Thinning Pearson type 3 distribution and modelling time series with Pearson type 3 marginal

Abstract

The Pearson type 3 distribution also called three-parameters gamma distribution is widely used in various scientific fields such as in reliability and hydrology. We focused our investigations on its applications in hydrologic frequency analysis for modelling annuals maximum flows. Its probability density function(p.d.f) is given by:

$$f(x,\nu,\lambda,\beta) = \frac{\beta^{\lambda}(x-\nu)^{\lambda-1}}{\Gamma(\lambda)} \exp\{-\beta(x-\nu)\} 1_{[0,+\infty[}, 0 < \nu \leq x,\ \lambda, \beta > 0 \tag{1}$$

where ν, λ, β denote respectively location, shape and scale parameters.
Our first goal consist to provide an stationary first order Markov AR(1) time series model based on the thinning operation of Joe (1996). To define this operator in the case of Pearson type 3 margin, the following proposition is required.

Proposition 1. *Let X denote a positive random variable distributed according to a Pearson type 3 distribution with location parameter ν, scale parameter β and shape parameter λ_X. Let Y denote a positive and independent random variable distributed according to a gamma distribution with scale parameter β and shape parameter λ_Y. The following assertions hold;*
1. *the random variable $Z = X + Y$ is distributed according to the Pearson type 3 distribution with location parameter ν, scale parameter β and shape parameter $\lambda = \lambda_X + \lambda_Y$.*
2. *the distribution of X given $Z = z$ is the same as that of the random variable $(z-\nu)B + \nu$ where $B \backsim Beta(\lambda_X, \lambda_Y)$.*

From this proposition, we define a convex thinning operator as following

Definition 1. *Let Z be a positive random variable distributed according to the Pearson type 3 distribution with location parameter ν, scale parameter β and shape parameter λ. Let $\alpha \in]0,1[$; the random variable $B(Z,\nu,\alpha)$ is called convex thinning of Z if its distribution conditional to $Z = z$ is the same as that of $(z-\nu)Beta(\alpha\lambda, (1-\alpha)\lambda) + \nu$.*

Furthermore, we proved that this convex thinning operator is distributed according to Person type 3 as stated the following proposition:

Proposition 2. *Let Z be a positive random variable distributed according to the Pearson type 3 distribution with location parameter ν, scale parameter β and shape parameter λ. Given $\alpha \in]0,1[$, one has:*
$B(Z,\nu,\alpha)$ is distributed according to a Pearson type 3 with location parameter ν, scale parameter β and shape parameter $\alpha\lambda$.

This last proposition allow us to formulate our model as follow.

An hierarchical auto-regressive model for time series with Pearson type 3 margin

The model

Let $(X_t)_{t\in\mathbb{N}}$ denote a time series such that:
1. $[X_t|X_{t-1}=x_{t-1}]=(x_{t-1}-\nu)U_t+\nu+\epsilon_t$;
2. $(U_t)_{t\in\mathbb{N}}$ is a sequence of independent random variables identically distributed according to a beta distribution with parameter $\alpha\lambda$ and $(1-\alpha)\lambda$, $\alpha\in]0,1[$ and U_t is independent of the sequence $(X_{t-s})_{s\geqslant 1}$;
3.$(\epsilon_t)_{t\in\mathbb{N}}$ is a sequence of random variables independent and identically distributed according to gamma distribution with scale parameter β and shape parameter $(1-\alpha)\lambda$;
4. the sequences $(U_t)_{t\in\mathbb{N}}$ and $(\epsilon_t)_{t\in\mathbb{N}}$ are mutually independent ;
5. X_0 is distributed according to a Pearson type 3 distribution with location parameter ν, scale parameter β and shape parameter λ.

Autocorrelation function

Inspired by the work of Joe(1996), we proved that the autocorrelation of lag j for this time series models $\rho(j)$ is given by:

$$\rho(j)=\frac{\alpha^j(\lambda\beta^2)}{(\lambda\beta^2)}=\alpha^j$$

Likelihood function and inference

Let $(x_t)_{t=0:T}$ be a time series generated according to the model stated above. Let $L(\theta|x_t,t=0:T)$ denote the likelihood function of the parameter vector $\theta={}^t(\nu,\beta,\lambda,\alpha)$. Using the markovity of the process, the likelihood function is given by

$$L(\theta|x_t,t=0:T)=f(x_0)\prod_{t=1}^{n}f(x_t|x_{t-1},\theta) \tag{2}$$

where

$$f(x_t|x_{t-1},\theta)=\int f(x_t|x_{t-1},u_t,\theta)g(u_t,\alpha,\lambda)du_t$$

However, the integral in the conditional distribution of x_t given x_{t-1} is not analytically tractable. To overcome this difficulty, we consider that $\{(x_t)_{t=0:T},(u_t)_{t=1:T}\}$ are the complete data where $(u_t)_{t=1:T}$ are missing data. Using Jensen inequality, a lower bound for the log-likelihood is obtained as follow:

$$\begin{aligned}\log(L(\theta|x_t,t=0:T))&\geq\log\frac{f(x_0|\theta)}{f(x_0|\theta')}+\log(L(\theta'|x_t,t=0:T))\\&-\sum_{t=1}^{T}\frac{1}{f(x_t|x_{t-1},\theta')}\int\log\{\frac{f(x_t,u_t|x_{t-1},\theta')}{f(x_t,u_t|x_{t-1},\theta)}\}f(u_t,x_t|x_{t-1},\theta')du_t\end{aligned}$$

The maximization of this lower bound is still difficult, since the marginal likelihood $f(x_t|x_{t-1},\theta')$ is unknown. However, note that the precedent relationship is equivalent to:

$$\begin{aligned}&\log(L(\theta|x_t,t=0:T))-\log(L(\theta'|x_t,t=0:T))\\&\geq\log\frac{f(x_0|\theta)}{f(x_0|\theta')}-\sum_{t=1}^{T}\frac{1}{f(x_t|x_{t-1},\theta')}\int\log\{\frac{f(x_t,u_t|x_{t-1},\theta')}{f(x_t,u_t|x_{t-1},\theta)}\}f(u_t,x_t|x_{t-1},\theta')du_t\\&=\log\frac{f(x_0|\theta)}{f(x_0|\theta')}-K(\theta|\theta')\\&=\nabla(\theta|\theta')\end{aligned}$$

To avoid the problem of unknown marginal likelihood, we propose Monte Carlo Generalised EM (GEM)algorithm. This is an iterative method consisting to increase the function to be maximized at each iteration. In our case, we have proved that it suffices to minimize the K-function.

Each iteration consists of an E-step and an M-step. The $(n+1)th$ E-step entails the calculation of $\nabla(\theta|\theta^{(n)})$ where $\theta^{(n)}$ denote the values of θ from the nth iteration. Monte Carlo simulation is used to estimate the quantity in the E-step. The new value $\theta^{(n+1)}$ is chosen to increase $\nabla(\theta|\theta^{(n)})$, then to minimise $K(\theta|\theta^{(n)})$.

Given a starting value $\theta^{(0)}$, iteration between the two steps of the GEM algorithm produces a sequence that, under regularity conditions(Dempster and al., 1977), converges to the maximum likelihood estimate $\hat{\theta}$.