Journal of Statistical and Econometric Methods, vol.1, no.1, 2012, 93-109
ISSN: 2241-0384 (print), 2241-0376 (online)
International Scientific Press, 2012

On Non-Gaussian AR(1) Inflation Modeling

Werner Hürlimann[1]

Abstract

A severe limitation of the original autoregressive process of order one or AR(1) process is the Gaussian nature of the assumed residual error distribution while the observed sample residual errors tend to be much more skewed and have a much higher kurtosis than is allowed by a normal distribution. Four non-Gaussian noise specifications are considered, namely the normal inverse Gaussian, the skew Student t, the normal Laplace and the reshaped Hermite-Gauss distributions. Besides predictive distributional properties of some of these AR(1) processes, an in-depth analysis of the fitting capabilities of these models is undertaken. For the Swiss consumer price index, it is shown that the AR(1) with normal Laplace (NL) noise has the best goodness-of-fit in a dual sense for four types of estimators. On the one hand the moment estimators of the NL residual error distribution yield the smallest Anderson-Darling, Cramér-von Mises and chi-square statistics, and on the other hand the minimum of these three statistics is also reached by the NL distribution.

Mathematics Subject Classification: 62M10, 62P05, 91B84
Keywords: Force of inflation, AR(1), Normal inverse Gaussian, Skew Student t, Normal Laplace, Hermite-Gauss, Bera-Jarque statistic, Cramér-von Mises statistic, Anderson-Darling statistic

---

[1] FRSGlobal Switzerland AG, e-mail: whurlimann@bluewin.ch

Article Info: *Received* : January 8, 2012. *Revised* : February 3, 2012
Published online : February 28, 2012

1 Introduction

The economic importance of the inflation phenomenon is ubiquitous. For example, in actuarial science the inflation rate has been identified as the driving component of a comprehensive stochastic investment model in [45], [46]. This model has been widely discussed (e.g. [28], [40]) and also extended and refined (e.g. [16], [44]).

A severe limitation of the AR(1) inflation model is the Gaussian nature of the residual error while the observed sample residual errors tend to be much more skewed and have a much higher kurtosis than is allowed by a normal distribution. In fact, the Bera-Jarque statistic (see [10]) yields a simple test of rejection for the normal distribution. A brief outline of the content follows.

Section 2 recalls the modeling features of the force of inflation. Four non-Gaussian noise specifications are retained in Section 3. These are the normal inverse Gaussian (NIG) and the skew Student t (ST) members in the class of generalized hyperbolic distributions ([8]), the normal Laplace (NL) distribution ([39]), and the new analytical reshaped Hermite-Gauss (HG) distribution in [47]. Predictive distributional properties are stated in Section 4. Estimation of the residual error distribution is discussed in Section 5. Besides the moment and the minimum chi-square estimators, we consider those obtained by minimizing the Cramér-von Mises and Anderson-Darling statistics. Finally, Section 6 discusses the fitting capabilities of the specified non-Gaussian distributions for the Swiss consumer price index and shows that the NL noise has the best goodness-of-fit in a dual sense. On the one hand the moment estimators of the NL residual error distribution yield the smallest Anderson-Darling, Cramér-von Mises and chi-square statistics, and on the other hand the minimum of these three statistics is also reached by the NL distribution.

2 The Force of Inflation as Non-Gaussian AR(1) Process

Assume that "inflation" is measured by a price index whose value at time t is denoted by Q_t. One is interested in the following related quantities:

$q_t = \ln\{Q_t / Q_{t-1}\}$	:	force of inflation over the time period $(t-1, t]$
$R_t = \exp\{q_t\}$	:	accumulation factor for inflation over the period $(t-1, t]$
μI	:	(long-term) average force of inflation
ρ	:	deviation of the force of inflation from its average

Note that $J_t = R_t - 1$ represents the (annual) rate of inflation over the time period $(t-1, t]$. Rather than modeling the index itself, one looks at the force of inflation, for various reasons:

(i) The starting point is arbitrary and only the proportional changes of the index matter.

(ii) The index is positive and its growth can be very quick and high.
(iii) The logarithm of the index $\ln\{Q_t\}$ is not expected to be stationary (i.e. no fluctuation around a fixed mean) while the series (q_t) might be almost stationary.
The simplest discrete time series used to model the force of inflation is an autoregressive process of order one or AR(1) process, described by the linear difference equation

$$X_k = \rho \cdot X_{k-1} + Z_k, \quad X_k = q_k - \mu I, \quad k = 0, \pm 1, \pm 2, \ldots, \tag{2.1}$$

where the noise (also called residual error or innovation) $(Z_k)_{k \in Z}$, is an independent and identically distributed (i.i.d.) real-valued sequence. The normal distribution specification $Z_k \sim N(0, \sigma^2)$, with error variance σ^2, yields the so-called Wilkie inflation model.

The study of various financial time series, including inflation data, has revealed a clear evidence of dependence, heavy tails and asymmetry (e.g. [34]). These features require the availability of a variety of AR(1) processes driven by i.i.d. noise with non-Gaussian distributions exhibiting non-zero skewness, high kurtosis and heavy-tails.

Before proceeding let us recall some standard results in time series analysis. Recursive substitution shows that a solution of (2.1) takes the form

$$X_k = \rho^N \cdot X_{k-N} + \sum_{j=0}^{N-1} \rho^j \cdot Z_{k-j}, \quad \forall\, N = 1, 2, \ldots, \tag{2.2}$$

where X_{k-N} is an initial value N time periods back from k. The limit as $N \to \infty$ of (2.2) exists if and only if $|\rho| < 1$, and then $X_k \sim X_\infty = \sum_{j=0}^{\infty} \rho^j \cdot Z_{k-j}$ is a stationary solution.

The parameters $(\mu I, \rho)$ are estimated in a classical way. For a sample of size n the ordinary least squares (OLS) estimators are given by

$$\hat{\mu} I = \overline{q} = \tfrac{1}{n} \sum_{k=1}^{n} q_k, \quad \hat{\rho} = \frac{\sum_{k=2}^{n} \left(q_{k-1} - \hat{\mu} I\right) \cdot \left(q_k - \hat{\mu} I\right)}{\sum_{k=2}^{n} \left(q_{k-1} - \hat{\mu} I\right)^2}. \tag{2.3}$$

In [20] it is shown that $\hat{\rho}$ converges almost surely to ρ and a confidence interval based on the inequalities of Bernstein-Fréchet type is derived.

3 Some Non-Gaussian Noise Specifications

We begin with a brief survey about non-Gaussian AR(1) processes, which have been considered so far, and introduce those candidates that are attractive both from an analytical computational point of view and felt to fit financial data especially well.

After the first non-Gaussian process appeared in [32], an elaborate study of these models followed in [25]. An inverse Gaussian model is found in [2], a skew Laplace one in [41], a normal-Laplace one and a generalization as well as references to many other ones are given in [30]. A recent unified framework is [33].

We have retained four non-Gaussian noise distributions. They all enable capturing of the extra skewness and kurtosis of financial data. The first three are the normal inverse Gaussian and the skew Student t members in the class of generalized hyperbolic distributions and the normal Laplace distribution. The fourth one is the new "Hermite-Gauss" distribution proposed by [47]. Alternatives, which possess natural multivariate generalizations but are not considered here, include the four parameter modified skew normal by [3], which extends the three parameter skew normal by [12] (see [7]), and the skew Student t of [26] (see also [23] and [29]). More complex models might also be considered, for example the GARCH type models introduced by [13] (see [4], Section II.4, for a readable introduction). We mention that a primary goal of the ARCH model by [22] has been inflation modeling.

3.1 Normal Inverse Gaussian (NIG)

The normal inverse Gaussian distribution is a special case of the class of generalized hyperbolic distributions introduced by [8]. It has been considered in a finance context by [9] and is nowadays widely used. Consider a NIG distributed random variable $X = \mu + \delta \cdot Z \sim NIG(\mu, \delta, \alpha, \beta),\ \mu \in (-\infty, \infty),\ \delta, \alpha > 0,\ |\beta| < \alpha$, with $Z \sim SNIG(\alpha, \beta)$ a standard NIG random variable with $\mu = 0, \delta = 1$. Its cumulant generating function equals

$$C(t) = \sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + t)^2} .$$

Setting $c = \beta / \alpha, |c| < 1$, the mean, variance, skewness and (excess) kurtosis of this distribution are given by

$$\begin{aligned} &\mu_X = \mu + \frac{\delta c}{\sqrt{1-c^2}}, \quad \sigma_X^2 = \frac{\delta^2}{\alpha\left(1-c^2\right)^{3/2}}, \\ &\gamma_X = \frac{\kappa_3}{\kappa_2^{3/2}} = \frac{3c}{\alpha^{1/2}\left(1-c^2\right)^{1/4}}, \quad \gamma_{2,X} = \frac{\kappa_4}{\kappa_2^2} = \frac{3(1+4c^2)}{\alpha\left(1-c^2\right)^{1/2}}, \quad |\gamma_X| \le \sqrt{\frac{3}{5}\gamma_{2,X}} . \end{aligned} \tag{3.1}$$

The distribution function satisfies the normal mean-variance mixture representation

$$F_X(x) = \int_0^\infty \Phi\left(\frac{x-\mu-\delta\beta t}{\delta\sqrt{t}}\right)\frac{1}{\sqrt{2\pi}}t^{-\frac{3}{2}}\exp\left\{-\frac{1}{2}\frac{\left(1-\sqrt{\alpha^2-\beta^2}\,t\right)^2}{t}\right\}dt, \qquad (3.2)$$

with $\Phi(x)$ the standard normal distribution.

3.2 Skew Student t (ST)

The probability density function of a skew Student t distributed random variable $X = \mu + \beta\cdot Y + \sqrt{Y}\cdot Z \sim t(\upsilon,\mu,\delta,\beta)$, with $Z \sim N(0,1)$ (standard normal), $Y \sim I\Gamma\left(\frac{\upsilon}{2},\frac{\delta^2}{2}\right)$ (inverse Gamma), (Y,Z) independent, is defined for $\beta \neq 0$ by (e.g. [1])

$$f_X(x) = \frac{2^{\frac{1-\upsilon}{2}}\delta^\upsilon|\beta|^{\frac{\upsilon+1}{2}}K_{\frac{\upsilon+1}{2}}\left(|\beta|q(x)\right)\exp\{\beta(x-\mu)\}}{\sqrt{\pi}\Gamma\left(\frac{\upsilon}{2}\right)q(x)^{\frac{\upsilon+1}{2}}},$$
$$q(x) = \sqrt{\delta^2+(x-\mu)^2}, \quad \delta > 0. \qquad (3.3)$$

The mean, variance, skewness and kurtosis parameters are obtained as follows:

$$\mu_X = \mu + \frac{\beta\delta^2}{\upsilon-2}, \quad \upsilon > 2, \quad \sigma_X^2 = \frac{\delta^2}{\upsilon-2} + \frac{2\beta^2\delta^4}{(\upsilon-2)^2(\upsilon-4)}, \quad \upsilon > 4,$$
$$\gamma_X = \frac{2(\upsilon-4)^{\frac{1}{2}}\beta\delta}{\left\{2\beta^2\delta^2+(\upsilon-2)(\upsilon-4)\right\}^{\frac{3}{2}}}\left\{3(\upsilon-2)+\frac{8\beta^2\delta^2}{\upsilon-6}\right\}, \quad \upsilon > 6, \qquad (3.4)$$
$$\gamma_{2,X} = \frac{6\left\{(\upsilon-2)^2(\upsilon-4)+\frac{16\beta^2\delta^2(\upsilon-2)(\upsilon-4)}{\upsilon-6}+\frac{8\beta^4\delta^4(5\upsilon-22)}{(\upsilon-6)(\upsilon-8)}\right\}}{\left\{2\beta^2\delta^2+(\upsilon-2)(\upsilon-4)\right\}^2}, \quad \upsilon > 8.$$

The distribution function satisfies the integral representation ([21], Theorem 1)

$$F_X(x) = \int_0^\infty \Phi\left(\frac{\gamma(x-\mu)}{\sigma_X\sqrt{2t}} - \sqrt{2t}\right)\cdot p\left(t;\upsilon,\frac{\gamma^2}{2}\right)dt,$$
$$p\left(x;\upsilon,\frac{\gamma^2}{2}\right) = \frac{\left(\frac{\upsilon\gamma^2}{4}\right)^{\frac{\upsilon}{2}}}{\Gamma\left(\frac{\upsilon}{2}\right)}\cdot x^{-\frac{1}{2}(\upsilon+2)}\cdot\exp\left\{-\frac{\upsilon\gamma^2}{4x}\right\}, \quad \gamma = \frac{(\upsilon-2)\cdot(\mu_X-\mu)}{\upsilon\cdot\sigma_X}. \qquad (3.5)$$

The symmetric special case $\beta = 0$ yields the well-known Student t distribution.

3.3 Normal-Laplace (NL)

The distribution function of a normal-Laplace random variable $X = \upsilon + \tau \cdot Z + \alpha^{-1} \cdot E_1 - \beta^{-1} \cdot E_2 \sim NL(\upsilon, \tau, \alpha, \beta)$, with $Z \sim N(0,1)$ (standard normal), $E_1, E_2 \sim Exp(1)$ (standard exponential), (Z, E_1, E_2) independent, is given by (e.g. [38], [39])

$$F_X(x) = \Phi\left(\frac{x-\nu}{\tau}\right) - \varphi\left(\frac{x-\nu}{\tau}\right)\left\{\frac{\beta}{\alpha+\beta} \cdot R\left(\alpha\tau - \frac{x-\nu}{\tau}\right) - \frac{\alpha}{\alpha+\beta} \cdot R\left(\beta\tau + \frac{x-\nu}{\tau}\right)\right\} \tag{3.6}$$

with $R(z) = \dfrac{\overline{\Phi}(z)}{\varphi(z)}$, $\overline{\Phi}(z) = 1 - \Phi(z)$, the Mill's ratio. The mean, variance, skewness and kurtosis parameters are obtained as follows:

$$\begin{aligned} &\mu_X = \nu + \alpha^{-1} - \beta^{-1}, \quad \sigma_X^2 = \tau^2 + \alpha^{-2} + \beta^{-2}, \\ &\gamma_X = \frac{2\alpha^{-3} - 2\beta^{-3}}{\sigma_X^3}, \quad \gamma_{2,X} = \frac{6\alpha^{-4} + 6\beta^{-4}}{\sigma_X^4}. \end{aligned} \tag{3.7}$$

This family includes the skew Laplace if $\tau = 0$ (e.g. [31]) and normal-exponential distributions as $\alpha \to \infty$ or $\beta \to \infty$. We note that the normal-Laplace is the convolution of a normal and skew Laplace random variable. Since the normal and Laplace distributions constitute Laplace's first and second law of errors, it is worthy to consider a convolution of the two error distributions for modeling the residual AR(1) error. Note that, in contrast to [30], we model the non-Gaussian AR(1) noise and not the marginal distributions.

3.4 Hermite-Gauss (HG)

The motivation behind this reshaped Gaussian curve is the desire to preserve the bell-shaped property of the Gaussian curve and simultaneously accounts for variable skewness and kurtosis. The main Theorem 4.1 in [47] states that the Hermite-Gauss function

$$f_X(x) = \left(1 + \tfrac{\gamma_X}{6} H_3\left(\tfrac{x-\mu_X}{\sigma_X}\right) + \tfrac{\gamma_{2,X}}{24} H_4\left(\tfrac{x-\mu_X}{\sigma_X}\right)\right) \cdot \tfrac{1}{\sigma_X} \varphi\left(\tfrac{x-\mu_X}{\sigma_X}\right), \tag{3.8}$$

with $\varphi(x)$ the standard normal density and $H_k(x)$ the Hermite polynomial of degree k yields a unimodal probability density function of a random variable with given mean μ_X, variance σ_X^2, skewness γ_X, and (excess) kurtosis $\gamma_{2,X}$, provided $\gamma_{2,X} < 2.4$ and γ_X is bounded above by a positive constant depending on $\gamma_{2,X}$. Integration of (3.8) yields

$$F_X(x) = G\left(\tfrac{x-\mu_X}{\sigma_X}\right), \quad G(x) = \Phi(x) - \left(\tfrac{\gamma_X}{6} + \tfrac{\gamma_{2,X}}{24} x\right)\left(x^2 - 1\right)\varphi(x). \tag{3.9}$$

This simple distribution does not always capture the high kurtosis observed on financial markets, which can easily exceed a value of 6. Fortunately, the inflation data in the practical illustration of Section 6 fulfills the parameter restriction, hence the HG distribution is a genuine candidate.

4 Predictive Distributions of AR(1) Processes

In the present Section we derive distributional properties of the specified AR(1) processes and their stationary solutions. For completeness let us begin with a known result.

Proposition 4.1 *Let* $Z_k \sim N(0,\sigma^2)$ *be the stationary noise distribution of the inflation model (2.1) and suppose that* $\rho \in (0,1)$. *Then, the predictive distribution of the force of inflation at time* $t+N, N>0$ *given the force of inflation at time* $t>0$ *is normally distributed such that*

$$q_{t+N} \mid q_t \sim N\left(\mu I + \rho^N (q_t - \mu I), \frac{(1-\rho^{2N})\sigma^2}{1-\rho^2}\right). \tag{4.1}$$

The stationary solution of the AR(1) process exists and is described by the marginal distribution

$$X_\infty \sim N\left(0, \sigma^2/(1-\rho^2)\right) \tag{4.2}$$

Proof. Setting $k = t+N$ in (2.2) one gets $X_{t+N} \mid X_t = \rho^N \cdot X_t + \sum_{j=0}^{N-1} \rho^j \cdot Z_{t+N-j}$, which implies (4.1). The stationary solution $X_\infty = \sum_{j=0}^{\infty} \rho^j \cdot Z_{k-j}$ has the distribution (4.2). ◊

Next, consider some non-Gaussian cases. Since [25] it is known that a distribution function is the stationary marginal distribution of an AR(1) process with stationary noise if and only if it belongs to the class of self-decomposable distributions. In particular, this means that the characteristic functions

$\psi_\infty(s), \psi_Z(s)$ of $X_\infty, Z \sim Z_k$ satisfy the identity $\psi_\infty(s) = \psi_\infty(\rho \cdot s) \cdot \psi_Z(s)$. Expressions for the predictive and stationary distributions of the AR(1) process with non-Gaussian NIG and NL distributed noise follow.

Proposition 4.2 *Given is an AR(1) process with NIG noise* $Z_k \sim NIG(0, \delta, \alpha, \beta)$ *and* $\rho \in (0,1)$. *The predictive distribution of the force of inflation at time* $t > 0$ *is NIG distributed such that*

$$q_{t+N} | q_t \sim NIG\left(\mu I + \rho^N (q_t - \mu I), \frac{(1-\rho^N)\delta}{1-\rho}, \alpha, \beta \right). \tag{4.3}$$

The stationary solution of the AR(1) process is described by the NIG marginal distribution

$$X_\infty \sim NIG\left(0, \frac{\delta}{1-\rho}, \alpha, \beta \right). \tag{4.4}$$

Proof. One has $Z_k = \delta \cdot S_k$ with $S_k \sim SNIG(\alpha, \beta)$ a standard NIG random variable. Since $\rho^j Z_k = \delta\rho^j \cdot S_k \sim NIG(0, \delta\rho^j, \alpha, \beta)$ and NIG is closed under convolution, i.e. $NIG(\mu_1, \delta_1, \alpha, \beta) * NIG(\mu_2, \delta_2, \alpha, \beta) \sim NIG(\mu_1 + \mu_2, \delta_1 + \delta_2, \alpha, \beta)$, one obtains (4.3) from the identity $X_{t+N} | X_t = \rho^N \cdot X_t + \sum_{j=0}^{N-1} \rho^j \cdot Z_{t+N-j}$. Letting $N \to \infty$ shows (4.4). ◊

Proposition 4.3 *Given is an AR(1) process with NL noise* $Z_k \sim NL(\upsilon = \beta^{-1} - \alpha^{-1}, \tau, \alpha, \beta)$ *and* $\rho \in (0,1)$. *The predictive distribution of the force of inflation at time* $t + N, N > 0$ *is the convolution of a normal and a sum of independent skew Laplace random variables such that*

$$q_{t+N} | q_t \sim N\left(\mu I + \rho^N (q_t - \mu I) + \frac{1-\rho^N}{1-\rho} \cdot \upsilon, \frac{1-\rho^{2N}}{1-\rho^2} \cdot \tau^2 \right) + sL_N, \tag{4.5}$$

where the distribution function $F_N(x)$ *of* sL_N *is given by*

$$F_N(x) = \begin{cases} \sum_{m=0}^{[(2N-1)/2]} c_{2m+1}^{2N-1} C_{2m+1}^{-1} \exp\left(-c_{2m+1}^{-1} x\right), & x < 0, \\ 1 - \sum_{m=0}^{[(2N-1)/2]} c_{2m}^{2N-1} C_{2m}^{-1} \exp\left(-c_{2m}^{-1} x\right), & x \geq 0, \end{cases} \tag{4.6}$$

with the coefficients defined by

$$c_j = \begin{cases} \alpha^{-1}\rho^{j/2}, & j = 2m, \\ -\beta^{-1}\rho^{(j-1)/2}, & j = 2m+1, \end{cases}$$

$$C_j = \prod_{k=0,\, k\neq j}^{2N-1} \left(c_j - c_k\right),\ j = 0,1,\ldots,2N-1. \tag{4.7}$$

The stationary solution of the AR(1) process is described by the convolution

$$X_\infty \sim N\left(\frac{\upsilon}{1-\rho}, \frac{\tau^2}{1-\rho^2}\right) + sL_\infty, \quad sL_\infty = \lim_{N\to\infty} sL_N. \tag{4.8}$$

Proof. Starting point is again the identity

$$X_{t+N} \left| X_t \right. = \rho^N \cdot X_t + Y_N, \quad Y_N = \sum_{j=0}^{N-1} \rho^j \cdot Z_{t+N-j}$$

with $Z_{t+N-j} = \upsilon + \tau \cdot G_{t+N-j} + \alpha^{-1} \cdot E_{1,t+N-j} - \beta^{-1} \cdot E_{2,t+N-j}$, $G_{t+N-j} \sim N(0,1)$, $E_{i,t+N-j} \sim Exp(1)$, $i = 1,2$. We obtain the independent sum $Y_N = sG_N + sL_N$ with $sG_N = N\left(\frac{1-\rho^N}{1-\rho} \cdot \upsilon, \frac{1-\rho^{2N}}{1-\rho^2} \cdot \tau^2\right)$, $sL_N = \sum_{j=0}^{N-1} \rho^j \cdot \left(\alpha^{-1} \cdot E_{1,t+N-j} - \beta^{-1} \cdot E_{2,t+N-j}\right)$, where the second term is a sum of independent skew Laplace random variables. Its distribution is derived using the result by [5] on the distribution of a linear combination of independent exponential random variables (see [41], Proposition 1). ◊

Remarks 4.1 First, the finite sum sL_N converges absolutely with probability one and in mean square to the infinite sum sL_∞ in (4.8) ([14], Prop. 3.1.1). In calculations, the infinite sum can be truncated after a few terms because the coefficients c_j decay at a geometric rate ([14], Chap. 3). Second, as follows from our numerical illustration in Section 6, the AR(1) with NL noise has the best goodness-of-fit among the considered non-Gaussian AR(1) candidates. For this reason and while the predictive distributions for ST and GH noises are more complex, they are not considered further.

5 Residual Error Estimation and Goodness-of-Fit

Given a data set of sample size n, the goodness-of-fit of the residual error distribution is based on a statistics, which measures the difference between the empirical distribution function $F_n(x)$ and the fitted distribution function $F(x)$. We use the Cramér-von Mises family of statistics defined by (e.g. [19], [17] and [15])

$$T = n \cdot \int_{-\infty}^{\infty} \left[F_n(x) - F(x) \right]^2 w(x) dF(x), \tag{5.1}$$

where $w(x)$ is a suitable weighting function. If $w(x) = 1$ one obtains the W^2 *Cramér-von Mises* statistic ([18], p.145-47, [35], p.316-35). If $w(x) = 1/\left[F(x)\overline{F}(x)\right]$ one gets the A^2 *Anderson-Darling* statistic (see [6]). Consider the order statistics of the error data such that $x_1 \le x_2 \le ... \le x_n$ and let $\hat{F}(x_i), i = 1,...,n,$ be the fitted values of the distribution function. Then one has the formulas

$$W^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left(\hat{F}(x_i) - \frac{2i-1}{n} \right)^2, \; A^2 = -n - \sum_{i=1}^{n} \frac{2i-1}{n} \cdot \ln\left\{ \hat{F}(x_i) \cdot \hat{\overline{F}}(x_{n-i+1}) \right\}. \tag{5.2}$$

One knows that A^2 yields one of the most powerful statistical test if the fitted distribution departs from the true distribution in the tails (e.g. [19]). In this situation, it is the recommended goodness-of-fit statistics. Now, the observed sample residual errors of our inflation data in Section 6 tend to be much more skewed and have a much higher kurtosis than is allowed by a normal distribution, which indicates that the fit in the tails matters. For this reason, we propose to estimate our non-Gaussian noise distributions by calculating the parameters, which minimize the Cramér-von Mises and Anderson-Darling goodness-of-fit statistics, simply called *minW* and *minA estimators*. This procedure automatically ensures optimal goodness-of-fit with respect to the corresponding fitting criteria.

We also apply a standard minimum chi-square estimation method to obtain *minχ2 estimators*, by minimizing the following goodness-of-fit statistic (introduced by [37]):

$$\chi^2 = \sum_{i=1}^{n-1} \frac{\left(\hat{F}(x_i) - \hat{F}(x_{i+1}) - (n-1)^{-1} \right)^2}{\hat{F}(x_i) - \hat{F}(x_{i+1})}. \tag{5.3}$$

Maximum likelihood estimation, which is asymptotically equivalent to the minimum chi-square method, as first proved by [24], is not considered here. It can be argued that minimum chi-square and not maximum likelihood is the basic principle of statistical inference (e.g. [11], [27]).

The method of moment estimators, abbreviated *MM estimators* (introduced by [36]) is also used. Though [24] argued in favor of minχ2 rather than MM estimators, we use them for comparative purposes. It is assumed that the distribution parameters can be written as functions of the moments. The method of moment replaces these moments by their unbiased empirical cumulant counterparts defined as follows (assume $n > 3$):

$$\hat{\kappa}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \hat{\kappa}_2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\kappa}_1)^2, \quad \hat{\kappa}_3 = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}(x_i - \hat{\kappa}_1)^3,$$
$$\hat{\kappa}_4 = \frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}(x_i - \hat{\kappa}_1)^4 - 3\cdot\frac{n}{(n-2)(n-3)}\left(\sum_{i=1}^{n}(x_i - \hat{\kappa}_1)^2\right)^2. \quad (5.4)$$

The empirical skewness and (excess) kurtosis are $\hat{\gamma}_X = \hat{\kappa}_3 / \hat{\sigma}_X^3$ and $\hat{\gamma}_{2,X} = \hat{\kappa}_4 / \hat{\sigma}_X^4$. To test normality one uses the Bera-Jarque statistic $JB = n\cdot\left(\hat{\gamma}_X / 6 + \hat{\gamma}_{2,X} / 24\right)$, which is asymptotically χ_2^2 distributed, and has a critical value of 5.99 for a 95% confidence level.

5.1 NIG MM estimators

Since the MM estimators are directly obtained from (3.2) their formulas are not stated.

5.2 ST MM estimators

Following [1] the parameters (μ, δ, β) are obtained as functions of the remaining parameter υ through the formulas

$$\kappa = \kappa(\upsilon) = \frac{1}{3\upsilon^2 - 2\upsilon - 32}\cdot \left(1 - \sqrt{1 - \frac{(3\upsilon^2 - 2\upsilon - 32)\cdot\left[12(5\upsilon - 22) - (\upsilon - 6)(\upsilon - 8)\gamma_{2,X}\right]}{216(\upsilon - 2)^2(\upsilon - 4)}}\right),$$
$$\beta = \beta(\upsilon) = \frac{\operatorname{sgn}(\gamma_X)}{6(\upsilon - 2)\sigma_X \kappa}\sqrt{\frac{1 - 6(\upsilon - 2)(\upsilon - 4)\kappa}{2(\upsilon - 4)}}, \quad (5.5)$$
$$\delta = \delta(\upsilon) = (\upsilon - 2)\sigma_X\sqrt{6(\upsilon - 4)\kappa}, \quad \mu = \mu(\upsilon) = \mu_X - \frac{\beta\delta^2}{\upsilon - 2},$$

where the unknown υ solves the skewness equation

$$(\upsilon - 6)\cdot|\gamma_X| = \left[4 - 6(\upsilon - 2)(\upsilon + 2)\kappa\right]\cdot\sqrt{2(\upsilon - 4)\left[1 - 6(\upsilon - 2)(\upsilon - 4)\kappa\right]} \quad (5.6)$$

5.3 NL MM estimators

The moment equations cannot always be solved for all feasible combinations of mean, variance, skewness and kurtosis. The following alternative estimation procedure applies. Given a NL distributed random variable (3.6), solve the skewness equation to get

$$\beta = \beta(\alpha) = \left(\alpha^{-3} - \tfrac{1}{2}\gamma_X \sigma_X^3\right)^{-1/3} \tag{5.7}$$

and insert into the kurtosis equation to solve for the parameter α:

$$6\alpha^{-4} + 6\left(\alpha^{-3} - \tfrac{1}{2}\gamma_X \sigma_X^3\right)^{4/3} - \gamma_{2,X}\sigma_X^4 = 0\,. \tag{5.8}$$

<u>Case 1</u>: $\sigma_X^2 - \alpha^{-2} - \beta^{-2} \geq 0$
The moment estimators exist and one has

$$\tau = \sqrt{\sigma_X^2 - \alpha^{-2} - \beta^{-2}}, \quad \upsilon = \mu_X - \alpha^{-1} + \beta^{-1}.$$

<u>Case 2</u>: $\sigma_X^2 - \alpha^{-2} - \beta^{-2} < 0$
The moment estimators do not exist. This means that the kurtosis is higher than the one that can be obtained from a Normal Laplace. In this situation the mean, variance and skewness can be calibrated by considering the special case $\tau = 0$ of a skew Laplace.
Let $X \sim skL(\upsilon, \alpha, \beta)$ be a skew Laplace distributed random variable. Assuming that $\alpha \geq \sigma_X^{-1}$, solve the variance equation to get $\beta = \beta(\alpha) = \left(\sigma_X^2 - \alpha^{-2}\right)^{-1/2}$, and insert into the skewness equation to solve for α:

$$2\alpha^{-3} - 2\left(\sigma_X^2 - \alpha^{-2}\right)^{3/2} - \gamma_X \sigma_X^3 = 0\,. \tag{5.9}$$

<u>Case 3</u>: $\alpha \geq \sigma_X^{-1}$
The remaining parameter is $\upsilon = \mu_X - \alpha^{-1} + \beta^{-1}$.

<u>Case 4</u>: $\alpha < \sigma_X^{-1}$
Consider the further special case of a symmetric Laplace $X \sim skL\left(\upsilon = \mu_X, \alpha = \frac{\sqrt{2}}{\sigma_X}, \beta = \alpha\right)$.

5.4 HG MM estimators

The MM estimators of $\mu_X, \sigma_X^2, \gamma_X, \gamma_{2,X}$ coincide with their unbiased empirical counterparts.

Finally, the calculation of the minW, minA and minχ2 estimators uses expressions of the distribution functions that can be computed within a reasonable computing time. Since not all software packages ensure this requirement for the NIG and ST distributions, a special software design must be implemented for them. Consider the integral representations (3.2) and (3.5), which are of the type

$$I = \int_0^\infty f(x)dx = \int_0^1 g(t)dt, \quad g(t) = t^{-2} f(t^{-1} - 1). \tag{5.10}$$

We apply the standard trapezoidal rule and Simpson's rule, as well as the recent modified Simpson's rule by [43], which yields a better error bound than Simpson's rule (see [42]). These numerical quadrature formulas can be used under the valid assumption $g(0) = g(1) = 0, g'(0) = g'(1) = 0$ to yield for an approximation of N -th order:

$$I = \int_0^1 g(t)dt \approx \frac{1}{N} \cdot \sum_{k=1}^{N-1} C_k^{rule} \cdot g\left(\frac{k}{N}\right) \tag{5.11}$$

where the coefficients C_k^{rule} depend on the chosen rule as follows:

Trapezoidal rule : $C_k^T = 1$

Simpson's rule : $C_k^S = \begin{cases} \frac{4}{3}, & k\ odd \\ \frac{2}{3}, & k\ even \end{cases}$

Modified Simpson's rule: $C_k^{MS} = \begin{cases} \frac{16}{15}, & k\ odd \\ \frac{14}{15}, & k\ even \end{cases}$

We use the modified Simpson's rule with $N = 500$ for the NIG and $N = 2'000$ for the ST.

6 Numerical Illustration

We discuss briefly the fitting capabilities of the specified distributions using the consumer price index (CPI) of Switzerland. It is downloadable from the website of the Federal Office of Statistics at

www.bfs.admin.ch/bfs/portal/de/index/themen/05/02/blank/key/basis_aktuell.html

We have chosen the time span between 1925 and 2009 on the June 1914 index base of 100 containing 85 yearly indices. The estimated AR(1) parameters (2.3) are $\hat{\mu} = 0.02165130$ and $\hat{\rho} = 0.66197187$. Since the Bera-Jarque statistic JB=18.41 is far away from its critical value 5.99, normality is rejected at the 95% confidence level. Table 6 shows that the AR(1) with NL noise has the best goodness-of-fit in the following dual sense. The MM estimators of the NL residual error distribution yield the smallest A^2, W^2, χ^2 statistics, and the minimum of these three statistics is also reached by the NL distribution.

Table 6: Fitting the AR(1) residual error distributions for the Swiss CPI

estimators	noise	goodness-of-fit statisstics A^2	W^2	χ^2	moments of fitted noise distribution μ_X	σ_X	γ_X	$\gamma_{2,X}$
MM	NIG	0.42653	0.08214	2.89125	0	0.02524	0.82549	2.0215
	ST	1.64622	0.28631	3.17875	0	0.02524	0.82549	2.0215
	NL	**0.40591**	**0.07709**	**2.87054**	**0**	**0.02524**	**0.82549**	**2.0215**
	HG	0.72216	0.12159	3.06631	0	0.02524	0.82549	2.0215
	Normal	1.05092	0.18188	3.03751	0	0.02524	0	0
minA	NIG	0.21197	0.02748	3.04653	0.00608	0.03225	2.02184	12.1267
	ST	0.2216	0.03099	2.99821	0.00086	0.01691	-	-
	NL	**0.20909**	0.02551	3.05475	**-0.00071**	**0.02599**	**0.29996**	**3.00734**
	HG	0.27466	0.03171	3.11513	-0.00051	0.02098	0.44541	2.48129
minW	NIG	0.24348	0.02371	3.16403	0.00632	0.03683	2.70594	21.7211
	ST	0.22594	0.02616	3.09535	0.04024	0.01612	-	-
	NL	0.23445	**0.02256**	3.03664	**-0.00145**	**0.02595**	**0.0302**	**3.00061**
	HG	0.31254	0.02611	3.3838	-0.0011	0.02003	0.19609	3.26934
minχ2	NIG	0.54573	0.09245	2.83694	0.00803	0.02598	0.68075	1.37474
	ST	0.65867	0.11111	2.89345	-0.00277	0.01866	-	-
	NL	0.54489	0.09177	**2.82304**	**-0.00196**	**0.02317**	**0.46443**	**1.04687**
	HG	0.50345	0.08174	2.86258	-0.00145	0.02275	0.36564	0.57171

References

[1] K. Aas and H. Haff, The generalized hyperbolic skew Student's t-distribution. *J. Fin. Econom.*, **4**(2), (2006), 275-309.

[2] B. Abraham and N. Balakrishna, Inverse Gaussian autoregressive models, *J. Time Series Analysis*, **20**(6), (1999), 605-618.

[3] C.J. Adcock and K. Shutes, *Portfolio selection based on the multivariate-skew normal distribution*, In: A. Skumilowski (Ed.), *Financial Modeling*, Progress Business Publ., Krakow, 2001.

[4] C. Alexander, *Market Risk Analysis, Volume II: Practical Financial Econometrics*, J. Wiley & Sons Ltd., Chichester, 2008.

[5] M.M. Ali and M. Obaidullah, Distribution of linear combination of exponential variates, *Comm. Statist.-Theory and Methods*, **11**(13), (1982), 1453-1463.

[6] T.W. Anderson and D.A. Darling, Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes, *Ann. Math. Statist.*, **23**(2), (1952), 193-212.

[7] A. Azzalini, A class of distributions which includes the normal ones, *Scand. J. Statist.*, **12**(2), (1985), 171-178.

[8] O.E. Barndorff-Nielsen, Hyperbolic distributions and distributions on hyperbolae, *Scand. J. Statist.*, **5**(3), (1978), 151-157.

[9] O.E. Barndorff-Nielsen, Normal inverse Gaussian distributions and stochastic volatility modeling, *Scand. J. Statist.*, **24**(1), (1997), 1-13.

[10] A.K. Bera and C.M. Jarque, A Test for Normality of Observations and Regression Residuals, *Int. Statist. Rev.*, **55**(2), (1987), 163-172.

[11] J. Berkson, Minimum chi-square, not maximum likelihood!, *Ann. Statist.*, **8**(3), (1980), 457-487.

[12] Z.W. Birnbaum, Effect of linear truncation on a multinormal population, *Ann. Math.* Statist., **21**(2), (1950), 272-279.

[13] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, *J. Econom.*, **31**(3), (1986), 307-327.

[14] P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods*, 2nd edition, Springer, New York, 1991.

[15] K. Burnecki, A. Misiorek and R. Weron, Loss Distributions, MPRA *Paper No.* **22163**, (2010).

[16] T. Chan, Some applications of Lévy processes to stochastic investment models for actuarial use, *ASTIN Bull.*, **28**(1), (1998), 77-93.

[17] P. Cizek, W. Härdle and R. Weron, *Statistical Tools for Finance and Insurance*, Springer, New York, 2005.

[18] H. Cramér, On the composition of elementary errors, Second Paper: Statistical Applications, *Skand. Akt.*, **1928**(1), (1928), 141-180.

[19] R.B. D'Agostino and M.A. Stephens, *Goodness-of-Fit Techniques*, Marcel Dekker, New York, 1986.

[20] A. Dahmani and M. Tari, Bernstein-Fréchet inequalities for the parameter of the order 1 autoregressive process, *Interstat*, **Nov.** (6), (2002), 1-6.

[21] S. Dokov, S.V. Stoyanov and S.T. Rachev, Computing VaR and AVaR of skewed-t distribution, *J. Appl. Functional Analysis*, **3**(1), (2008), 189-208.

[22] R. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, (1982), 987-1007.

[23] C. Fernandez and M.F.J. Steel, On Bayesian modeling of fat tails and skewness, *J. Amer. Statist. Assoc.*, **93**(441), (1998), 359-371.

[24] R. Fisher, R., The conditions under which chi-square measures the discrepancy between observation and hypothesis, *J. Royal Statist. Soc.*, **87**, (1924), 442-450.

[25] D.P. Gaver and P.A.W. Lewis, First order autoregressive Gamma sequences and point processes, *Adv. Appl. Prob.*, **12**(3), (1980), 727-745.

[26] B.E. Hansen, Autoregressive conditional density estimation, *Int. Econ. Rev.*, **35**(3), (1994), 705-730.

[27] R.R. Harris and G.K. Kanji, On the use of minimum chi-square estimation, *The Statistician*, **32**, (1983), 379-394.

[28] P.P. Huber, A review of Wilkie's stochastic asset model, *British Act. J.,* **3**, (1996), 181-210.

[29] E. Jondeau, S.-H. Poon and M. Rockinger, *Financial modeling under non-Gaussian distributions*, Springer Finance, London, 2007.

[30] K.K. Jose, T. Lishamol and J. Sreekumar, Autoregressive processes with normal-Laplace marginal, *Statist. Prob. Letters*, **78**(15), (2008), 2456-2462.

[31] S. Kotz, T.J. Kozubowski and K. Podgorski, *The Laplace distribution and generalizations: a revisit with applications to Communications, Economics, Engineering, and Finance*, Birkhäuser, Boston, 2001.

[32] A.J. Lawrance, Some autoregressive models for point processes, *Proc. Bolyai Math. Soc. Colloq. point processes and queuing problems*, **24**, (1978), 257-275.

[33] T. Lishamol and K.K. Jose, A unified framework for Gaussian and non-Gaussian AR(1) modeling, *J. Prob. Statist. Sci.*, **8**(1), (2010), 97-105.

[34] Y. Malevergne and D. Sornette, *Extreme Financial Risks – From Dependence to Risk Management*, Springer, Berlin, 2006.

[35] R. von Mises, *Wahrscheinlichkeitsrechnung und Ihre Anwendung in der Statistik und Theoretischen Physik*, Deuticke, Leipzig, 1931.

[36] K. Pearson, Contribution to the mathematical theory of evolution, *Phil. Trans. Royal Soc. London, Ser. A*, **185**, (1894), 71-110.

[37] K. Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Philosophical Magazine Series*, **50**, (1900), 157-175.

[38] W.J. Reed, *The normal Laplace distribution and its relatives*, In: N. Balakrishna, E. Castillo and J.M. Sarabia (Eds.), *Advances in Distribution Theory, Order Statistics and Inference, Statistics for Industry and Technology*, Birkhäuser, Boston, 2006.

[39] W.J. Reed and M. Jorgensen, The double Pareto-lognormal distribution – a new parametric model for size distributions, *Comm. Statist.-Theory and Methods*, **33**(8), (2004), 1733-1753.

[40] S. Sahin, A. Cairns, T. Kleinow and A.D. Wilkie, Revisiting the Wilkie investment model, In: *Proc. 18th Int. AFIR Colloquium*, Rom, available at www.actuaries.org, 2008.

[41] A. Trindade, Y. Zhu and B. Andrews, Time series models with asymmetric Laplace innovations, *J. Statist. Comp. Simulation*, **80**, (2010), 1317-1333.

[42] N. Ujevic, New error bounds for the Simpson's quadrature rule and applications, *Computers & Math. with Appl.*, **53**(1), (2007), 64-72.

[43] N. Ujevic and A.J. Roberts, A corrected quadrature formula and applications, The Australian & New Zealand Industrial and Appl. Math. J., **45**(E), (2004), E41-E56.

[44] S.P. Whitten and R.G. Thomas, A non-linear stochastic asset model for actuarial use, *British Actuar. J.*, **5**, (1999), 919-953.

[45] A.D. Wilkie, A Stochastic Investment Model for Actuarial Use, *Transactions of Faculty of Actuaries*, **39**, (1986), 341-403.
[46] A.D. Wilkie, More on a Stochastic Model for Actuarial Use, *British Actuar. J.*, **1**, (1995), 777-964.
[47] M.G. Zoia, Tailoring the Gaussian law for excess kurtosis and skewness by Hermite polynomials, *Comm. Statist. Theor. Meth.*, **39**, (2010), 52-64.