned}$$

$$(II) = \quad \tfrac{1}{M_n}\textstyle\sum_{t=1}^{M_n}\left\{(Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n))\right\}$$

$$-\tfrac{1}{M_n}\textstyle\sum_{t=1}^{M_n}\left\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\right\},$$

$$\begin{aligned}(III)= \quad & \tfrac{1}{M_n}\textstyle\sum_{t=1}^{M_n}\left\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right\}\\ & -\tfrac{1}{M_n}\textstyle\sum_{t=1}^{M_n}\left\{\kappa_0(\lambda_t(\boldsymbol{\theta}_0))-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right\},\end{aligned}$$

and

$$(IV)=\frac{1}{\sqrt{M_n}}C.$$

Hereafter, we show that $(I),\ (II)=o_p(1)$, $(III)\to^p \mathrm{E}\left(\kappa_0(\lambda_t(\boldsymbol{\theta}_0))-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right)\neq 0$, and $(IV)=o(1)$, which imply our conclusion.

Clearly, we have that $(IV)=o(1)$. We can easily see that

$$(I),\ (II)=o_p(1),$$

for both cases $M_n=o(n)$ and $M_n=n$, by the similar approach to the proof of Theorem 3.3 (a) and (b) since we assume that

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\tilde{\lambda}_t(\boldsymbol{\theta})-\lambda_t(\boldsymbol{\theta})\right|\leq V\rho^t \quad \text{a.s.} \quad \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\kappa_0(\tilde{\lambda}(\boldsymbol{\theta}))-\kappa_0(\lambda(\boldsymbol{\theta}))|\leq V\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\rho^t \quad a.s.$$

in Assumption 2.1 and we have that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0)=O_p(1)$$

from Theorem 2. By the ergodic theorem, it holds that

$$(III)=\mathrm{E}\left((Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right)+\mathrm{E}\left(\kappa_0(\lambda_t(\boldsymbol{\theta}_0))-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right)+o_p(1)$$

$$
\begin{aligned}
=&\mathrm{E}\left(\mathrm{E}\left((Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa(\lambda_t(\boldsymbol{\theta}_0))|\mathcal{F}_{t-1}\right)\right)\\
&+\mathrm{E}\left[\kappa_0(\lambda_t(\boldsymbol{\theta}_0))-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right]+o_p(1)\\
=&\mathrm{E}\left[\kappa_0(\lambda_t(\boldsymbol{\theta}_0))-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right]+o_p(1).
\end{aligned}
$$

Therefore, we have that

$$
A_n,\ A_n' \to^p A, \quad n\to\infty,
$$

where $A=\mathrm{E}\left(\kappa_0(\lambda_t(\boldsymbol{\theta}_0))-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right)$. Note that the assumption

$$
\mathrm{E}\left(\kappa_0(\lambda_t(\boldsymbol{\theta}_0))-\kappa(\lambda_t(\boldsymbol{\theta}_0))\right)\neq 0
$$

ensures to apply the continuous mapping theorem to deduce that

$$
\mathbb{I}_{\{A_n>0\}} \to^p \mathbb{I}_{\{A>0\}} \quad \text{and} \quad \mathbb{I}_{\{A_n'<0\}} \to^p \mathbb{I}_{\{A<0\}}, \quad n\to\infty.
$$

Using the dominated convergence theorem, we obtain that

$$
\begin{aligned}
\lim_{n\to\infty} P\left(|T_n|>C\right) &= \lim_{n\to\infty}\left\{P\left(A_n>0\right)+P\left(A_n'>0\right)\right\}\\
&= \lim_{n\to\infty}\left\{\mathrm{E}\left(\mathbb{I}_{\{A_n>0\}}\right)+\mathrm{E}\left(\mathbb{I}_{\{A_n'<0\}}\right)\right\}\\
&= \mathrm{E}\left(\mathbb{I}_{\{A>0\}}\right)+\mathrm{E}\left(\mathbb{I}_{\{A<0\}}\right)\\
&= 1,
\end{aligned}
$$

which ends the proof. □

Proof of Theorem 5. Denote $\{P_n\}$ as the sequence of probability distributions for $H_{1,n}$ Under $H_{1,n}$, we have that

$$T_n = (I) + (II) + (III),$$

where

$$\begin{aligned}(I) = \quad & \tfrac{1}{\sqrt{M_n}} \textstyle\sum_{t=1}^{M_n} \left\{(Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))\right\} \\ & -\tfrac{1}{\sqrt{M_n}} \textstyle\sum_{t=1}^{M_n} \left\{(Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n))\right\},\end{aligned}$$

$$\begin{aligned}(II) = \quad & \tfrac{1}{\sqrt{M_n}} \textstyle\sum_{t=1}^{M_n} \left\{(Z_t - \lambda_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\lambda_t(\hat{\boldsymbol{\theta}}_n))\right\} \\ & -\tfrac{1}{\sqrt{M_n}} \textstyle\sum_{t=1}^{M_n} \left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0))\right\}\end{aligned}$$

and

$$\begin{aligned}(III) = \quad & \tfrac{1}{\sqrt{M_n}} \textstyle\sum_{t=1}^{M_n} \left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\} \\ & -\tfrac{1}{\sqrt{M_n}} \textstyle\sum_{t=1}^{M_n} \left\{\kappa_0(\lambda_t(\boldsymbol{\theta}_0)) - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\}\end{aligned}$$

for both cases $M_n = o(n)$ and $M_n = n$. As well as the proof of Theorem 4, we can see that $(I) = o_{p_n}(1)$ and (II) satisfies that

$$(II) = \begin{cases} o_{p_n}(1) & M_n = o(n), \\ L^\top J^{-1} \tfrac{1}{\sqrt{n}} \textstyle\sum_{t=1}^{n} \tfrac{\partial}{\partial \boldsymbol{\theta}} \ell(Z_t, \lambda_t(\boldsymbol{\theta}_0)) + o_{p_n}(1) & M_n = n. \end{cases}$$

For (III), it holds that

$$(III) = \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\} + \frac{1}{M_n} \sum_{t=1}^{M_n} h(\lambda_t(\boldsymbol{\theta}_0)).$$

The ergodic theorem implies that

$$\frac{1}{M_n}\sum_{t=1}^{M_n} h(\lambda_t(\boldsymbol{\theta}_0)) = \mathrm{E}(h(\lambda_t(\boldsymbol{\theta}_0))) + o_{p_n}(1).$$

Therefore, under $H_{1,n}$ and $M_n = o(n)$, we have that

$$T_n = \frac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\} + \mathrm{E}(h(\lambda_t(\boldsymbol{\theta}_0))) + o_{p_n}(1).$$

Noting that

$$\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\}_{1\le t\le n}$$

is a martingale difference array, that

$$\begin{aligned}
&\frac{1}{M_n}\sum_{t=1}^{M_n}\mathrm{E}\left(\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\}^2 |\mathcal{F}_{t-1}\right)\\
&= \frac{1}{M_n}\sum_{t=1}^{M_n}\mathrm{E}\left(\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\}^2 |\mathcal{F}_{t-1}\right)\\
&\quad + \frac{1}{M_n}\sum_{t=1}^{M_n}\mathrm{E}\left(\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\}\frac{1}{\sqrt{n}}h(\lambda_t(\boldsymbol{\theta}_0))|\mathcal{F}_{t-1}\right)\\
&\quad + \frac{1}{M_n}\sum_{t=1}^{M_n}\mathrm{E}\left(\frac{1}{M_n}|h(\lambda_t(\boldsymbol{\theta}_0))|^2|\mathcal{F}_{t-1}\right)\\
&= \sigma^2 + o_{p_n}(1),
\end{aligned}$$

and that the Lindeberg condition holds under our assumptions, we have that

$$T_n \Rightarrow N(\mathrm{E}(h(\lambda_t(\boldsymbol{\theta}_0))), \sigma^2)$$

by using the central limit theorem for martingale difference arrays (see, e.g., Pollard (1984)). When $M_n = n$, it holds that

$$T_n = (1,1)\begin{pmatrix} L^\top J^{-1} & O \\ O & 1 \end{pmatrix}\left(\sum_{t=1}^n \zeta_{t,n}\right) + \mathrm{E}(h(\lambda_t(\boldsymbol{\theta}_0))) + o_{p_n}(1),$$

where

$$\zeta_{t,n} = \begin{pmatrix} \frac{1}{\sqrt{n}}\frac{\partial}{\partial\boldsymbol{\theta}}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0)) \\ \frac{1}{\sqrt{n}}\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_n(\lambda_t(\boldsymbol{\theta}_0))\right\} \end{pmatrix}.$$

Noting that $\{\zeta_{t,n}\}_{1\le t\le n}$ is an $\mathbb{R}^{d+1}$-valued martingale difference array, we can apply the central limit theorem as well as the case $M_n = o(n)$ and obtain that

$$T_n \Rightarrow N(\mathrm{E}(h(\lambda_t(\boldsymbol{\theta}_0))), \tilde{\sigma}^2),$$

which implies our conclusion. □

S2 Additional examples of goodness of fit tests

This section serves goodness-of-fit tests for binomial and gamma distributions.

Goodness of fit test for binomial distribution. If we choose $\eta = \log(p/(1-p))$, $A(\eta) = m\log(1+\exp(\eta))$, and $h(z) = {}_m\mathrm{C}_z$, Z_η follows the binomial distribution with an unknown parameter p and a known parameter m. Then, the mean and variance are $\lambda := mp$ and $mp(1-p) = \lambda(1-\lambda/m)$, respectively. Here, we need to restrict the range of the intensity λ to $(0,m]$ to ensure $\mathrm{Var}(Z_t) \geq 0$. Under the appropriate moment conditions and the null G_0, the test statistic T_n^{bin}, defined as

$$T_n^{\mathrm{bin}} := \begin{cases} \hat{\sigma}^{-1} \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - (-\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) + r) \frac{\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)}{r} \right\} & M_n = o(n), \\ \hat{\hat{\sigma}}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - (-\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) + r) \frac{\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)}{r} \right\} & M_n = n, \end{cases}$$

where

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} \Bigg((6+2r)\tilde{\lambda}_t^4(\hat{\boldsymbol{\theta}}_n) + (12r + 4r^2)\tilde{\lambda}_t^3(\hat{\boldsymbol{\theta}}_n) + (7r^2 + 2r^3)\tilde{\lambda}_t^2(\hat{\boldsymbol{\theta}}_n) + r^3 \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) \Bigg) / r^3$$

and $\hat{\hat{\sigma}}^2$ is defined in Theorem 3 for $\kappa_0(\lambda_t(\boldsymbol{\theta})) = \lambda_t(\boldsymbol{\theta}) \left(1 - \lambda_t(\boldsymbol{\theta})/m\right)$, converges to the standard normal distribution, as $n \to \infty$.

Note that the Bernoulli distribution is the special case which corresponds to $m = 1$.

Goodness of fit test for gamma distribution. By choosing $\eta = -\beta$, $A(\eta) = -\log(-\eta)^{\alpha}$, and $h(z) = x^{\alpha-1}/\Gamma(\alpha)$, Z_η follows the Gamma distribution with a known shape parameter a and unknown rate β. Then the mean and variance are given by $\lambda := \alpha/\beta$ and $\alpha/\beta^2 = \lambda^2/\alpha$, respectively. Under appropriate moment condition and the null G_0, the test statistic T_n^{gam}, defined as

$$T_n^{\mathrm{gam}} := \begin{cases} \hat{\sigma}^{-1} \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \frac{\tilde{\lambda}_t^2(\hat{\boldsymbol{\theta}}_n)}{\alpha} \right\} & M_n = o(n) \\ \hat{\hat{\sigma}}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \frac{\tilde{\lambda}_t^2(\hat{\boldsymbol{\theta}}_n)}{\alpha} \right\} & M_n = n, \end{cases}$$

where

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} \frac{(2\alpha + 6)\, \tilde{\lambda}_t^4(\hat{\boldsymbol{\theta}}_n)}{\alpha^3}$$

and $\hat{\hat{\sigma}}^2$ is defined in Theorem 3 for $\kappa_0(\lambda_t(\boldsymbol{\theta})) = \lambda_t(\boldsymbol{\theta})^2/\alpha$, converges to the standard normal distribution as $n \to \infty$. When $\alpha = 1$, we have the goodness-of-fit test for the exponential distribution.

S3 Higher moments for several conditional distributions

This section provides the explicit expression of higher moments to calculate the asymptotic variances for several conditional distributions.

S3.1 Poisson distribution

Let Z be a random variable whose distribution is Poisson with a parameter l. The moment generating function is given by $e^{l(e^t-1)}$. Subsequently, the higher moments are given by

$$\mathrm{E}\,(Z) = l,$$

$$\begin{aligned}\mathrm{E}\left(Z^2\right) =& l(l+1),\\ \mathrm{E}\left(Z^3\right) =& l(l^2+3l+1),\\ \mathrm{E}\left(Z^4\right) =& l(l^3+6l^2+7l+1).\end{aligned}$$

A simple calculation gives

$$\mathrm{E}\left((Z-l)^2-l\right)^2 = 2l^2+l.$$

S3.2 Binomial distribution

Let Z be a random variable whose distribution is binomial with parameters p and m. Since the moment generating function is given by $(pe^t+(1-p))^n$, we have

$$\begin{aligned}\mathrm{E}\left(Z\right) =& mp = l,\\ \mathrm{E}\left(Z^2\right) =& mp(mp-p+1) = l + l^2 - l^2/m,\\ \mathrm{E}\left(Z^3\right) =& mp(m^2p^2+(1-3m)p^2+(3m-1)p+p^2-2p+1)\\ =& l+3l^2+l^3+(2l^3)/m^2-(3l^2)/m-(3l^3)/m,\\ \mathrm{E}\left(Z^4\right) =& mp\{m^3p^3+(-6m^2+4m-1)p^3+(6m^2-4m+1)p^2\\ &+(7m-4)p^3+(8-14m)p^2+(7m-4)p-p^3+3p^2-3p+1\}\\ =& -(6l^4)/m^3+(11l^4)/m^2-(6l^4)/m+l^4+(12l^3)/m^2-(18l^3)/m\\ &+6l^3-(7l^2)/m+7l^2+l.\end{aligned}$$

Hence, it holds that

$$\begin{aligned}\mathrm{E}\left((Z-l)^2-(l-l^2/m)\right)^2 =& l^4(-6/m^3+2/m^2)+l^3(-4/m+12/m^2)\\ &+l^2(2-7/m)+l.\end{aligned}$$

S3.3 Negative binomial distribution

Let Z be a random variable whose distribution is negative binomial with parameters r and p. The moment generating function is given $(p/(1-(1-p)e^t))^r$ for $t<-\log(1-p)$, and the higher order moments can be calculated as

$$\begin{aligned}\mathrm{E}\left(Z\right) =& -((p-1)r)/p = l,\\ \mathrm{E}\left(Z^2\right) =& ((p-1)r((p-1)r-1))/p^2 = (l^2r+l^2+lr)/r,\\ \mathrm{E}\left(Z^3\right) =& -((p-1)r((p^2-2p+1)r^2+((3-3p)r-p+1)+1))/p^3\\ =& (2l^3+3l^2r+3l^3r+lr^2+3l^2r^2+l^3r^2)/r^2,\\ \mathrm{E}\left(Z^4\right) =& (p-1)r\{(p^3-3p^2+3p-1)r^3+(-6p^2+12p-6)r^2\\ &+(-4p^2+8p-4)r-p^2+2p-1+((7p-7)r+4p-4)-1\}/p^4\\ =& (l^4r^3+6l^4r^2+11l^4r+6l^4+6l^3r^3+18l^3r^2+12l^3r+7l^2r^3+7l^2r^2\\ &+lr^3)/r^3.\end{aligned}$$

We can see that

$$\mathrm{E}\left(((Z-l)^2-(l^2+lr)/r)\right)^2 = (6l^4)/r^3+(2l^4)/r^2+(12l^3)/r^2+(4l^3)/r$$

$$+ (7l^2)/r + 2l^2 + l.$$

S3.4 Gamma distribution

Let Z be a random variable whose distribution is gamma distribution with a shape parameter a and a rate parameter b. Then, the moment generating function is given by $(b/(b-t))^a$ for $t < b$, which yields that

$$\begin{aligned}
\mathrm{E}(Z) =& a/b = l, \\
\mathrm{E}\left(Z^2\right) =& a(a+1)/b^2 = ((1+a)l^2)/a, \\
\mathrm{E}\left(Z^3\right) =& a(a+1)(a+2)/b^3 = ((a+1)(a+2)l^3)/a^2, \\
\mathrm{E}\left(Z^4\right) =& a(a+1)(a+2)(a+3)/b^4 = (a+1)(a+2)(a+3)l^4/a^3.
\end{aligned}$$

It is easy to see that

$$\mathrm{E}\left((Z-l)^2 - l^2/a)\right)^2 = (6l^4)/a^3 + (2l^4)/a^2.$$

Bibliography

Francq, C. and J. M. Zakoian (2010). *GARCH models: structure, statistical inference and financial applications.* John Wiley & Sons.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.