Actuarially Market Consistent Valuation of Catastrophe Bonds

Saeid Safarveisi*

Actuarial Research Group, AFI, Faculty of Economics and Business, KU Leuven, Leuven, Belgium

saeid.safarveisi@kuleuven.be

Dixon Domfeh

Department of Computer Science, Georgia Institute of Technology, Atlanta, United States

dnkwantabisa3@gatech.edu

Arpita Chatterjee

Department of Mathematical Sciences, Georgia Southern University, Georgia, United States

achatterjee@georgiasouthern.edu

Version: June 15, 2022

Abstract

[To be completed ...]

Keywords:

*Corresponding author

1 Introduction

Dealing with monetary losses resulting from insurance-related, non-financial risks has been the main concern of insurance companies. This matter becomes even more crucial when the risks under consideration are linked to extreme events, in which case insurance companies are exposed to higher risk layers. In addition to the traditional reinsurance solution, the risk management strategy suggests the use of insurance-linked securities which provides insurers with market-based protection by transferring their risks to the capital market. Through this mechanism, insurance companies can raise capital and capacity. One category of insurance-linked securities that covers non-life risks and has received the most attention in academic research and industry is catastrophe bonds or, shortly, CAT bonds whose payoff function is contingent on occurring catastrophe events, e.g., earthquakes, floods, and wildfire. The valuation of such contracts is of great importance as it determines a fair value at which the insurer and the investor are willing to enter the deal. However, from a technical point of view, the procedure of finding the price is not straightforward in light of the existence of both insurance and financial risks during the valuation process.

2 Model assumption and valuation framework

In this section we build on our valuation framework for CAT bond pricing based on the two-step valuation approach established by Pelsser and Stadje (2014), which was developed later by Tang and Zhongyi (2019).

The probability space on which all random variables will be defined is described as follows: We model the uncertainty in the market by using a filtered probability space $(\Omega, \mathcal{F}, \underline{\mathcal{F}}, \mathbb{P})$ where the flow of information is modeled by an increasing family of sub sigma-fields of $\mathcal{F}$, $\underline{\mathcal{F}} = (\mathcal{F}_t)_{(t \geq 0)}$, satisfying all usual conditions. We also assume the existence of the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, \underline{\mathcal{F}})$. As it is postulated that financial risks and actuarial risks are independent under both measures (for the sake of more simplicity in the computation procedure), a complete configuration of the probability space can then be represented by $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$, $\mathbb{P}(\omega) = \mathbb{P}_1(\omega_1) \times \mathbb{P}_2(\omega_2)$, and $\mathbb{Q}(\omega) = \mathbb{Q}_1(\omega_1) \times \mathbb{Q}_2(\omega_2)$, for $\omega = (\omega_1, \omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2$, where the first component refers to the financial risk and the second component refers to the insurance risk. A detailed study of the exact form of above-mentioned filtrations associated with insurance and financial risks under the independence assumption can be found in Safarveisi et al. (2021).

The classical pricing formula for asset pricing mentions that under arbitrage-free assumption, there exist an equivalent martingale measure $\mathbb{Q}$ (in the sense of absolute continuity) for the reference measure $\mathbb{P}$ such that any price process $\{V(t) : 0 \leq t \leq T\}$ discounted at a risk-free rate is a martingale. Hence, the present value of a claim at time t can be written as follows:

$$V^{\mathbb{Q}}(t) = \mathbb{E}^{\mathbb{Q}}[D(t,T)V(T)|\mathcal{F}_t] \tag{2.1}$$

where $D(t,T) = exp\{\int_t^T r_s ds\}$ is called the discount factor in which r_s is a risk-free interest rate. Furthermore, $V(T)$ specifies the claim's value at time T, which is a random variable. It is well-known that a complete market leads to a unique price while an incomplete market results in an infinite number of prices, in which case $V^{\mathbb{Q}}(t)$ is supposed to be in the interval $(\inf_{\mathbb{Q}\in\mathcal{M}} V^{\mathbb{Q}}(t), \sup_{\mathbb{Q}\in\mathcal{M}} V^{\mathbb{Q}}(t))$, where $\mathcal{M}$ is the class of all possible equivalent measures under which the price process becomes a martingale. In the sequel, we assume that the claim's payoff at maturity time T represented by $V(T)$ only consists of insurance risk. As the discount factor $D(t,T)$ is defined as a function of interest rate that itself is a macroeconomics element, it can be seen as a part connected to the financial risk. Therefore, using the independence assumption, we can rewrite the relation (2.1) in this way:

$$V^{\mathbb{Q}}(t) = \mathbb{E}^{\mathbb{Q}_1}[D(t,T)|\mathcal{F}_t^1] \times \mathbb{E}^{\mathbb{Q}_2}[V(T)|\mathcal{F}_t^2] \tag{2.2}$$

where $\mathcal{F}_t^1$ and $\mathcal{F}_t^2$ are natural filtrations associated with financial and insurance risks, respectively.

3 Equivalent measure for financial risk

We suppose that the dynamics of interest rate process r_t under measure $\mathbb{P}_1$ is governed by the Cox-Ingersoll-Ross (CIR) model (Cox et al. 1985a), of which stochastic differential equation (SDE)is given by

$$dr_t^{\mathbb{P}_1} = \theta(m - r_t)dt + \sigma\sqrt{r_t}dW_t^{\mathbb{P}_1}, \tag{3.1}$$

where $\mathbb{P}_1$ is defined on measurable space $(\Omega_1, \mathcal{F}_1)$, $W_t^{\mathbb{P}_1}$ is a Brownian motion process, $\theta > 0$ is the mean-reverting force measurement (i.e., the speed of mean-reverting), $m > 0$ is a long-run interest rate mean, and $\sigma > 0$ is the volatility parameter for the interest rate. The Feller condition $2m\theta > \sigma^2$ guarantees that r_t is almost surely strictly positive (Feller 1951). Notice here that we may want to use interchangeably, for instance, $\mathbb{P}_1$ as a superscript or with a dash for a process in order to stress that the given process is described under measure $\mathbb{P}_1$.

To price the financial risk contained in relation (2.2), the dynamics of interest rate r_t under a risk-neutral measure $\mathbb{Q}_1$, namely $r_t^{\mathbb{Q}_1}$, should be derived. The mean idea is to employ the well-known Girsanov's Theorem which states that if $W_t^{\mathbb{P}_1}$ is a Brownian motion process, and we shift the process by $\int_0^t \frac{\lambda}{\sigma}\sqrt{r_s^{\mathbb{P}_1}}ds$, then the shifted process is again a Brownian motion process under risk-neutral probability measure $\mathbb{Q}_1$, where $\mathbb{P}_1$ and $\mathbb{Q}_1$ are linked together via a likelihood ratio, also known as Radon-Nikodym derivative formula, defined as below

$$\frac{d\mathbb{Q}_1}{d\mathbb{P}_1}\Big|_{\mathcal{F}_t^1} = \exp\left\{\int_0^t \frac{\lambda}{\sigma}\sqrt{r_s^{\mathbb{P}_1}}dW_s^{\mathbb{P}_1} - \frac{1}{2}\int_0^t \frac{\lambda^2}{\sigma^2} r_s^{\mathbb{P}_1} ds\right\} \tag{3.2}$$

where λ is a constant which determines the market price of risk. In the CIR model, the market price of risk has the form $\lambda_t = \frac{-\lambda\sqrt{r_t}}{\sigma}$[1], which is a popular choice used as a kernel function when applying Girsanov´Theorem. It is easy to show that the dynamics of r_t under risk-neutral measure $\mathbb{Q}_1$ is given by

$$dr_t^{\mathbb{Q}_1} = \theta^*(m^* - r_t^{\mathbb{Q}_1})dt + \sigma\sqrt{r_t^{\mathbb{Q}_1}}dW_t^{\mathbb{Q}_1} \tag{3.3}$$

with new parameters θ^* and m^*, which are defined by

$$\theta^* = \theta + \lambda,\ m^* = \frac{\theta m}{\theta + \lambda} \tag{3.4}$$

4 Equivalent measure for insurance risk

In order to price the insurance risk contained in the payoff function $V(T)$ presented in (2.2), we first need to identify the underlying risk process that captures the respective insurance risk. For this purpose, we start with typical definitions of a CAT bond's payoff function.

4.1 CAT bond's structure

In this paper, we consider a one-period framework for the CAT bond, of which conditional payments take place only once during the length of the contract, as opposed to a multi-period

[1]The market price of risk here represents the market´expectation about the rate evolution - a positive value of λ is interpreted as the situation where a significant decrease of rates may happen, while a negative value of which shows an significant increase of the rates.

framework where multiple conditional payments are made periodically up to and including the maturity time T. For more information about the latter type, readers can refer to, for example, Safarveisi et al. (2021). For the former type, we follow a stylized structure provided by Ma nd Ma (2013).

A zero-coupon CAT bond delivers a face value Z at maturity time T if the trigger condition $L_T \leq D$ happens, where D stands for a contractually defined threshold level and L_T is a loss index reflecting losses from natural catastrophes over the time interval $[0, T]$, otherwise a fraction $p \in (0, 1)$ of the face value is paid to the bondholder, that is,

$$P^1_{CAT}(T) = \begin{cases} Z & \text{if } L_T \leq D \\ qZ & \text{if } L_T > D \end{cases} \tag{4.1}$$

The other type that is in our interest is the coupon CAT bond whose payment includes the principal Z plus a coupon C at maturity time T provided that no triggering event, in that the event $L_T > D$, occurs. Upon happening the triggering event at a specific time $\tau \leq T$, the coupon-related part of the CAT's payment is fully and immediately deducted as an compensation for paying the promised insurance protection. Therefore, the final payment at maturity time T contains only the principal Z, that is,

$$P^2_{CAT}(T) = \begin{cases} Z + C & \text{if } L_T \leq D \\ Z & \text{if } L_T > D \end{cases} \tag{4.2}$$

According to the foregoing structures selected for the CAT bond contract[2], the payment is contingent on the behaviour of the index-linked catastrophe loss L_T, which is modelled by a counting process and a sequence of positive independent and identically random variables. The counting process defines the number of claims over time interval $[0, T]$, and random variables denote the losses resulting from natural disasters. Such a CAT bond contract whose trigger mechanism is characterized by the underlying loss index L_T is called a non-indemnity-type CAT bond. In many applications, the loss index L_T is chosen to be an aggregate loss process. This model-based loss index plays a fundamental role when calculating the insurance premium in actuarial science, and one can make different assumptions for modeling purposes. Moreover, from (4.1) and (4.2), it is obvious that L_T is the main information to determine the price of a CAT bond contract. In subsequence sections we discuss further the assumptions that one can make in order to model L_T.

[2] Also known as index-based binary CAT bond contract

4.2 Underlying risk process model

Throughout this subsection, the probability space under consideration is $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$. We denote by $X = \{X_n\}_{n\in\mathbb{N}}$ a sequence of $\mathbb{P}_2$-i.i.d positive real-valued random variables named claim size process, $N = \{N_t\}_{t\in\mathbb{R}_+}$ a counting process or claim number, $T = \{T_n\}_{n\in\mathbb{N}_0}$ a claim arrival time process, and $W = \{W_n\}_{n\in\mathbb{N}}$ a claim interarrival process. In fact, W is said to be inter-arrival times between consequtively arriving events, i.e., $W_n = T_n - T_{n-1}$. The counting process N is called a renewal counting process with parameter δ if the interarrival process W is P_2-i.i.d with distribution $\mathbf{K}^\delta$. In particular, when $\delta > 0$ and $\mathbf{K}^\delta = \mathbf{EXP}(\delta)$, namely Exponential distribution with parameter δ, the counting process N becomes a homogeneous Poisson process[3] with intensity δ where δ can be interpreted as the rate of arrivals. So, in a homogeneous Poisson process, W_n's are independent and exponentially distributed with constant parameter δ. The situation where parameter δ varies with time, i.e., it is a deterministic function of time, the counting process N is a non-homogeneous Poisson process for which W_n's are not independent and exponentially distributed any more. In case δ itself is a random variable, N turns to a mixed renewal counting process. This process can be seen as a special case of a stochastic renewal counting process where δ is itself a stochastic process (also called Cox process). In this paper, we restrict our work to the renewal counting process with fixed parameter δ.

In what followed in the previous section together with the above-mentioned definitions, the underlying risk process $L = \{L_t\}_{t\in\mathbb{R}_+}$ is said to be the aggregate claims process induced by (N, X), which is defined as $L_t = \sum_{n=1}^{N_t} X_n$ for any $t \geq 0$. Accordingly, if N is a renewal counting process with parameter δ, and independent of X, the aggregate claims process is called a $\mathbb{P}_2$-compound renewal process (CRP for short) specified by $\mathbf{K}^\delta$ and $\mathbb{P}_{X_1}$ (henceforth written as $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$)[4]). An example of the CRP is called compound Poisson process (CPP for short) in which the counting process N is supposed to be a Poisson process, i.e., L_t is a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$)[5].

4.3 Equivalent measure under renewal risk model

The fact that the underlying risk process L_t follows a compound renewal process induced by (N, X), implies that we need to change the distribution of (N, X) as a result of a change of measures. The new obtained pricing measure is characterized by market prices of insurance

[3] Other well-known examples are negative binomial process and general inverse Gaussian process, of which $\mathbf{K}^\delta$ is distributed as Gamma and general inverse Gaussian, respectively.

[4] For simplicity in notation, we use $\mathbb{P}^2_{X_1}$ to denote the distribution of X_1 under measure $\mathbb{P}_2$.

[5] Note here that a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$) is in fact a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta = \mathbf{EXP}(\delta)$, $\mathbb{P}^2_{X_1}$)

risks associated with the claim number and claim size, which reflect the risk averseness of an economic agent in the market. Characterization of such an equivalent measure can be conducted based on the information that L_t provides for us. Macheras and Tzaninis (2020) introduced a class of all equivalent measures such that a compound renewal process under reference measure remains a compound renewal process under its corresponding equivalent measure. More formally, let $\mathcal{F}_t^2 = \mathcal{F}_t^L$ be the natural filtration generated by random process L_t, in symbol we write $\mathcal{F}_t^L = \sigma(L_s, s \leq t)$ which means sigma-algebra generated by process L_s. Denote by $\mathbf{\Lambda}^{\rho(\delta)}$ the distribution function of interarrival process W with parameter $\rho(\delta)$ under measure $\mathbb{Q}_2$, where ρ is an arbitrary function. We already know that the distribution function of W under measure $\mathbb{P}_2$ is denoted by $\mathbf{K}^\delta$. It can be shown that the Radon-Nikodym derivative satisfying the property of preserving the structure of a renewal compound process under both measures is of the following form

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \Big[\prod_{j=1}^{N_t} h^{-1}(\gamma(X_j)) \times \frac{d\mathbb{Q}_{W_1}^2}{d\mathbb{P}_{W_1}^2}(W_j)\Big] \times \frac{1-\mathbf{\Lambda}^{\rho(\delta)}(t-T_{N_t})}{1-\mathbf{K}^\delta(t-T_{N_t})} \tag{4.3}$$

where h and γ are real-valued Borel measurable mapping from $(0,\infty)$ to $\mathbb{R}$ such that $\mathbb{E}^{\mathbb{P}_2}[h^{-1}(\gamma(X_j))] = 1$, $\mathbb{E}^{\mathbb{P}_2}[X_1^l h^{-1}(\gamma(X_j))] < \infty$ (for $l = 1, 2$), and $\frac{d\mathbb{Q}_{W_1}^2}{d\mathbb{P}_{W_1}^2}$ is the Radon-Nikodym derivative of distribution W_1 under measure $\mathbb{Q}_2$ with respect to distribution W_1 under measure $\mathbb{P}_2$. In practice, we set $h = \ln$ (natural logarithm) and $\gamma = h(f)$ with f being a Radon-Nikodym derivative of $\mathbb{Q}_{X_1}^2$ with respect to $\mathbb{P}_{X_1}^2$, in symbol we write $f = \frac{d\mathbb{Q}_{X_1}^2}{d\mathbb{P}_{X_1}^2}$ where $\mathbb{Q}_{X_1}^2$ and $\mathbb{P}_{X_1}^2$ are assumed to be equivalent measures. We provide a rough proof of (4.3) inspired by Macheras and Tzaninis (2020).

Proof. Assume that $\mathcal{F}^W = \{\mathcal{F}_n^W\}_{n\in\mathbb{N}_0}$ and $\mathcal{F}^X = \{\mathcal{F}_n^X\}_{n\in\mathbb{N}_0}$ are the natural filtration of W and X, respectively. It can be shown that the following holds true (for unexplained details, interested readers can refer to Macheras and Tzaninis (2020)):

$$\forall A \in \mathcal{F}_t^L \quad \exists B_k \in \sigma(\mathcal{F}_t^W \cup \mathcal{F}_t^X)(\text{for every} k \in \mathbb{N}_0)\ s.t\ A \cap \{N_t = k\} = B_k \cap \{N_t = k\} \tag{4.4}$$

Using (4.4) we yield

$$\mathbb{Q}_2(A) = \sum_{k=0}^{\infty} \mathbb{Q}_2\big(B_k \cap \{N_t = k\}\big) = \sum_{k=0}^{\infty} \mathbb{Q}_2\big(B_k \cap \{T_k \leq t\} \cap \{W_{k+1} > t - T_k\}\big) \tag{4.5}$$

For a fixed but arbitrary $n \in \mathbb{N}_0$, we define

$$G_n = \bigcap_{j=1}^{n} (W_j^{-1}(E_j) \cap X_j^{-1}(F_j)) \cap \{W_{n+1} > t - T_n\} \tag{4.6}$$

where E_j, $F_j \in \mathcal{B}(\mathbb{R}_+)$ for any $j \in \{1, 2, \cdots, n\}$. We then have that

$$\begin{aligned}
\mathbb{Q}_2(G_n) &= \mathbb{E}^{\mathbb{Q}_2}\left[I_{E_1}(W_1) I_{F_1}(X_1) \cdots I_{E_n}(W_n) I_{F_n}(X_n) I_{\{W_{n+1} > t - T_n\}}\right] \\
&= \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{Q}_2}[I_{E_j}(W_j)]\right] \times \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{Q}_2}[I_{F_j}(X_j)]\right] \times \mathbb{E}^{\mathbb{Q}_2}[I_{\{W_{n+1} > t - T_n\}}] \\
&= \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{P}_2}\left[\frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}(W_j) I_{E_j}(W_j)\right]\right] \times \left[\prod_{j=1}^{n} \mathbb{E}^{\mathbb{P}_2}\left[\frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}(X_j) I_{F_j}(X_j)\right]\right] \\
&\quad \times \frac{\mathbb{E}^{\mathbb{Q}_2}[I_{\{W_{n+1} > t - T_n\}}]}{\mathbb{E}^{\mathbb{P}_2}[I_{\{W_{n+1} > t - T_n\}}]} \times \mathbb{E}^{\mathbb{P}_2}[I_{\{W_{n+1} > t - T_n\}}] \\
&= \mathbb{E}^{\mathbb{P}_2}\left[I_{G_n} \prod_{j=1}^{n} \left(\frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}(W_j) \frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}(X_j)\right) \frac{\mathbb{E}^{\mathbb{Q}_2}[I_{\{W_{n+1} > t - T_n\}}]}{\mathbb{E}^{\mathbb{P}_2}[I_{\{W_{n+1} > t - T_n\}}]}\right]
\end{aligned} \tag{4.7}$$

A monotone class argument allows one to employ (4.7) in (4.5) for any $k \in \mathbb{N}_0$, which leads us to derive the following Radon-Nikodym derivative.

$$\left.\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\right|_{\mathcal{F}_t^2} = \left[\prod_{j=1}^{N_t} \left(\frac{d\mathbb{Q}^2_{W_1}}{d\mathbb{P}^2_{W_1}}(W_j) \frac{d\mathbb{Q}^2_{X_1}}{d\mathbb{P}^2_{X_1}}(X_j)\right)\right] \times \frac{\mathbb{Q}_2(W_{n+1} > t - T_{N_t})}{\mathbb{P}_2(W_{n+1} > t - T_{N_t})} \tag{4.8}$$

■

In the following examples, we explain more on the derivation of Radon-Nikodym derivative using relation (4.3).

Example 4.3.1 *Let $\mathbb{P}_2$ and $\mathbb{Q}_2$ be probability measures such that $\{L_t\}_{t \in \mathbb{R}_+}$ is a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$). This implies that the counting process N is a Poisson process, and hence the interarrival process W, which is independent of N, contains a sequence of independent and exponentially distributed random variables under both measures. Therefore, we can write*

$$\begin{aligned}
\left.\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\right|_{\mathcal{F}_t^2} &= \left[\prod_{j=1}^{N_t} e^{\gamma(X_j)} \times \frac{\rho(\delta) e^{-\rho(\delta) W_j}}{\delta e^{-\delta W_j}}\right] \times \frac{e^{-\rho(\delta)(t - T_{N_t})}}{e^{\delta(t - T_{N_t})}} \\
&= e^{\sum_{j=1}^{N_t} \gamma(X_j)} \times \left(\frac{\rho(\delta)}{\delta}\right)^{N_t} \times e^{-(\rho(\delta) - \delta) \sum_{j=1}^{N_t} W_j} \times e^{-t(\rho(\delta) - \delta)} \times e^{T_{N_t}(\rho(\delta) - \delta)} \\
&= e^{\sum_{j=1}^{N_t} \gamma(X_j)} \times \left(\frac{\rho(\delta)}{\delta}\right)^{N_t} \times e^{-t(\rho(\delta) - \delta)}
\end{aligned} \tag{4.9}$$

where the last line is due to the fact that $T_{N_t} = \sum_{j=1}^{N_t} W_j$.

Relation (4.9) can be reformulated through the following notation which was introduced by Macheras and Tzaninis (2020): Define a real-valued Borel measurable function $\beta_\delta(x) = \gamma(x) + \alpha_\delta$ such that $\alpha_\delta = \ln \rho(\delta) + \ln \mathbb{E}^{\mathbb{P}_2}[W_1]$. Knowing the fact that W_1 is $\mathbb{P}_2$-**EXP**(δ) and putting $\rho(\delta) = \frac{e^{\alpha_\delta}}{\mathbb{E}^{\mathbb{P}_2}[W_1]}$ which itself leads to $\alpha_\delta = \ln(\frac{\rho(\delta)}{\delta})$ as $\mathbb{E}^{\mathbb{P}_2}[W_1] = 1/\delta$, relation (4.9) turns into

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \exp\left\{\sum_{j=1}^{N_t} \beta(X_j) - \delta t \mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)} - 1]\right\}, \tag{4.10}$$

Relation (4.10) is the main result that was proved by Delbaen and Haezendonck (1989). This is however not surprising since a compound Poisson process is a special case of a compound renewal process, and one can generate subclasses by means of general framework proposed by Macheras and Tzaninis (2020). We now intend to find the distribution of claim number and claim size inducing aggregate process $\{L_t\}_{t\in\mathbb{R}_+}$ under new measure $\mathbb{Q}_2$. First, $\mathbb{E}^{\mathbb{P}_2}[\exp\{\gamma(X_1)\}] = 1$ together with $\beta_\delta(x) - \alpha_\delta = \gamma(x)$ yield

$$\mathbb{E}^{\mathbb{P}_2}[\exp\{\beta(X_1)\}] = \exp\{\alpha_\delta\} = \frac{\rho(\delta)}{\delta} \tag{4.11}$$

and so we have that

$$\rho(\delta) = \delta \mathbb{E}^{\mathbb{P}_2}[\exp\{\beta(X_1)\}] \tag{4.12}$$

Second, for all $A \in \mathcal{B}(\mathbb{R}_+)$ where $\mathcal{B}(\mathbb{R}_+)$ is defined to be the Borel sets on $\mathbb{R}_+$, we can write

$$Q_{X_1}^2(A) = \mathbb{E}^{\mathbb{P}_2}\left[I_A \frac{d\mathbb{Q}_{X_1}^2}{d\mathbb{P}_{X_1}^2}\right] = \mathbb{E}^{\mathbb{P}_2}\left[I_A e^{\gamma(X_1)}\right] = \mathbb{E}^{\mathbb{P}_2}\left[I_A \frac{e^{\beta(X_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]}\right] = \int_A \frac{e^{\beta(x_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]} d\mathbb{P}_{X_1}^2 \tag{4.13}$$

From (4.12) and (4.13), we conclude that process L_t is a $\mathbb{Q}_2$-CPP$\big(\delta \mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}], \frac{e^{\beta(x_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]}\mathbb{P}_{X_1}^2\big)$. As proposed by Muermann (2003), an alternative representation of relations (4.10) and (4.13) can be expressed in this way: Define $\kappa = \mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]$ and $\nu(x_1) = \frac{e^{\beta(x_1)}}{\mathbb{E}^{\mathbb{P}_2}[e^{\beta(X_1)}]}$. Then, we have that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} = \exp\left\{\sum_{j=1}^{N_t} \ln(\kappa\nu(X_j)) + \delta t(1-\kappa)\right\}, \tag{4.14}$$

where κ and $v(.)$ can be interpreted as the market prices of claim number risk and claim size risk, respectively. Hence, the characterization of Radon-Nikody derivative (4.14) yields that under measure $\mathbb{P}_2$, L_t is a $\mathbb{P}_2$-CPP(δ, $\mathbb{P}_{X_1}^2$) while under measure $\mathbb{Q}_2$ is a $\mathbb{Q}_2$-CPP$\big(\delta\kappa, \nu(x_1)\mathbb{P}_{X_1}^2\big)$.

Example 4.3.2 *Let* $\mathbb{P}_2$ *and* $\mathbb{Q}_2$ *be probability measures such that* $\{L_t\}_{t\in\mathbb{R}_+}$ *is a* $\mathbb{P}_2$-*CRP(*$\mathbf{K}^\delta$*,*

$\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \boldsymbol{Ga}(\delta)$[6], where $\delta = (\eta_1, \eta_2)$ and η_1 is assumed to be a positive integer. Relation (4.3) gives us that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}^2_t} = e^{\sum_{j=1}^{N_t} \gamma(X_j)} \times \left(\frac{\rho(\delta)\Gamma(\eta_1)}{\eta_2^{\eta_1}}\right)^{N_t} \times \left(\prod_{j=1}^{N_t} \frac{1}{W_j^{\eta_1 - 1}}\right) \times \frac{e^{-t(\rho(\delta)-\eta_2)}}{\sum_{i=0}^{\eta_1 - 1} \frac{(\eta_2(t-T_{N_t}))^i}{i!}} \tag{4.15}$$

Analogously to the previous example, one can show that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}^2_t} = \exp\left\{ \sum_{j=1}^{N_t} ln(\kappa\nu(X_j)) + N_t ln\left(\frac{\eta_2\Gamma(\eta_1)}{\eta_2^{\eta_1}\eta_1}\right) - (\eta_1 - 1)\sum_{j=1}^{N_t} ln(W_j) + \frac{t\eta_2}{\eta_1}(\eta_1 - \kappa) - ln\left(\sum_{i=0}^{\eta_1 - 1} \frac{\left(\eta_2(t - T_{N_t})\right)^i}{i!}\right) \right\} \tag{4.16}$$

where κ and $\nu(.)$ are as before. In this example, L_t is $\mathbb{Q}_2$-CPP$\left(\frac{\eta_2}{\eta_1}\kappa,\ \nu(x_1)\mathbb{P}^2_{X_1}\right)$.

Example 4.3.3 *Let $\mathbb{P}_2$ and $\mathbb{Q}_2$ be probability measures such that $\{L_t\}_{t\in\mathbb{R}_+}$ is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \boldsymbol{WE}(\delta)$[7], where $\delta = (\eta_3, \eta_4)$. Then, we have that*

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}^2_t} = e^{\sum_{j=1}^{N_t} \gamma(X_j)} \times \left(\frac{\eta_4^{\eta_3}\rho(\delta)}{\eta_3}\right)^{N_t} \times \left(\prod_{j=1}^{N_t} \frac{1}{W_j^{\eta_3 - 1}}\right) \times \left(\prod_{j=1}^{N_t} e^{-\rho(\delta)W_j + \frac{W_j^{\eta_3}}{\eta_4^{\eta_3}}}\right) \times \frac{e^{\rho(\delta)(t - T_{N_t})}}{e^{\frac{-(t-T_{N_t})^{\eta_3}}{\eta_4^{\eta_3}}}} \tag{4.17}$$

Similarly, it can be shown that

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}^2_t} = \exp\left\{ \sum_{j=1}^{N_t} ln(\kappa\nu(X_j)) + N_t ln\left(\frac{\eta_4^{\eta_3}}{\eta_4\Gamma(1+\frac{1}{\eta_3})\eta_3}\right) - (\eta_3 - 1)\sum_{j=1}^{N_t} ln(W_j) + \frac{1}{\eta_4^{\eta_3}}\sum_{j=1}^{N_t} W_j^{\eta_3} - \frac{t\kappa}{\eta_4\Gamma(1+\frac{1}{\eta_3})} + \frac{(t - T_{N_t})^{\eta_3}}{\eta_4^{\eta_3}} \right\} \tag{4.18}$$

---

[6] $\mathbf{Ga}(\delta)$ represents the distribution function of a Gamma distribution with the density given by

$$f(x) = \frac{\eta_2^{\eta_1}}{\Gamma(\eta_1)} x^{\eta_1 - 1} e^{-\eta_2 x} \quad (x \geq 0)$$

[7] $\mathbf{WE}(\delta)$ represents the distribution function of a Weibull distribution with the density given by

$$f(x) = \frac{\eta_3}{\eta_4^{\eta_3}} x^{\eta_3 - 1} e^{-(\frac{x}{\eta_4})^{\eta_3}} \quad (x \geq 0)$$

for which L_t *is* $\mathbb{Q}_2$*-CPP*$\big(\ \eta_4\Gamma(1+\frac{1}{\eta_3})\kappa,\ \nu(x_1)\mathbb{P}^2_{X_1}\big)$.

Example 4.3.4 *Let* $\mathbb{P}_2$ *and* $\mathbb{Q}_2$ *be probability measures such that* $\{L_t\}_{t\in\mathbb{R}_+}$ *is a* $\mathbb{P}_2$*-CRP(*$\mathbf{K}^\delta$*,* $\mathbb{P}^2_{X_1}$*) and* $\mathbb{Q}_2$*-CPP(*$\rho(\delta)$*,* $\mathbb{Q}^2_{X_1}$*) with* $\mathbf{K}^\delta$ = ***MIX-EXP***(δ)[8]*, where* $\delta = (\phi, \eta_5, \eta_6)$*. Then, we have that*

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}^2_t} = \exp\left\{\sum_{j=1}^{N_t}\gamma(X_j)\right\}\times\left(\prod_{j=1}^{N_t}\frac{\rho(\delta)e^{-\rho(\delta)W_j}}{\phi\eta_5e^{-\eta_5W_j}+(1-\phi)\eta_6e^{-\eta_6W_j}}\right)\times\frac{e^{\rho(\delta)(t-T_{N_t})}}{\Delta_1} \tag{4.19}$$

where $\Delta_1 = 1-[\phi(1-e^{-\eta_5(t-T_{N_t})})+(1-\phi)(1-e^{-\eta_6(t-T_{N_t})})]$*. Following the same procedure discussed earlier, we have that*

$$\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}^2_t} = \exp\left\{\sum_{j=1}^{N_t}ln(\kappa\nu(X_j)) - N_t ln\left(\frac{\phi}{\eta_5}+\frac{(1-\phi)}{\eta_6}\right) - \frac{\kappa t}{\frac{\phi}{\eta_5}+\frac{(1-\phi)}{\eta_6}} - ln(\Delta_1)\right\} \tag{4.20}$$

for which L_t *is* $\mathbb{Q}_2$*-CPP*$\big(\big(\frac{\phi}{\eta_5}+\frac{(1-\phi)}{\eta_6}\big)^{-1}\kappa,\ \nu(x_1)\mathbb{P}^2_{X_1}\big)$.

In the examples mentioned so far, we have considered a CPP for L_t under measure $\mathbb{Q}_2$, in which case we observed how, with the help of notations defined by Macheras and Tzaninis (2020), the Radon-Nikodym derivative could be represented in terms of parameters contained in the distributions that were assumed under the real-world measure. In fact, when L_t is a CPP under $\mathbb{Q}_2$, by means of these notations, it is possible to find an alternative representation based on the real-world parameters for the intensity parameter contained in the Poisson process under measure $\mathbb{Q}$, which itself helps to express the Radon-Nikodym solidly based on the real-world parameters. However, finding such an representation appears to be more complicate when we assume a CRP for L_t under measure $\mathbb{Q}_2$. One reason is that the definition of α_δ can be only used in case of one-dimensional space for parameter $\rho(\delta)$. In the next example, we will show this hurdle.

Example 4.3.5 *Let* $\mathbb{P}_2$ *and* $\mathbb{Q}_2$ *be probability measures such that* $\{L_t\}_{t\in\mathbb{R}_+}$ *is a* $\mathbb{P}_2$*-CRP(*$\mathbf{K}^\delta$*,* $\mathbb{P}^2_{X_1}$*) and* $\mathbb{Q}_2$*-CRP(*$\mathbf{\Lambda}^{\rho(\delta)}$*,* $\mathbb{Q}^2_{X_1}$*) with* $\mathbf{K}^\delta$ = ***Ga***(δ) *and* $\mathbf{\Lambda}^{\rho(\delta)}$ = ***Ga***$(\rho(\delta))$*, where* $\delta = (\xi_1, \xi_2)$

[8]**MIX-EXP**(δ) represents the distribution function of a mixture of two exponential distributions with the density given by

$$f(x) = \phi\eta_5e^{-\eta_5x}+(1-\phi)\eta_6e^{-\eta_6x} \quad (x\geq 0),$$

with $\phi\in(0,1)$.

and $\rho(\delta) = (\epsilon_1, \epsilon_2)$. Applying relation (4.3), we get that

$$\begin{aligned}
\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_t^2} &= \left[\prod_{j=1}^{N_t} e^{\gamma(X_j)} \times \frac{\frac{\epsilon_2^{\epsilon_1}}{\Gamma(\epsilon_1)} W_j^{\epsilon_1 - 1} e^{-\epsilon_2 W_j}}{\frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)} W_j^{\xi_1 - 1} e^{-\xi_2 W_j}}\right] \times \frac{\sum_{i=0}^{\epsilon_1 - 1} \frac{(\epsilon_2 (t - T_{N_t}))^i}{i!} e^{-\epsilon_2 (t - T_{N_t})}}{\sum_{i=0}^{\xi_1 - 1} \frac{(\xi_2 (t - T_{N_t}))^i}{i!} e^{-\xi_2 (t - T_{N_t})}} \\
&= \exp\left\{\sum_{j=1}^{N_t} \gamma(X_j)\right\} \times \left(\frac{\epsilon_2^{\epsilon_1}\Gamma(\xi_1)}{\xi_2^{\xi_1}\Gamma(\epsilon_1)}\right)^{N_t} \times \left[\prod_{j=1}^{N_t} W_j^{\epsilon_1 - \xi_1}\right] \times \frac{\sum_{i=0}^{\epsilon_1 - 1} \frac{(\epsilon_2 (t - T_{N_t}))^i}{i!}}{\sum_{i=0}^{\xi_1 - 1} \frac{(\xi_2 (t - T_{N_t}))^i}{i!}}
\end{aligned} \tag{4.21}$$

where ξ_1 and ϵ_1 are said to be positive integers.

In this example, we observe that $\rho(\delta)$ is of two-dimension, which does not allow us to find an alternative representation for the Radon-Nikodym derivative in terms of only real-world parameters.

5 CAT bond price

In this section we aim to adopt the actuarially consistent valuation principle (ACVP) introduced by Muermann (2003) to price catastrophe derivatives. To discuss the ACVP concept, we first start with the framework of Delbaen and Haezendonck (1989) in which martingale approach is used to compute the insurance premium from a financial perspective.

Suppose that an insurance company owns a reinsurance contract whereby he is permitted to sell-off the risk $L_T - L_t$ of the remaining period $(t, T]$ for a premium that is predictable at time t and denoted by p_t. Denote by Z_t the underlying risk process which can represents either the price process of a financial asset in the insurance business, in which case it can be defined as $Z_t = L_t - p_t$, or the company's liability at time t given by $Z_t = p_t - L_t$, where the latter can be interpreted as the net earning from insurance business up to time t. The assumption that the market where such a policy is taken over is sufficiently liquid implies there should be risk-neutral probability measure $\mathbb{Q}$ under which $\{Z_t : 0 \le t \le T\}$ is a martingale. We therefore for $s \le t$ can write

$$\mathbb{E}^{\mathbb{Q}}[L_t - p_t | \mathcal{F}_s] = L_s - p_s \tag{5.1}$$

and hence,

$$\mathbb{E}^{\mathbb{Q}}[L_T - L_t | \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[L_T - p_T + p_T - L_t | \mathcal{F}_t] = L_t - p_t + p_T - L_t = p_T - p_t = p^* \tag{5.2}$$

where p^* can be thought of as the dual predictable projection of the random process $L_T - L_t$.

We use p^* as the premium definition, which is defined as

$$p^* = \mathbb{E}^{\mathbb{Q}}[L_T - L_t|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}\left[\sum_{j=1}^{N(T-t)} X_j \middle| \mathcal{F}_t\right] \tag{5.3}$$

We now consider CAT bond contracts whose pay-off function are given by (4.1) and (4.2) and that have the same underlying risk process L_t as in the reinsurance contract. According to the pricing formula in (2.2), the financial price of the zero-coupon CAT bond is given by

$$V_1^{\mathbb{Q}}(t) = \mathbb{E}^{\mathbb{Q}_1}[D(t,T)|\mathcal{F}_t^1] \times \mathbb{E}^{\mathbb{Q}_2}[P_{CAT}^1(T)|\mathcal{F}_t^2] \tag{5.4}$$

and the financial price of the coupon payment CAT bond is given by

$$V_2^{\mathbb{Q}}(t) = \mathbb{E}^{\mathbb{Q}_1}[D(t,T)|\mathcal{F}_t^1] \times \mathbb{E}^{\mathbb{Q}_2}[P_{CAT}^2(T)|\mathcal{F}_t^2] \tag{5.5}$$

where the first term in (5.4) and (5.5) denotes the zero-coupon bond price expressed by

$$P(t,T,r(t),\theta^*,m^*,\sigma) = \mathbb{E}^{\mathbb{Q}_1}[D(t,T)|\mathcal{F}_t^1] = A(t,T)e^{-B(t,T)r(t)}, \tag{5.6}$$

where $A(t,T)$ and $B(t,T)$ are given as below:

$$\begin{aligned} A(t,T) &= \left[\frac{2\vartheta e^{(\vartheta+\theta^*)(T-t)/2}}{(\vartheta+\theta^*)(e^{\vartheta(T-t)}-1)+2\vartheta}\right]^{\frac{2m^*\theta^*}{\sigma^2}}, \\ B(t,T) &= \frac{2(e^{\vartheta(T-t)}-1)}{(\vartheta+\theta^*)(e^{\vartheta(T-t)}-1)+2\vartheta}, \\ \vartheta &= \sqrt{(\theta^*)^2+2\sigma^2}. \end{aligned}$$

(see, e.g., Brigo and Mercurio **?**). Suppose that the above CAT derivatives and the reinsurance contract defined on the same underlying risk process L_t are traded simultaneously. The ACVP states that in order to preclude arbitrage-opportunity in the market, the CAT derivatives must be priced in relation to the observable premium of their corresponding insurance contracts. More precisely, the financial prices represented by (5.5) and (5.6) should be expressed in terms of premium formula in (5.4). In the next subsection, we assume different premium principles to derive CAT bond price under different model assumptions discussed in the preceding section.

5.1 Expected value principle

We first consider the case of example 4.3.1 where L_t is $\mathbb{P}_2$-CPP(δ, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP$\big(\delta\kappa$, $\nu(x_1)\mathbb{P}^2_{X_1}\big)$. The expected value principle is constructed by taking $\beta(x) = \alpha$, where α is a constant. Then, the insurance premium in (5.3) becomes

$$\begin{aligned} p^*_{1,1} = \mathbb{E}^{\mathbb{Q}}\left[\sum_{j=1}^{N(T-t)} X_j \Big| \mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{Q}_2}[N_1]\mathbb{E}^{\mathbb{Q}_2}[X_1] \times (T-t) = \delta\kappa\mathbb{E}^{\mathbb{P}_2}[X_1.\nu(X_1)] \times (T-t) \\ &= e^{\alpha}\delta\mathbb{E}^{\mathbb{P}_2}[X_1] \times (T-t) \end{aligned} \tag{5.7}$$

In example 4.3.2, it is assumed L_t is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \mathbf{Ga}(\delta)$. Therefore,

$$p^*_{1,2} = \mathbb{E}^{\mathbb{Q}_2}[N_1]\mathbb{E}^{\mathbb{Q}_2}[X_1] \times (T-t) = e^{\alpha}\frac{\eta_2}{\eta_1}\mathbb{E}^{\mathbb{P}_2}[X_1] \times (T-t) \tag{5.8}$$

In example 4.3.3, L_t is considered to be a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \mathbf{WE}(\delta)$. Hence,

$$p^*_{1,3} = e^{\alpha}\eta_4\Gamma(1+\frac{1}{\eta_3})\mathbb{E}^{\mathbb{P}_2}[X_1] \times (T-t) \tag{5.9}$$

In example 4.3.4, L_t is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CPP($\rho(\delta)$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \mathbf{MIX\text{-}EXP}(\delta)$, for which the insurance premium is given by

$$p^*_{1,4} = e^{\alpha}\big(\frac{\phi}{\eta_5} + \frac{(1-\phi)}{\eta_6}\big)^{-1}\mathbb{E}^{\mathbb{P}_2}[X_1] \times (T-t) \tag{5.10}$$

We see that under the expected value principle, the expected value of claim size remains unchanged.

5.2 Variance principle

For the variance principle, we choose $\beta(x) = ln(a+bx)$ with $b > 0$ and $a = 1 - b\mathbb{E}^{\mathbb{P}_2}[X_1] > 0$. Then, from examples 4.3.1, 4.3.2, 4.3.3, and 4.3.4, we respectively deduce the following insurance premiums:

$$
\begin{aligned}
p^*_{2,1} &= \delta(\mathbb{E}^{\mathbb{P}_2}[X_1] + b\mathbb{V}ar^{\mathbb{P}_2}[X_1]) \times (T-t) & (5.11)\\
p^*_{2,2} &= \frac{\eta_2}{\eta_1}(\mathbb{E}^{\mathbb{P}_2}[X_1] + b\mathbb{V}ar^{\mathbb{P}_2}[X_1]) \times (T-t) & (5.12)\\
p^*_{2,3} &= \eta_4\Gamma(1+\frac{1}{\eta_3})(\mathbb{E}^{\mathbb{P}_2}[X_1] + b\mathbb{V}ar^{\mathbb{P}_2}[X_1]) \times (T-t) & (5.13)\\
p^*_{2,4} &= (\frac{\phi}{\eta_5} + \frac{(1-\phi)}{\eta_6})^{-1}(\mathbb{E}^{\mathbb{P}_2}[X_1] + b\mathbb{V}ar^{\mathbb{P}_2}[X_1]) \times (T-t) & (5.14)
\end{aligned}
$$

5.3 Esscher principle

Choosing $\beta(x) = cx - ln(\mathbb{E}^{\mathbb{P}_2}[e^{cX_1}])$ with $c > 0$, we obtain the Esscher transform for the expected claim size under the new measure. From 4.3.1, 4.3.2, 4.3.3, and 4.3.4, we then derive the insurance premium,

$$
\begin{aligned}
p^*_{3,1} &= \delta\frac{\mathbb{E}^{\mathbb{P}_2}[X_1\exp\{cX_1\}]}{\mathbb{E}^{\mathbb{P}_2}[\exp\{cX_1\}]} \times (T-t) & (5.15)\\
p^*_{3,2} &= \frac{\eta_2}{\eta_1}\frac{\mathbb{E}^{\mathbb{P}_2}[X_1\exp\{cX_1\}]}{\mathbb{E}^{\mathbb{P}_2}[\exp\{cX_1\}]} \times (T-t) & (5.16)\\
p^*_{3,3} &= \eta_4\Gamma(1+\frac{1}{\eta_3})\frac{\mathbb{E}^{\mathbb{P}_2}[X_1\exp\{cX_1\}]}{\mathbb{E}^{\mathbb{P}_2}[\exp\{cX_1\}]} \times (T-t) & (5.17)\\
p^*_{3,4} &= (\frac{\phi}{\eta_5} + \frac{(1-\phi)}{\eta_6})^{-1}\frac{\mathbb{E}^{\mathbb{P}_2}[X_1\exp\{cX_1\}]}{\mathbb{E}^{\mathbb{P}_2}[\exp\{cX_1\}]} \times (T-t) & (5.18)
\end{aligned}
$$

5.4 Examples of renewal process under both measures

Our next intention is to compute the insurance premium in situations where the aggregate loss process L_t is a CRP under both measures. We first consider example 4.3.5, where L_t is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CRP($\mathbf{\Lambda}^{\rho(\delta)}$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta = \mathbf{Ga}(\delta)$ and $\mathbf{\Lambda}^{\rho(\delta)} = \mathbf{Ga}(\rho(\delta))$. As we discussed earlier, it is not possible to acquire a representation based on κ and $\nu(.)$. We then proceed with an alternative way.

Recall that the distribution of renewal process N_t under measure $\mathbb{Q}_2$ is given by

$$
\begin{aligned}
\mathbb{Q}_2(N_t = n) &= \mathbb{Q}_2(\{N_t \geq n\} \cap \{N_t \geq n+1\}^c)\\
&= \mathbb{Q}_2(\{N_t \geq n\}) - \mathbb{Q}_2(\{N_t \geq n+1\})\\
&= \mathbb{Q}_2(T_n \leq t) - \mathbb{Q}_2(T_{n+1} \leq t), & (5.19)
\end{aligned}
$$

where $T_n = \sum_{i=1}^{n} W_i$ (i.e., the arrival time of n-th event) can be written as the sum of n interarrival times W_i's, each of which follows a gamma distribution with parameter $n\epsilon_1$ and ϵ_2. Similarly, T_{n+1} is a Gamma with parameters $(n+1)\epsilon_1$ and ϵ_2. Hence, we have that

$$\mathbb{Q}_2(N_t = n) = \sum_{s=n\epsilon_1}^{+\infty} \frac{e^{-\epsilon_2 t}(\epsilon_2 t)^s}{s!} - \sum_{s=(n+1)\epsilon_1}^{+\infty} \frac{e^{-\epsilon_2 t}(\epsilon_2 t)^s}{s!} = \sum_{s=n\epsilon_1}^{n\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 t}(\epsilon_2 t)^s}{s!}, \tag{5.20}$$

which means that the event that exactly n events have occurred in the renewal process may be thought of as the event that between $n\epsilon_1$ and $n\epsilon_1 + \epsilon_1 - 1$ have occurred in the underlying Poisson process with intensity parameter ϵ_2. Accordingly, in example 4.3.5, we can say that the problem of renewal process can be turned into the problem of Poisson process where the counting process N_t takes its values between $n\epsilon_1$ and $n\epsilon_1 + \epsilon_1 - 1$ for $n = 0, 1, 2, \cdots$. We finally have that

$$p^*_{4,1} = \mathbb{E}^{\mathbb{Q}_2}[N_{T-t}]\mathbb{E}^{\mathbb{Q}_2}[X_1] = \left(\sum_{n=0}^{\infty} \sum_{s=n\epsilon_1}^{n\epsilon_1+\epsilon_1-1} n \frac{e^{-\epsilon_2(T-t)}(\epsilon_2(T-t))^s}{s!} \right) \mathbb{E}^{\mathbb{Q}_2}[X_1] \tag{5.21}$$

Example 5.4.1 *Another interesting case is when L_t is a $\mathbb{P}_2$-CRP($\mathbf{K}^\delta$, $\mathbb{P}^2_{X_1}$) and $\mathbb{Q}_2$-CRP($\mathbf{\Lambda}^{\rho(\delta)}$, $\mathbb{Q}^2_{X_1}$) with $\mathbf{K}^\delta$ = **MIX-EXP**(δ) and $\mathbf{\Lambda}^{\rho(\delta)}$ = **MIX-EXP**$(\rho(\delta))$, where $\delta = (\phi, \eta_5, \eta_6)$ and $\rho(\delta) = (\phi^*, \eta_7, \eta_8)$. Applying Laplace transform can easily show that the insurance premium in this case can be written as follows:*

$$p^*_{5,1} = \frac{\eta_7\eta_8(T-t)}{(1-\phi^*)\eta_7 + \phi^*\eta_8} + \frac{\phi^*(1-\phi^*)(\eta_7-\eta_8)^2}{[(1-\phi^*)\eta_7 + \phi^*\eta_8]^2} \left[1 - e^{[(1-\phi^*)\eta_7+\phi^*\eta_8](T-t)}\right] \mathbb{E}^{\mathbb{Q}_2}[X_1] \tag{5.22}$$

5.5 Actuarially consistent valuation

In this section, we use the ACVP to combine the insurance premium and the financial price of a CAT bond contract specified in the insurance/reinsurance market and the capital market, respectively. To simplify our calculation, we apply the well-known principle of a single big jump to approximate the distribution function corresponding to sum of random variables (see, e.g., Foss et al. 2007). It is stated that for sufficiently large value y, and i.i.d random variables $Y_1, Y_2, \cdots, Y_n$ with a common heavy-tailed distribution G, one can show

$$\mathbb{P}(Y_1 + Y_2 + \cdots + Y_n > y) \approx \mathbb{P}(\max\{Y_1, Y_2, \cdots, Y_n\} > y) \tag{5.23}$$

In this case, we say that the two random variables $\sum_{i=1}^{n} Y_i$ and $\max\{Y_1, Y_2, \cdots, Y_n\}$ are tail-equivalent.

According to (5.4), the price of a zero-coupon CAT bond with maturity time T, pay-off function (4.1) and model assumptions stated in example (4.3.1) is

$$V_1^{\mathbb{Q}}(t) = P(t, T, r(t), \theta^*, m^*, \sigma)\Big[Z + (Z - qZ)\mathbb{Q}_2(L_T > D)\Big] \tag{5.24}$$

where $P(t, T, r(t), \theta^*, m^*, \sigma)$ is given by (5.6), and

$$\begin{aligned}
\mathbb{Q}_2(L_T > D) &= e^{-\delta\kappa T}\sum_{k=0}^{\infty}\frac{(\delta\kappa T)^k}{k!}\mathbb{Q}_2\left(\sum_{i=1}^{k} X_i > D\right) \\
&\approx e^{-\delta\kappa T}\sum_{k=0}^{\infty}\frac{(\delta\kappa T)^k}{k!}\mathbb{Q}_2(\max\{X_1, \cdots, X_k\} > D) \\
&= e^{-\delta\kappa T}\sum_{k=0}^{\infty}\frac{(\delta\kappa T)^k}{k!}\left(1 - \mathbb{Q}_2(X_1 \le D)^k\right) \\
&= e^{-\delta\kappa T}\sum_{k=0}^{\infty}\frac{(\delta\kappa T)^k}{k!} - e^{-\delta\kappa T}\sum_{k=0}^{\infty}\frac{(\delta\kappa T)^k}{k!}\mathbb{Q}_2(X_1 \le D)^k \\
&= 1 - e^{-\delta\kappa T\mathbb{Q}_2(X_1 > D)} \qquad (5.25) \\
&= 1 - e^{-\delta\kappa T\nu(D)\mathbb{P}_2(X_1 > D)} \qquad (5.26)
\end{aligned}$$

We can then write the price in terms of market prices of interest rate, claim frequency and claim size as below:

$$V_1^{\mathbb{Q}}(t) = V_1(t, \lambda, \kappa, \nu, T) = A(t, T)e^{-B(t,T)r(t)}\left[Z + (Z - qZ)\left(1 - e^{-\delta\kappa T\nu(D)\mathbb{P}_2(X_1 > D)}\right)\right] \tag{5.27}$$

Similarly, the price of the coupon payment CAT bond with pay-off function (4.2) becomes

$$V_2^{\mathbb{Q}}(t) = V_1(t, \lambda, \kappa, \nu, T) = A(t, T)e^{-B(t,T)r(t)}\left[(Z + C) - C\left(1 - e^{-\delta\kappa T\nu(D)\mathbb{P}_2(X_1 > D)}\right)\right] \tag{5.28}$$

Denote by $V_1^A(t, T, \lambda, \kappa, \nu, p_{i,j}^*)$ and $V_2^A(t, T, \lambda, \kappa, \nu, p_{i,j}^*)$ the actuarial consistent representation of $V_1^{\mathbb{Q}}(t)$ and $V_2^{\mathbb{Q}}(t)$, which are derived according to

Expected value principle:

$$
\begin{aligned}
V_1^A(t,T,\lambda,\kappa,\nu,p^*_{1,1}) &= A(t,T)e^{-B(t,T)r(t)}\Big[Z+(Z-qZ)\Big(1-e^{-p^*_{1,1}\frac{T\mathbb{P}_2(X_1>D)}{(T-t)e^{\alpha}\mathbb{E}^{\mathbb{P}_2}[X_1]}}\Big)\Big] \\
V_2^A(t,T,\lambda,\kappa,\nu,p^*_{1,1}) &= A(t,T)e^{-B(t,T)r(t)}\Big[(Z+C)-C\Big(1-e^{-p^*_{1,1}\frac{T\mathbb{P}_2(X_1>D)}{(T-t)e^{\alpha}\mathbb{E}^{\mathbb{P}_2}[X_1]}}\Big)\Big]
\end{aligned}
\tag{5.29}
$$

Variance principle:

$$
\begin{aligned}
V_1^A(t,T,\lambda,\kappa,\nu,p^*_{1,2}) &= A(t,T)e^{-B(t,T)r(t)}\Big[Z+(Z-qZ)\Big(1-e^{-p^*_{1,2}\frac{T(1-b\mathbb{E}^{\mathbb{P}_2}[X_1]+bD)\mathbb{P}_2(X_1>D)}{(T-t)(\mathbb{E}^{\mathbb{P}_2}[X_1]+b\mathbb{V}ar^{\mathbb{P}_2}[X_1])}}\Big)\Big] \\
V_2^A(t,T,\lambda,\kappa,\nu,p^*_{1,2}) &= A(t,T)e^{-B(t,T)r(t)}\Big[(Z+C)-C\Big(1-e^{-p^*_{1,2}\frac{T(1-b\mathbb{E}^{\mathbb{P}_2}[X_1]+bD)\mathbb{P}_2(X_1>D)}{(T-t)(\mathbb{E}^{\mathbb{P}_2}[X_1]+b\mathbb{V}ar^{\mathbb{P}_2}[X_1])}}\Big)\Big]
\end{aligned}
\tag{5.30}
$$

Esscher principle:

$$
\begin{aligned}
V_1^A(t,T,\lambda,\kappa,\nu,p^*_{1,3}) &= A(t,T)e^{-B(t,T)r(t)}\Big[Z+(Z-qZ)\Big(1-e^{-p^*_{1,3}\frac{T\mathbb{E}^{\mathbb{P}_2}[e^{cX_1}]\mathbb{P}_2(X_1>D)}{(T-t)e^{cD}}}\Big)\Big] \\
V_2^A(t,T,\lambda,\kappa,\nu,p^*_{1,3}) &= A(t,T)e^{-B(t,T)r(t)}\Big[(Z+C)-C\Big(1-e^{-p^*_{1,3}\frac{T\mathbb{E}^{\mathbb{P}_2}[e^{cX_1}]\mathbb{P}_2(X_1>D)}{(T-t)e^{cD}}}\Big)\Big]
\end{aligned}
\tag{5.31}
$$

As it is straightforward to find the CAT bond price formula for examples 4.3.2, 4.3.3, and 4.3.4, we do not discuss further here. But to be able to find their corresponding prices in practice, we first need to estimate physical parameters of postulated models (i.e., claim size, claim frequency, inter-arrival time) based on historical data. The next step is to extract (α, c, d) from insurance premium formulas by having a knowledge of insurance premium values charged by the insurance company. The market price of interest rate λ however should be extracted from the Capital market. Plugging all information gathered from historical data, insurance market, and capital market gives us the price we look for. We now take our attention to examples 4.3.5 and 5.4.1 where L_t is assumed to be CRP under both measures. For such a situation, the CAT bond price will also rely on risk-neutral parameters contained in inter-arrival time distribution, which can not be estimated easily from information available in the capital market. In this paper, we find the estimation of inter-arrival parameters under risk-neutral model by using Bayesian framework.

5.5.1 ACVP under Bayesian framework

The foundation of Bayesian approach considers the specification of a sampling model from which our data is drawn and a marginal distribution of unknown parameters contained in the target distribution, called a prior distribution. Using Bayes'rule, the information about unknown parameters is updated through the conditional model of observed data $\mathbf{w}$ and the prior model of unknown parameters to achieve a new distribution called a posteriori. This transition from a prior distribution to a posterior distribution can be seen as the situation where we move from the physical measure to a risk-neutral measure. Then, the inverse probability of unknown parameters under physical measure (i.e., posterior distribution) can be employed to find a reasonable estimation for corresponding risk-neutral parameters. The Bayesian estimation is one that minimize the posteriori expected of a given loss function. The commonly used loss functions are squared error loss function, absolute value error loss function, and weighted squared error loss function. Here, we consider a loss function based on the Kullback-Leibler divergence (KLD). By its standard definition, KLD defines the relative distance between two absolutely continuous measures $\mathbb{P}$ and $\mathbb{Q}$ in the entropy sense[9]. Let $f(x)$ be a density function for a continuous random variable X, characterized by parameter Θ. From Bayesian perspective, a necessary form for a proper loss function under KLD is then given by

$$\text{KL}(\Theta \parallel \hat{\Theta}) = \text{KL}\big(f(x;\Theta) \parallel f(x;\hat{\Theta})\big) = \int_{\mathcal{A}} \log \frac{f(x;\Theta)}{f(x;\hat{\Theta})} f(x;\Theta)dx, \tag{5.32}$$

where we call $\text{KL}(\Theta \parallel \hat{\Theta})$ as the Kullback error loss function (KEL) corresponding to the density function f. The Bayesian estimation is then defined as the estimator Θ^B that minimizes KLD between the actual parameter of interest Θ and its possible estimation $\hat{\Theta}$, that is,

$$\Theta^B = \arg\min_{\hat{\Theta}} \mathbb{E}^{\mathbb{P}}[\text{KL}(\Theta \parallel \hat{\Theta}) \,|\, \mathbf{x}] \tag{5.33}$$

We consider the problem of example 4.3.5 in which under physical measure the inter-arrival time distribution $\mathbf{W}$ is Gamma distribution with shape and rate parameters ξ_1 and ξ_2, while under risk-neutral measure is Gamma distribution with shape and rate parameters ϵ_1 and ϵ_2. Let $\mathbf{W}_m = (W_1, W_2, \cdots, W_m)$ be a complete sample from $\mathbf{W}$. We specify the likelihood and prior of the model as follows:

$$\begin{aligned} \mathbf{W}_m|\xi_1,\xi_2 &\sim \text{Ga}(\xi_1,\xi_2), \quad \xi_1 > 0, \quad \xi_2 > 0, \\ (\xi_1,\xi_2) &\sim \pi(\delta) \end{aligned}$$

[9]In an incomplete market where the class of all possible equivalent measures is not unique, a desired measure is selected to minimize KLD, called the minimal entropy martingale measure.

where

$$f(\mathbf{w}_m|\xi_1,\xi_2) = \frac{\xi_2^{m\xi_1}}{\Gamma^m(\xi_1)}\Big(\prod_{i=1}^{m} w_i^{\xi_1-1}\Big)\exp\Big\{-\xi_2\sum_{i=1}^{m} w_i\Big\} \tag{5.34}$$

$$\pi(\delta) \propto \frac{\xi_2}{\Gamma(\xi_1)}\exp\left\{(\xi_1-1)\frac{\psi(\xi_1)}{\Gamma(\xi_1)}-\xi_1\right\} \tag{5.35}$$

with $\psi(\xi_1) = \frac{\partial}{\partial\xi_1} log\Gamma(\xi_1) = \frac{\Gamma'(\xi_1)}{\Gamma(\xi_1)}$ being the diGamma function. The selected non-informative prior distribution above is derived by the justified maximal data information prior (JMDIP) which maximizes the prior average information in the data density minus the information in the prior density. A simulation study conducted by Moala et al. (2013) has shown that for Gamma distribution JMDIP outperforms in a class of non-informative priors such as Jeffrey's Prior, Reference Prior, and Tibishirani's Prior. Assuming the above likelihood and prior distribution, the joint posterior distribution for parameters ξ_1 and ξ_2 is given by,

$$\begin{aligned} f(\xi_1,\xi_2|\mathbf{w}_m) &\propto f(\mathbf{w}_m|\xi_1,\xi_2)\times\pi(\delta) \\ &\propto \frac{\xi_2^{m\xi_1}}{\Gamma^m(\xi_1)}\Big(\prod_{i=1}^{m} w_i^{\xi_1-1}\Big)\exp\Big\{-\xi_2\sum_{i=1}^{m} w_i\Big\}\times\frac{\xi_2}{\Gamma(\xi_1)}\exp\left\{(\xi_1-1)\frac{\psi(\xi_1)}{\Gamma(\xi_1)}-\xi_1\right\} \end{aligned} \tag{5.36}$$

Using relation (5.32), the KEL corresponding to $\text{Ga}(\xi_1,\xi_2)$ with density $f(x;\xi_1,\xi_2)$ is given by:

$$\begin{aligned} \text{KL}(\Theta \parallel \hat{\Theta}) &= \int_0^\infty \log\frac{f(x;\xi_1,\xi_2)}{f(x;\hat{\xi_1},\hat{\xi_2})} f(x;\xi_1,\xi_2)dx, \\ &= \log\frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)} - \log\frac{\hat{\xi_2}^{\hat{\xi_1}}}{\Gamma(\hat{\xi_1})} + (\hat{\xi_2}-\xi_2)\frac{\xi_1}{\xi_2} - (\hat{\xi_1}-\xi_1)\big(\psi(\xi_1)-\log\xi_2\big) \end{aligned} \tag{5.37}$$

According to relation (5.33), the Bayes estimate of ξ_1 and ξ_2 can be found by differentiating the following equation with respect to ξ_1 and ξ_2 and setting equal to zero,

$$\int_0^\infty\int_0^\infty \left[\log\frac{\xi_2^{\xi_1}}{\Gamma(\xi_1)} - \log\frac{\hat{\xi_2}^{\hat{\xi_1}}}{\Gamma(\hat{\xi_1})} + (\hat{\xi_2}-\xi_2)\frac{\xi_1}{\xi_2} - (\hat{\xi_1}-\xi_1)\big(\psi(\xi_1)-\log\xi_2\big)\right] f(\xi_1,\xi_2|\mathbf{w}_m)d\xi_1 d\xi_2, \tag{5.38}$$

which results in

$$\hat{\xi_2} = \frac{\hat{\xi_1}}{\mathbb{E}^{\mathbb{P}}[\frac{\xi_1}{\xi_2}|\mathbf{w}]}, \quad \psi(\hat{\xi_1}) = \mathbb{E}^{\mathbb{P}}\big[\psi(\xi_1) - \log \xi_2 |\mathbf{w}\big] + \log \hat{\xi_2}. \tag{5.39}$$

A Markov Chain Monte Carlo (MCMC) simulation according to Gibbs sampling scheme can be used to generate samples from the posterior distribution, but as the exact analytical form of posterior joint and marginal distributions are not available, it is more convenient to apply Metropolis-Hastings (MH) algorithm as prescribed by Moala et al. (2013). To generate samples from ξ_1 and ξ_2, we run Algorithm 1 where $f(\xi_1^{(t)}, \xi_2^{(t-1)}|\mathbf{w})$ is given by (5.36), and $\mathbf{Ga}(\frac{\xi_1^{(t-1)}}{c}, c)$ and $\mathbf{Ga}(\frac{\xi_2^{(t-1)}}{d}, d)$ are proposal distribution of which hype-parameters c and d are selected such that a good mixing of the chains and the convergence of the MCMC samples of parameters are obtained. The generated sample hereby is applied to find Bayes estimators of ξ_1 and ξ_2, denoted by ξ_1^B and ξ_2^B, through (5.39). These bayesian estimations of real-world parameters are then regarded as estimations of their corresponding risk-neutral parameters, that is, we set $\hat{\epsilon_1} = \xi_1^B$ and $\hat{\epsilon_2} = \xi_2^B$.

Algorithm 1 Metropolis-Hasting sampling for Gamma distribution

Initialize: $\xi_1^{(0)}$ and $\xi_2^{(0)}$
for $t = 1, 2, \cdots$ **do**

Generate new value $\xi_1^{(t)}$ from $\mathbf{Ga}(\frac{\xi_1^{(t-1)}}{c}, c)$, and accept it with the following probability known as the MH ratio:

$$u(\xi_1^{(t-1)}, \xi_1^{(t)}) = \min\left\{1, \frac{\mathbf{Ga}(\frac{\xi_1^{(t-1)}}{c}, c) f(\xi_1^{(t)}, \xi_2^{(t-1)}|\mathbf{w})}{\mathbf{Ga}(\frac{\xi_1^{(t)}}{c}, c) f(\xi_1^{(t-1)}, \xi_2^{(t-1)}|\mathbf{w})}\right\}$$

Generate new value $\xi_2^{(t)}$ from $\mathbf{Ga}(\frac{\xi_2^{(t-1)}}{d}, d)$, and accept it with the following probability:

$$u(\xi_2^{(t-1)}, \xi_2^{(t)}) = \min\left\{1, \frac{\mathbf{Ga}(\frac{\xi_2^{(t-1)}}{d}, d) f(\xi_1^{(t)}, \xi_2^{(t)}|\mathbf{w})}{\mathbf{Ga}(\frac{\xi_2^{(t)}}{d}, d) f(\xi_1^{(t)}, \xi_2^{(t-1)}|\mathbf{w})}\right\}$$

end for

Let us now consider the problem of example 5.4.1 in which under physical measure the inter-arrival time distribution $\mathbf{W}$ is a mixture of two exponential distributions characterized by parameters (ϕ, η_5, η_6), while under risk-neutral measure is a mixture of two exponential distributions characterized by parameters (ϕ^*, η_7, η_8). There are two ways to specify the likelihood function of a mixture model. To illustrate this clearly, we start with the basic definition. Generally, a random variable Y with a g-component mixture density, can be generated in this way. Suppose that $\boldsymbol{Z}$ be the set of g latent variables where the k-th element of $\boldsymbol{Z}$ is defined to be zero or one, according to whether the component of origin of Y in the mixture is equal to k or not. More precisely, consider $\boldsymbol{Z}_j$ as the latent variable vector for the j-th sample point. Therefore, $\boldsymbol{Z}_j$ can be a vector like this:

$$\boldsymbol{Z}_j = (Z_{1j}, Z_{2j}, \cdots, Z_{gj}) \quad \text{or} \quad \boldsymbol{Z}_j = (Z_{kj}); \quad k = 1, 2, \cdots, g \tag{5.40}$$

where $Z_{kj} = 1$ if the j-th sample point belongs to the k-th component (or cluster), otherwise $Z_{kj} = 0$. We also know that each element of vector $\boldsymbol{Z}_j$ occurs independently with probabilities $\pi_1, \pi_2, \cdots, \pi_g$. So, it is easy to see that $\boldsymbol{Z}_j$ is distributed according to a multinomial distribution consisting of one draw on g categories with probabilities $\pi_1, \pi_2, \cdots, \pi_g$, i.e., $\boldsymbol{Z}_j \sim Multi(1, \boldsymbol{\pi})$, where $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_g)$. The probability density function is then given by:

$$p(\boldsymbol{z}_j) = \mathbb{P}(\boldsymbol{Z}_j = \boldsymbol{z}_j) = \mathbb{P}(Z_{1j} = 1)^{z_{1j}} \times \mathbb{P}(Z_{2j} = 1)^{z_{2j}} \times \cdots \times \mathbb{P}(Z_{gj} = 1)^{z_{gj}} = \prod_{k=1}^{g} \pi_k^{z_{kj}} \tag{5.41}$$

For example, if we know that the j-th sample point belongs to the the second component, then for the vector $\boldsymbol{Z}_j$ we observe the vector $\boldsymbol{z}_j = (z_{1j}, z_{2j}, \cdots, z_{jg}) = (0, 1, 0, \cdots, 0)$ with the probability function:

$$p(\boldsymbol{z}_j) = \mathbb{P}\big(\boldsymbol{Z}_j = (0, 1, 0, \cdots, 0)\big) = \pi_2 \tag{5.42}$$

Let $\boldsymbol{z}_j = k$ denotes the fact that the j-th sample point fall in the k-th component, i.e., the k-th element of vector $\boldsymbol{z}_j$ is equal to one and the others are equal to zero. Based on this definition, we make another assumption which says that the conditional density of y_j given $\boldsymbol{z}_j = k$ is $f_k(y_j)$. Then, following the same logic used for the $p(\boldsymbol{z}_j)$, we can conclude that

$$f(y_j|\boldsymbol{z}_j) = \prod_{k=1}^{g} f_k(y_j)^{z_{kj}} \tag{5.43}$$

In order to obtain the mixture model of $f(y)$, we just need to apply the Bayes rule by summing up the terms on $\boldsymbol{z}$ to get the probability density function of $f(y)$. That is,

$$f(y) = \sum_{k=1}^{g} f(y, \boldsymbol{z}) = \sum_{k=1}^{g} f(y|\boldsymbol{z})p(\boldsymbol{z}) = \sum_{k=1}^{g} \pi_k f_k(y) \tag{5.44}$$

where $f_k(y)$ are called the component densities of the mixture model and the π_k are mixing weights that satisfy:

$$\sum_{k=1}^{g} \pi_k = 1 \quad \text{and} \quad 0 \leqslant \pi_k \leqslant 1 \quad (k = 1, 2, \cdots, g) \tag{5.45}$$

Suppose that $\mathcal{Y} = (y_1, y_2, ..., y_n)$ is the vector of n observed data points. Using (5.44), the likelihood function is then given by

$$f(\mathcal{Y}; \Theta) = \prod_{i=1}^{n} \left\{ \sum_{k=1}^{g} \pi_k f_k(y_i; \Theta_k) \right\} \tag{5.46}$$

Another useful decomposition of (5.46) is based on latent variables. We consider the vector $\mathcal{X} = (\mathcal{Y}, \mathcal{C})$ involving latent variables Z_i defined in (5.40), for which $\mathcal{C} = (Z_1, Z_2, ..., Z_n)$, where $Z_i = (Z_{ki})$ with $\mathbb{P}(Z_{ki} = 1) = \pi_k$, for $i = 1, 2, ..., n$ and $k = 1, 2, ..., g$. Therefore, the complete version of the likelihood function becomes

$$\begin{aligned} f(\mathcal{X}; \Theta) = f(\mathcal{Y}, \mathcal{C}; \boldsymbol{\Theta}) &= \prod_{i=1}^{n} f(y_i, Z_i; \boldsymbol{\Theta}) = \prod_{i=1}^{n} f(y_i | Z_i; \boldsymbol{\Theta}) p(Z_i) \\ &= \prod_{i=1}^{n} \left[\prod_{k=1}^{g} f_k(y_i; \Theta_k)^{Z_{ki}} \right] \times \prod_{k=1}^{g} \pi_k^{Z_{ki}} \\ &= \prod_{i=1}^{n} \prod_{k=1}^{g} \left[\pi_k f_k(y_i; \Theta_k) \right]^{Z_{ki}} \end{aligned} \tag{5.47}$$

We adopt the above-mentioned setting to construct our conditional distribution for Bayesian framework under prior distributions selected for unknown parameters contained in the mixture model. For simplicity in notation and avoidance of confusion, suppose that $\mathbf{z} = (z_1, z_2, \cdots, z_n)$ is a vector of latent variables where z_i, corresponding to the i-th sample point, takes value $k \in \{1, 2, \cdots, g\}$ with probability $p(z_i = k) = w_k$. The prior of mixing weights $\mathbf{w} = (w_1, w_2, \cdots, w_g)$ is chosen to be a standard Dirichlet prior, that is,

$$\mathbf{w} \sim \text{Dirichlet}_g(\alpha_1, \alpha_2, \cdots, \alpha_g) = \frac{\Gamma(\alpha)}{\prod_{k=1}^{g} \Gamma(\alpha_k)} \prod_{k=1}^{g} w_k^{\alpha_k - 1} \tag{5.48}$$

where $\alpha_1, \alpha_2, \cdots, \alpha_g$ are non-negative hyper-parameters, and $\alpha = \sum_{k=1}^{g} \alpha_k$. Inspired by (5.41), one can write the conditional distribution for the cluster allocations in this way:

$$p(\mathbf{z} | \mathbf{w}) = \prod_{k=1}^{g} w_k^{n_k} \tag{5.49}$$

where $n_k = \sum_{i=1}^{n} I_{\{k\}}(z_i)$ is the number of samples attributed to the k-th cluster. By marginalising (5.49) with respect to $\mathbf{w}$, we get the marginal distribution of $\mathbf{z}$,

$$p(\mathbf{z}) = \frac{\Gamma(\alpha)}{\Gamma(\alpha+n)} \prod_{k=1}^{g} \frac{\Gamma(\alpha_k + n_k)}{\Gamma(\alpha_k)} \tag{5.50}$$

Motivated by (5.47), the joint distribution for $\mathbf{y} = \mathcal{Y}$ and $\mathbf{z}$ is conveniently proportional with

$$p(\mathbf{y}, \mathbf{z}|\Theta) \propto \prod_{k=1}^{g} \left\{ \Gamma(\alpha_k + n_k) \prod_{i:z_i=k} f_k(y_i|\Theta_k) \right\} \tag{5.51}$$

It is assumed that the vector $\mathbf{w}$ is independent of $\mathbf{y}$, which can be considered reasonable in a mixture model. Therefore, from (5.48), (5.49), (5.50), we are able to find the conditional distribution $p(\mathbf{w}|\mathbf{z})$, i.e., the posterior distribution of $\mathbf{w}$,

$$p(\mathbf{w}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{w})p(\mathbf{w})}{p(\mathbf{z})} = \text{Dirichlet}_g(\alpha_1 + n_1, \alpha_2 + n_2, \cdots, \alpha_g + n_g) \tag{5.52}$$

In addition, it is possible to find the posterior distribution of Θ by associating a conjugate prior with each parameter Θ_k in (5.51). In other words,

$$p(\Theta|\mathbf{y}, \mathbf{z}) \propto \prod_{k=1}^{g} \left\{ \prod_{i:z_i=k} f_k(y_i|\Theta_k) \right\} \pi(\Theta_k) \tag{5.53}$$

We are now in a position to specify our non-parametric Bayesian framework for the mixture of two exponential distributions as follows:

$$\begin{aligned} y_i|z_i = k &\sim \mathbf{EXP}(\lambda_k) \\ \mathbf{w} = (w_1, w_2) &\sim \text{Dirichlet}_2(\alpha_1, \alpha_2) \\ \lambda_k &\sim \mathbf{Ga}(\tau_k, \Psi_k); \quad i = 1, 2, \cdots, n, \quad k = 1, 2. \end{aligned}$$

where $(\alpha_k, \tau_k, \Psi_k)$ are known hyper-parameters. The corresponding Gibbs samplers with states $t = 1, 2, \cdots$ can be written as follows:

Algorithm 2 Gibbs sampling for an exponential mixture

Initialize: $\mathbf{w}^{(0)} = (w_1^{(0)}, w_2^{(2)})$ and $\Theta^{(0)} = (\lambda_1^{(0)}, \lambda_2^{(0)})$
for $t = 1, 2, 3, \cdots$ **do**
Generate $z_i^{(t)}$ $(i = 1, \cdots, n, k = 1, 2)$ from

$$\mathbb{P}(z_i^{(t)} = k | w_k^{(t-1)}, \lambda_k^{(t-1)}, y_i) \propto w_k^{(t-1)} f(y_i | \lambda_k^{(t-1)}) = w_k^{(t-1)} \lambda_k^{(t-1)} \exp\{-\lambda_k^{(t-1)} y_i\}$$

Compute $n_k^{(t)} = \sum_{i=1}^{n} I_{\{k\}}(z_i)$ and $(s_k^y)^{(t)} = \sum_{i=1}^{n} I_{\{k\}}(z_i)\, y_i$
Generate $\mathbf{w}^{(t)}$ from Dirichlet$_2(\alpha_1 + n_1^{(t)}, \alpha_2 + n_2^{(t)})$
Generate $\lambda_k^{(t)}$ from $\mathbf{Ga}(\tau_k + n_k^{(t)}, \Psi_k + (s_k^y)^{(t)})$
end for

5.5.2 ACVP in case of RP under $\mathbb{P}_2$ and $\mathbb{Q}_2$

This subsection devotes to finding the CAT bond price according to the ACVP for examples 4.3.5 and 5.4.1. Under assumptions presented in example 4.3.5, we first need to find the distribution function of aggregate loss process L_t,

$$\begin{aligned}
\mathbb{Q}_2(L_T > D) &= \sum_{k=0}^{\infty} \mathbb{Q}_2(N_T = k) \mathbb{Q}_2\left(\sum_{i=1}^{k} X_i > D\right) \\
&= \sum_{k=0}^{\infty} \sum_{s=k\epsilon_1}^{k\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 T}(\epsilon_2 T)^s}{s!} \mathbb{Q}_2\left(\sum_{i=1}^{k} X_i > D\right) \\
&\approx \sum_{k=0}^{\infty} \sum_{s=k\epsilon_1}^{k\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 T}(\epsilon_2 T)^s}{s!} \mathbb{Q}_2(\max\{X_1, \cdots, X_k\} > D) \\
&= \sum_{k=0}^{\infty} \sum_{s=k\epsilon_1}^{k\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 T}(\epsilon_2 T)^s}{s!} \left(1 - \mathbb{Q}_2(X_1 \leq D)^k\right) \\
&= \sum_{k=0}^{\infty} \sum_{s=k\epsilon_1}^{k\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 T}(\epsilon_2 T)^s}{s!} - \sum_{k=0}^{\infty} \sum_{s=k\epsilon_1}^{k\epsilon_1+\epsilon_1-1} \frac{e^{-\epsilon_2 T}(\epsilon_2 T)^s}{s!} \mathbb{Q}_2(X_1 \leq D)^k
\end{aligned} \tag{5.54}$$

Unlike previous cases where the distribution function associated with the loss random variables X_i was transformed by factor $\nu(.)$ when considering a CPP under measure $\mathbb{Q}_2$, we do not know how to change the distribution function induced by random variable X_i when considering a CRP under measure $\mathbb{Q}_2$. To deal with this issue, a distortion function g can be selected to create a linkage between a risk-neutral distribution and its actual distribution, meaning that $\mathbb{Q}_2(X_i \leq D) = g(\mathbb{P}_2(X_i \leq D))$, more relevant information concerning a general class of distortion functions with application for CAT bond pricing can be found in Godin at al. (2019).

Note here that since the exact form of Radon-Nikodym derivative under CRP is available, we can alternatively compute (5.54) by using the change of measure technique, that is

$$\mathbb{Q}_2(L_T > D) = \mathbb{E}^{\mathbb{Q}_2}[I_{\{L_T > D\}}] = \mathbb{E}^{\mathbb{P}_2}\left[\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_T^2} I_{\{L_T > D\}}\right] \tag{5.55}$$

where $\frac{d\mathbb{Q}_2}{d\mathbb{P}_2}\Big|_{\mathcal{F}_T^2}$ is given by (4.21). Applying a simulation technique gives an approximation for relation (5.55).

In aiming to find the distribution function associated with L_t under the assumptions presented in example (5.4.1), we first obtain the probability function of claim number N_t which itself follows a counting renewal process. To this purpose, it is then essential, according to relation (5.19), to derive the distribution function associated with the n-th arrival time T_n which is sum of i.i.d random variables following a mixture of two exponential distributions **MIX-EXP**$(\rho(\delta))$. We express our problem in the following lemma with a general format.

Lemma 5.5.1 *Let random variable Y follows a mixture of two exponential distributions, that is*

$$Y \sim \pi_1 \boldsymbol{EXP}(\lambda_1) + \pi_2 \boldsymbol{EXP}(\lambda_2), \tag{5.56}$$

where π_i are mixing weights with the property that $\pi_1 + \pi_2 = 1$. Furthermore, assume that $\{Y_1, Y_2, \cdots, Y_n\}$ is a sample of i.i.d random variables drawn from $f_Y(y)$. Then, the probability density function of $Z = \sum_{j=1}^{n} Y_j$ is given by

$$f_Z(z) = \sum_{k=0}^{n} \binom{n}{k} \pi_1^k \pi_2^{n-k} \frac{\lambda_1^k \lambda_2^{n-k}}{\Gamma(n)} e^{-\lambda_1 z} z^{n-1} {}_1F(n-k, n, (\lambda_1 - \lambda_2)z) \tag{5.57}$$

with ${}_1F(a, b, c)$ being confluent hypergeometric function of the first kind defined as below:

$${}_1F(a; c; Z) = \begin{cases} \sum_{i=1}^{\infty} \frac{(a)_i}{(b)_i} \frac{c^i}{i!} & \textit{hypergeometric series representation} \\ \frac{\Gamma(b)}{\Gamma(b-1)\Gamma(a)} \int_0^1 e^{ct} t^{a-1} (1-t)^{b-a-1} dt & \textit{Integral representation} \end{cases} \tag{5.58}$$

where the Pochhammer symbol $(a)_i = \frac{\Gamma(a+i)}{\Gamma(a)} = a(a+1)\cdots(a+i-1)$.

Proof. Define random vector $\boldsymbol{\Upsilon} = (\Upsilon_1, \Upsilon_2, \cdots, \Upsilon_n)$ consisting of n i.i.d Bernoulli variable $\Upsilon_i \sim \mathbf{Ber}(\pi_1)$ which are associated with each Y_i in this way:

$$Y_i | \Upsilon_i = k \sim \mathbf{EXP}(\lambda_{2-k}); \quad \text{for} \quad k = 0, 1. \tag{5.59}$$

We define $\Upsilon^* = \sum_{i=1}^{n} \Upsilon_i \sim \mathbf{Bin}(n, \pi_1)$. Conditioning on $\boldsymbol{\Upsilon}^*$, we can write

$$Z = \sum_{j=1}^{n} Y_j | \Upsilon^* = \left(\sum_{j:\Upsilon_j=0} Y_j + \sum_{j:\Upsilon_j=1} Y_j \right) \Big| \Upsilon^*, \tag{5.60}$$

from which one can conclude that $Z|\boldsymbol{\Upsilon}^* \sim Z_1 + Z_2$, where

$$Z_1 \sim \mathbf{Ga}(\Upsilon^*, \lambda_1) \quad \text{and} \quad Z_2 \sim \mathbf{Ga}(n - \Upsilon^*, \lambda_2) \tag{5.61}$$

This shows that distribution of Z given Υ^* can be written as the sum of two independent random variables distributed as gamma with different shape and rate parameters. Denote by GDC(a, b, c, d; x) the gamma distribution convolution consisting of $\mathbf{Ga}(a, b)$ and $\mathbf{Ga}(c, d)$, which is given by:

$$\text{GDC}(a, b, c, d; x) = \begin{cases} \frac{b^a d^c}{\Gamma(a+c)} e^{-bx} x^{a+c-1} {}_1F(c, a+c, (b-d)x), & x > 0 \\ 0, & x \leq 0 \end{cases}, \tag{5.62}$$

(see, e.g., Wesolowski et al. (2016) and Di Salvo (2006)). Finally, marginalizing in Z leads to

$$\begin{aligned} f_Z(z) &= \sum_{k=0}^{n} f(z|\Upsilon^*) p(\Upsilon^* = k) \\ &= \sum_{k=0}^{n} \binom{n}{k} \pi_1^k \pi_2^{n-k} \text{GDC}(k, \lambda_1, n-k, \lambda_2; z), \end{aligned}$$

which can be viewed as a mixture of $(n+1)$ components with mixing weights equal to $\binom{n}{k} \pi_1^k \pi_2^{n-k}$.

■

Uisng (5.19) and (6.2), we derive the probability function of the renewal process N_t in example 5.4.1,

$$\begin{aligned} \mathbb{Q}_2(N_t = n) &= \mathbb{Q}_2(T_n \leq t) - \mathbb{Q}_2(T_{n+1} \leq t) \\ &= \sum_{k=0}^{n} \binom{n}{k} \pi_1^k \pi_2^{n-k} \text{CGDC}(k, \lambda_1, n-k, \lambda_2; t) \\ &\quad - \sum_{k=0}^{n+1} \binom{n+1}{k} \pi_1^k \pi_2^{(n+1)-k} \text{CGDC}(k, \lambda_1, (n+1)-k, \lambda_2; t) \end{aligned}$$

where the notation CGDC denotes the cumulative distribution function of GDC. Eventually, we have that

$$
\begin{aligned}
\mathbb{Q}_2(L_T > D) &= \sum_{n=0}^{\infty} \mathbb{Q}_2(N_T = n)\left(1 - \mathbb{Q}_2(X_1 \le D)^k\right) \\
&= \sum_{n=0}^{\infty} \left\{ \sum_{s=0}^{n} \binom{n}{s} \pi_1^s \pi_2^{n-s} \mathrm{CGDC}(s, \lambda_1, n-s, \lambda_2; T) \right. \\
&\quad \left. - \sum_{s=0}^{n+1} \binom{n+1}{s} \pi_1^s \pi_2^{(n+1)-s} \mathrm{CGDC}(s, \lambda_1, (n+1)-s, \lambda_2; T) \right\} \left(1 - \mathbb{Q}_2(X_1 \le D)^n\right) \\
&= \sum_{n=0}^{\infty} \left\{ \sum_{s=0}^{n} \binom{n}{s} \pi_1^s \pi_2^{n-s} \mathrm{CGDC}(s, \lambda_1, n-s, \lambda_2; T) \right. \\
&\quad \left. - \sum_{s=0}^{n+1} \binom{n+1}{s} \pi_1^s \pi_2^{(n+1)-s} \mathrm{CGDC}(s, \lambda_1, (n+1)-s, \lambda_2; T) \right\} \\
&\quad - \sum_{n=0}^{\infty} \mathbb{Q}_2(X_1 \le D)^n \left\{ \sum_{s=0}^{n} \binom{n}{s} \pi_1^s \pi_2^{n-s} \mathrm{CGDC}(s, \lambda_1, n-s, \lambda_2; T) \right. \\
&\quad \left. - \sum_{s=0}^{n+1} \binom{n+1}{s} \pi_1^s \pi_2^{(n+1)-s} \mathrm{CGDC}(s, \lambda_1, (n+1)-s, \lambda_2; T) \right\}
\end{aligned}
$$

where $\pi_1 = \phi^*$, $\pi_2 = 1 - \phi^*$, $\lambda_1 = \eta_7$, and $\lambda_2 = \eta_8$.

6 Numerical illustration

6.1 Parameter calibration of the CIR model

To calibrate the parameters of the CIR model (3.1), we take the daily historical yields on the 3-months US treasury bills from January 2, 1990, to May 25, 2022[10] as depicted in Figure 2.

In this section, we use maximum likelihood estimation (MLE) method for parameters of the CIR model (3.1), of which the transition probability density is given by

$$f(r_t|r_u) = c p_{\chi^2_{(d_1,d_2)}}(c r_t), \quad t > u, \tag{6.1}$$

where $c = \frac{4\theta}{\sigma^2(1-e^{-\theta(t-u)})}$ and $p_{\chi^2_{(d_1,d_2)}}(.)$ denotes the probability density of a non-central chi-square distribution with the degree of freedom $d_1 = \frac{4\theta m}{\sigma^2}$ and the non-central parameter $d_2 =$

[10]Note here that it is not essential to use the same time period for the catastrophe loss observations, as we already assumed that insurance risk and financial risks are independent.

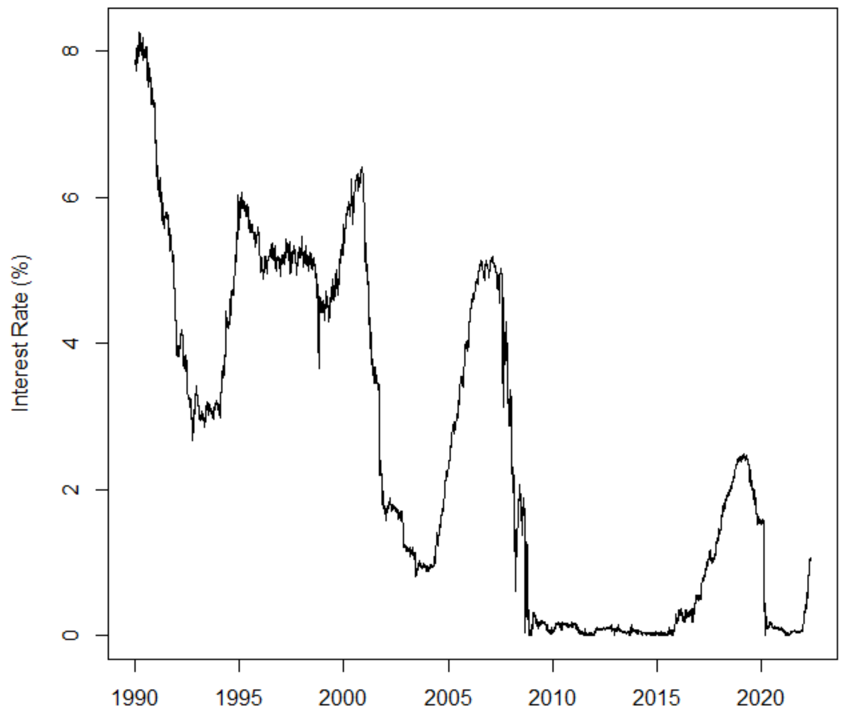

Figure 1: Historical data on interest rates observed from January 2, 1990, to May 25, 2022.

$\frac{4\theta e^{-\theta(t-u)}}{\sigma^2(1-e^{-\theta(t-u)})}r_u$ (see, e.g., Fergusson 2019). If the data $r^* = \{r_1, r_2, \cdots, r_n\}$ is given according to the historical data, then the log-likelihood function can be written as:

$$l(\theta, m, \sigma; r^*) = \sum_{i=1}^{n} ln(c) + \sum_{i=1}^{n} ln(p_{\chi^2_{(d_1,d_2)}}(cr_i|r_{i-1})) \tag{6.2}$$

In the aiming of finding the MLE of (6.2), a numerical optimization technique can be applied. For that, initial values for parameters are essential to be specified to start the iteration process embedded in the algorithm. In this paper, the initial values are obtained using ordinary least square estimation (OLSE) method. The main idea to achieve approximate estimates of parameters contained in the CIR model on the basis of OLSE method is to find a regression version that can describe a discretized version of the CIR model derived by the Euler discretization technique. According to the Euler scheme, the SDE associated with the CIR model (3.1) can be represented as follows:

$$r_{t_{i+1}} - r_{t_i} = \theta(m - r_{t_i})\Delta t_i + \sigma\sqrt{r_{t_i}}\Delta W_i \tag{6.3}$$

where $\Delta t_i = (t_{i+1} - t_i)/250$[11] and $\Delta W_i = W_{t_{i+1}} - W_{t_i}$, for $i = 0, 1, 2, \cdots, n-1$. After performing some simple algebraic manipulations, one can realize that the matrix representation of the regression model corresponding to the discretized version (6.3) is of the form:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix}}_{\boldsymbol{Y}} = \underbrace{\begin{bmatrix} \sqrt{\frac{\Delta t_1}{r_1}} & \sqrt{\Delta t_1 r_1} \\ \sqrt{\frac{\Delta t_2}{r_2}} & \sqrt{\Delta t_2 r_2} \\ \vdots & \vdots \\ \sqrt{\frac{\Delta t_{n-1}}{r_{n-1}}} & \sqrt{\Delta t_{n-1} r_{n-1}} \end{bmatrix}}_{\boldsymbol{Z}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \sigma \mathcal{N}_1(0,1) \\ \sigma \mathcal{N}_2(0,1) \\ \vdots \\ \sigma \mathcal{N}_{n-1}(0,1) \end{bmatrix}}_{\boldsymbol{\epsilon}} \tag{6.4}$$

where $y_i = \frac{r_{t_{i+1}} - r_{t_i}}{\sqrt{\Delta t_i r_{t_i}}}$, $\beta_1 = \theta m$, $\beta_2 = -\theta$, and $\mathcal{N}_i(0,1)$ are independent random variables, each having standard normal distribution. Then, the OLSE of $\boldsymbol{\beta}$ and σ become

$$\hat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{Z}^T \boldsymbol{Z})^{-1} \boldsymbol{Z}^T \boldsymbol{Y}, \quad \widehat{\sigma^2} = \frac{1}{n-2} ||\boldsymbol{Y} - \boldsymbol{Z}\hat{\boldsymbol{\beta}}_{OLS}||^2 \tag{6.5}$$

where $||.||$ denotes the Euclidean distance. We use these estimates as initial values for a numerical optimization of the likelihood function (6.2). Table 1 represents the final results for the OLS and ML estimations of the CIR parameters in (3.1).

Method (Under physical measure)	$\hat{\theta}$	$\hat{m}$	$\hat{\sigma}$
OLSE (Initial values)	0.277	0.018	0.077
MLE (Optimal values)	0.254	0.011	0.074

Table 1: Estimates of parameters of the CIR model under physical measure.

The risk-neutralized parameters represented in (3.4) can be found by knowing the estimated physical parameters as well as the constant parameter λ which determines the market price of interest rate risk. Following the same idea in Ahmad and Wilmott (2006), we find an estimation for parameter λ empirically. We summarize our steps as follows:

For simplicity in notation, suppose that $P(t, T, r_t)$ denotes the zero-coupon bond price at time t with maturity time T. In the first step, we consider the partial differential equation (PDE)

[11] Since a daily base observation are used for the interest rate, the time step of the Euler scheme is computed as the time difference between consecutive points divided by the number of working days per year, which as a standard convention is said to be 250 days.

associated with the CIR model, which is given by

$$\frac{\partial P}{\partial t}(t,T,r_t)+[\theta(m-r_t)+\lambda r_t]\frac{\partial P}{\partial r_t}(t,T,r_t)+\frac{1}{2}\sigma^2 r_t\frac{\partial^2 P}{\partial r_t^2}(t,T,r_t)-r_tP(t,T,r_t)=0 \quad (6.6)$$

In the second step, the Taylor series expansion of the zero-coupon bond price around $t = T$ with final condition $P(T, T, r_T) = 1$ is derived by the following polynomial representation:

$$P(t,T,r_t)=\sum_{j=0}^{\infty}c_j(r_t)(T-t)^j \quad (6.7)$$

Substituting (6.7) into (6.6) gives the following recursive relation for $c_j(r_t)$:

$$c_{j+1}(r_t)=\frac{1}{j+1}\left\{[\theta(m-r_t)-\lambda r_t]c_j'(r_t)+\frac{1}{2}\sigma^2 r_t c_j''(r_t)-r_tc_j(r_t)\right\}, \quad j=0,1,\cdots \quad (6.8)$$

where $c_j'(r_t)$ and $c_j''(r_t)$ are the first and second derivatives of function $c_j(r_t)$ with respect to r_t. Using (6.8), an approximation for the zero-coupon bond price is given by

$$P(t,T,r_t)\approx 1-r_t(T-t)+\frac{1}{2}\Big\{(\lambda+\theta)r_t-\theta m-r_t^2\Big\}(T-t)^2+\cdots, \quad \text{as} \quad t\longrightarrow T \quad (6.9)$$

From this we have

$$-\frac{\log(P(t,T,r_t))}{(T-t)}\approx -r_t+\frac{1}{2}\Big\{\theta m-(\lambda+\theta)r_t\Big\}(T-t)+\cdots, \quad \text{as} \quad t\longrightarrow T \quad (6.10)$$

Relation (6.10) shows that the slope of the yield curve at the short end is equal to $\frac{1}{2}\Big\{\theta m - (\lambda+\theta)r_t\Big\}$. This is useful as through observing the slope of the yield curve and short rate each day, we can compare the empirical slope with its corresponding analytical slope in order to back out a time series for λ. We report a value for parameter λ by taking an average of this generated time series. Our data set contains daily rates on US treasury bills with maturity times 3-months, 6-months, and 1-year. The 3-month maturity rate is considered to be the short rate. The 6-months and 1-year maturity rates are used for the calculation of yield curve´s slope per day. Applying the above idea, we finally found out that $\lambda = -0.033$. This value is then applied to find the risk-neutral parameters, see Table (2).

Risk neutral parameters	θ^*	m^*	σ^*
Estimation	0.221	0.013	0.074

Table 2: Estimates of parameters of the CIR model under risk-neutral measure.

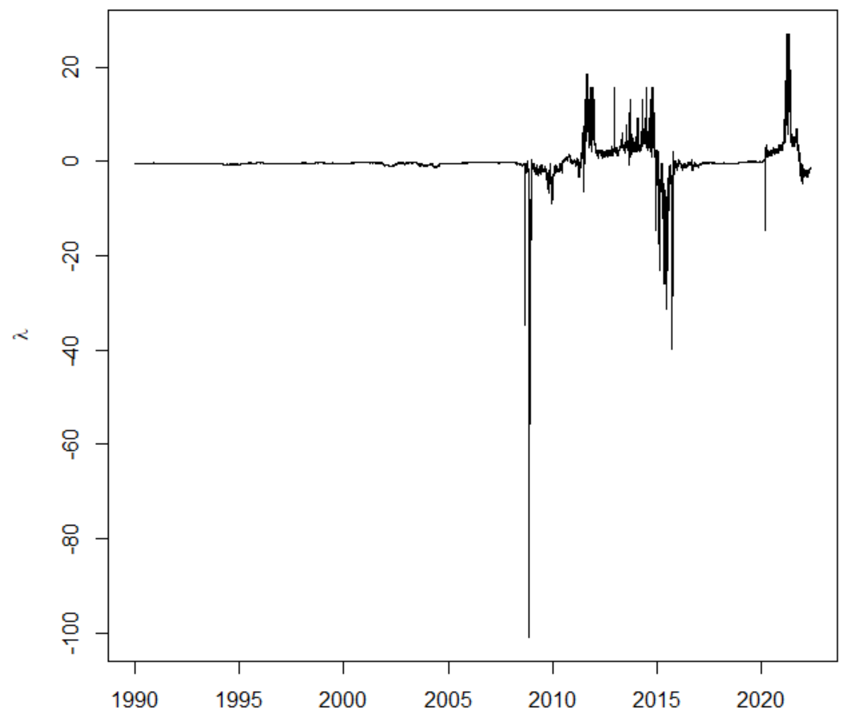

Figure 2: λ time series

References

Feller, W. 1951. Two singular diffusion problems. Ann. of Math. (2), 54:173–182. (Cited on pages 290 and 551).

Cox, J. C., J. J. E. Ingersoll, and S. A. Ross. 1985a. A theory of the term structure of interest rates. Econometrica, 385–407.

Macheras, N.D. and Tzaninis, S.M., 2020. A characterization of equivalent martingale measures in a renewal risk model with applications to premium calculation principles. Modern Stochastics: Theory and Applications, 7(1), pp.43-60

Delbaen, F. and Haezendonck, J., 1989. A martingale approach to premium calculation principles in an arbitrage free market. Insurance: Mathematics and Economics, 8(4), pp.269-277.

Muermann, A., 2003. Actuarially consistent valuation of catastrophe derivatives. The Wharton Financial Institution Center Working Paper Series, pp.03-18

Foss, S., Konstantopoulos, T. and Zachary, S., 2007. Discrete and continuous time modulated random walks with heavy-tailed increments. Journal of Theoretical Probability, 20(3), pp.581-612

ANTONIO MOALA, F.E.R.N.A.N.D.O., LUIZ RAMOS, P.E.D.R.O. and ALBERTO ACHCAR, J.O.R.G.E., 2013. Bayesian inference for two-parameter gamma distribution assuming different noninformative priors. Revista Colombiana de Estadística, 36(2), pp.319-336.

Godin, F., VS. Lai, and DA. Trottier. 2019. A general class of distortion operators for pricing contingent claims with applications to CAT bonds. Scandinavian Actuarial Journal 2019(7):558-84.

Wesolowski, C.A., Wanasundara, S.N., Wesolowski, M.J., 2016. A gamma-distribution convolution model of 99mTc-MIBI thyroid time-activity curves. EJNMMI Phys 3, 31 (2016).

Di Salvo, F., 2006. The exact distribution of the weighted convolution of two gamma distributions. In Atti della XLIII Riunione Scientifica SIS (Acts of the 43 Annual Meeting of the Italian Scientific Society). Università di Torino: Società Italiana di Scientifica (pp. 511-4).

Fergusson, K., 2019. Asymptotics of bond yields and volatilities for extended CIR models under the real-world measure. Scandinavian Actuarial Journal, 1-36.

Ahmad, R. and Wilmott, P., 2006. The market price of interest-rate risk: Measuring and modelling fear and greed in the fixed-income markets. Wilmott magazine, pp.64-70.