Stochastic asset price models: analytical background and calibration methods

M. Eriksson

December 25, 2025

Contents

1 Introduction

This document is intended as a technical appendix to the `stochastic-pricing` project, providing a theoretical background to the formulas used and their Python implementation. Namely, we define the stochastic differential equations (SDEs) for each model, derive suitable discretizations for simulation, and then present calibration methods based on moments, regression, and realized variance. We also derive the formulas for value-at-risk (VaR) and expected shortfall (ES) used in the risk analysis.

2 Preliminaries: Itô's lemma and its application to GBM

Itô calculus is often used to discretize differential equations, and so we will begin by recalling Itô's lemma in one dimension.

2.1 Itô's lemma (one-dimensional)

Let X_t satisfy the Itô SDE

$$\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + b(X_t, t)\,\mathrm{d}W_t, \tag{2.1}$$

where $(W_t)_{t\geq 0}$ denotes a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Then the process $Y_t = f(X_t, t)$, where $f : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ is $C^{2,1}$ (twice continuously differentiable in x, once in t), satisfies

$$\mathrm{d}Y_t = \left(\frac{\partial f}{\partial t} + a\,\frac{\partial f}{\partial x} + \frac{1}{2}b^2\,\frac{\partial^2 f}{\partial x^2}\right)\mathrm{d}t + b\,\frac{\partial f}{\partial x}\,\mathrm{d}W_t. \tag{2.2}$$

2.2 Geometric Brownian motion and its distribution of log-returns

Geometric Brownian motion (GBM) is defined by the SDE

$$\mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t, \tag{2.3}$$

where $\mu \in \mathbb{R}$ is the drift and $\sigma > 0$ is the volatility.

In order to obtain formulas for the discretized price and log-returns, we consider the quantity $X_t = \log S_t$. We define $f(x) = \log x$ and $X_t = f(S_t)$, such that $f'(x) = 1/x$ and $f''(x) = -1/x^2$. Comparing (2.3) with (2.1), we identify $a(S_t, t) = \mu S_t$ and $b(S_t, t) = \sigma S_t$. Applying Itô's lemma (2.2), we then have

$$\begin{aligned}
\mathrm{d}X_t &= f'(S_t)\,dS_t + \frac{1}{2}f''(S_t)\,(\mathrm{d}S_t)^2 && (2.4)\\
&= \frac{1}{S_t}(\mu S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t) + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)(\sigma^2 S_t^2\,\mathrm{d}t) && (2.5)\\
&= \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t - \frac{1}{2}\sigma^2\,\mathrm{d}t && (2.6)\\
&= \left(\mu - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t. && (2.7)
\end{aligned}$$

Hence, we obtain a differential equation for the log-price as

$$\mathrm{d}(\log S_t) = \left(\mu - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t, \tag{2.8}$$

which will be used to obtain the update rule for the discretized price and log-returns. By integrating (2.8) from t to $t + \Delta t$, we obtain

$$\log S_{t+\Delta t} - \log S_t = \left(\mu - \frac{1}{2}\sigma^2\right) \Delta t + \sigma(W_{t+\Delta t} - W_t), \tag{2.9}$$

which corresponds to log-return $r_{t+\Delta t}$, i.e.,

$$r_{t+\Delta t} := \log \frac{S_{t+\Delta t}}{S_t} = \left(\mu - \frac{1}{2}\sigma^2\right) \Delta t + \sigma(W_{t+\Delta t} - W_t). \tag{2.10}$$

Since the Brownian stochastic noise term W_t is normally distributed, and $W_{t+\Delta t} - W_t \sim N(0, \Delta t)$, the log-returns then follow the normal distribution

$$r_{t+\Delta t} \sim N\left(\left(\mu - \frac{1}{2}\sigma^2\right) \Delta t,\ \sigma^2 \Delta t\right). \tag{2.11}$$

Lastly, from (2.9) the update rule for the log-price can be equivalently expressed as

$$S_{t+\Delta t} = S_t \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right) \Delta t + \sigma\sqrt{\Delta t}\, Z\right], \tag{2.12}$$

where $Z \sim N(0, 1)$.

3 Models and discretization

In this section we summarize the SDEs and discretizations for the models implemented.

3.1 Geometric Brownian motion

As in (2.3), the GBM SDE reads

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t, \tag{3.1}$$

with exact log-Euler discretization

$$S_{t+\Delta t} = S_t \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right) \Delta t + \sigma\sqrt{\Delta t}\, Z\right]. \tag{3.2}$$

3.2 Ornstein–Uhlenbeck process

The OU process is defined by

$$\mathrm{d}X_t = (\mu - X_t)\theta \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \tag{3.3}$$

with mean-reversion speed $\theta > 0$, long-run mean μ, and volatility parameter $\sigma > 0$. For simulation we use the Euler discretization scheme:

$$X_{t+\Delta t} = X_t + (\mu - X_t)\theta\Delta t + \sigma\sqrt{\Delta t}\, Z. \tag{3.4}$$

Unlike the other models, the OU dynamics are applied to log-prices $X_t = \log S_t$, rather than prices S_t directly.

3.3 Heston stochastic volatility model

The Heston model couples the asset price S_t with a stochastic variance process v_t:

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sqrt{v_t}\, S_t \,\mathrm{d}W_t^{(1)}, \tag{3.5}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\,\mathrm{d}t + \xi\sqrt{v_t}\,\mathrm{d}W_t^{(2)}, \tag{3.6}$$

where $\kappa > 0$ is the variance mean-reversion speed, $\theta > 0$ is the long-run variance, $\xi > 0$ is the volatility-of-volatility, and the noise terms are correlated as

$$\mathrm{d}W_t^{(1)}\mathrm{d}W_t^{(2)} = \rho\,\mathrm{d}t,$$

with $\rho \in [-1, 1]$ parametrizing the strength of the correlation.

The price and variance are discretized as:

$$S_{t+\Delta t} = S_t \exp\left[\left(\mu - \frac{1}{2}v_t\right)\Delta t + \sqrt{v_t\Delta t}\, Z_1\right], \tag{3.7}$$

$$v_{t+\Delta t} = \max\Big(v_t + \kappa(\theta - v_t)\Delta t + \xi\sqrt{v_t\Delta t}\, Z_2, 0\Big), \tag{3.8}$$

where $\max(\,\cdot\,, 0)$ ensures that the variance is positive, and the noise variables (Z_1, Z_2) are constructed to have correlation ρ.

3.4 Merton jump-diffusion model

The Merton model augments GBM with a compound Poisson jump term:

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t + S_t(J - 1)\,\mathrm{d}N_t, \tag{3.9}$$

where N_t is a Poisson process with intensity $\lambda_j > 0$, and J is a random jump multiplier, often parameterized as

$$\ln J \sim N(\mu_j, \sigma_j^2).$$

Over a small interval Δt, the number of jumps $N_{\Delta t} := N_{t+\Delta t} - N_t$ satisfies $N_{\Delta t} \sim \mathrm{Poisson}(\lambda_j \Delta t)$.

The discretization is given by

$$S_{t+\Delta t} = S_t \exp\Big[\big(\mu - \tfrac{1}{2}\sigma^2\big)\,\Delta t + \sigma\sqrt{\Delta t}\, Z\Big] \prod_{k=1}^{N_{\Delta t}} e^{Y_k}, \tag{3.10}$$

where $Y_k = \ln J_k \sim N(\mu_j, \sigma_j^2)$ is the (log-)jump size random variable.

4 Calibration methods

The model parameters are calibrated using real market prices by matching model-implied quantities (moments, distributions, etc.) with observed data. In this project we use method-of-moments and regresion-based estimators, as they are easy to implement, robust, and provide closed-form parameter updates. Maximum likelihood estimation is a natural next step but requires more involved optimization and is left for future extensions.

4.1 GBM model calibration via log-returns

We recall from (2.10) that under GBM, the log-returns follow the distribution

$$r_t = \log \frac{S_t}{S_{t-1}} \sim N\left(\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t,\ \sigma^2 \Delta t\right),$$

i.e., the theoretical return expectation value (mean) $\mathbb{E}[r_t]$ and variance $\mathrm{Var}(r_t)$ are given by

$$\mathbb{E}[r_t] = \left(\mu - \frac{1}{2}\sigma^2\right)\Delta t, \tag{4.1}$$

$$\mathrm{Var}(r_t) = \sigma^2 \Delta t. \tag{4.2}$$

From data, we compute the sample mean $\hat{\mathbb{E}}[r]$ and sample variance $\widehat{\mathrm{Var}}(r)$. We then estimate the model parameters μ and σ by *assuming* that

$$\hat{\mathbb{E}}[r] \approx \mathbb{E}[r_t], \qquad \widehat{\mathrm{Var}}(r) \approx \mathrm{Var}(r_t). \tag{4.3}$$

Letting r_t denote the observed log-returns between fixed time-steps Δt, we define the sample mean and (unbiased) sample variance as

$$\hat{\mathbb{E}}[r] := \frac{1}{T}\sum_{t=1}^{T} r_t, \tag{4.4}$$

$$\widehat{\mathrm{Var}}(r) := \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \hat{\mathbb{E}}[r])^2. \tag{4.5}$$

When the end-time T is large, then $\hat{\mathbb{E}}[r]$ and $\widehat{\mathrm{Var}}(r)$ are consistent estimators of the true mean and variance. By comparing with the theoretical moments in (4.1) and (4.2), we obtain

$$\hat{\mu} = \frac{\hat{\mathbb{E}}[r]}{\Delta t} + \frac{1}{2}\hat{\sigma}^2, \tag{4.6}$$

$$\hat{\sigma}^2 = \frac{\widehat{\mathrm{Var}}(r)}{\Delta t}, \tag{4.7}$$

where the hat denotes estimation. These are the moment-based estimators used in the implementation.

4.2 OU model calibration via linear regression

For the OU process

$$\mathrm{d}X_t = \theta(\mu - X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

we consider the Euler discretization

$$X_{t+\Delta t} = X_t + \theta(\mu - X_t)\Delta t + \sigma\sqrt{\Delta t}\, Z_t, \tag{4.8}$$

where $Z_t \sim N(0,1)$. This equation satisfies a linear regression structure, which can be more clearly seen by rewriting it as

$$X_{t+\Delta t} = a + bX_t + \eta_t, \tag{4.9}$$

where

$$b = 1 - \theta\Delta t, \tag{4.10}$$

$$a = \theta\mu\Delta t, \tag{4.11}$$

$$\eta_t = \sigma\sqrt{\Delta t}\, Z_t. \tag{4.12}$$

Given observed log-prices X_t, with $t = 0, \dots, T-1$, the standard ordinary least square (OLS) method then sets

$$\hat{b} = \frac{\mathrm{Cov}(X_t, X_{t+\Delta t})}{\mathrm{Var}(X_t)}, \tag{4.13}$$

$$\hat{a} = \mathbb{E}[X_{t+\Delta t}] - \hat{b}\,\mathbb{E}[X_t]. \tag{4.14}$$

The residuals $\hat{\eta}_t = X_{t+\Delta t} - (\hat{a} + \hat{b}X_t)$ provide an estimate of $\mathrm{Var}(\eta_t)$:

$$\widehat{\mathrm{Var}}(\eta_t) = \frac{1}{T-1}\sum_{t=0}^{T-1}(\hat{\eta}_t - \bar{\hat{\eta}})^2.$$

The continuous-time parameters are then recovered as

$$\hat{\theta} = \frac{1 - \hat{b}}{\Delta t}, \tag{4.15}$$

$$\hat{\mu} = \frac{\hat{a}}{\hat{\theta}\Delta t}, \tag{4.16}$$

$$\hat{\sigma} = \frac{\sqrt{\widehat{\mathrm{Var}}(\eta_t)}}{\sqrt{\Delta t}}. \tag{4.17}$$

We note that since $X_t = \log S_t$ is used, the OU calibration is applied in log-price space.

4.3 Heston model calibration via realized variance and OU fit

The Heston *variance* dynamics are given by

$$\mathrm{d}v_t = \kappa(\theta - v_t)\,\mathrm{d}t + \xi\sqrt{v_t}\,\mathrm{d}W_t^{(2)}. \tag{4.18}$$

We outline a method-of-moments style calibration based on:

- estimating a discrete variance proxy $\hat{v}_t$ from log-returns (realized variance),
- fitting an OU-like regression to $\hat{v}_t$ to obtain κ and θ,
- estimating ξ from the residual variance,
- estimating the correlation ρ between price and variance shocks.

4.3.1 Realized variance approximation

For sufficiently small Δt, the Heston *price* dynamics

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sqrt{v_t} S_t \,\mathrm{d}W_t^{(1)}$$

imply for log-returns

$$r_t := \log \frac{S_{t+\Delta t}}{S_t} \approx \left(\mu - \tfrac{1}{2} v_t\right) \Delta t + \sqrt{v_t \Delta t}\, Z_t^{(1)},$$

where $Z_t^{(1)} \sim N(0,1)$. If Δt is small and v_t does not change much over $[t, t+\Delta t]$, the drift term is negligible relative to the diffusion, and we can approximate

$$r_t \approx \sqrt{v_t \Delta t}\, Z_t^{(1)}. \tag{4.19}$$

Given the theoretical variance v_t, the conditional expectation value and variance are

$$\mathbb{E}[r_t \mid v_t] \approx 0, \qquad \mathrm{Var}(r_t \mid v_t) \approx v_t \Delta t.$$

A natural estimator for v_t is then the "realized variance"

$$\hat{v}_t = \frac{r_t^2}{\Delta t}. \tag{4.20}$$

Using the approximation (4.19), the conditional expectation value reads

$$\mathbb{E}[\hat{v}_t \mid v_t] = \frac{1}{\Delta t} \mathbb{E}[r_t^2 \mid v_t] \approx \frac{1}{\Delta t} v_t \Delta t \mathbb{E}[Z_t^2] = v_t,$$

which tell us that $\hat{v}_t$ is (approximately) an unbiased estimator of v_t.

In practice, $\hat{v}_t$ is sometimes smoothed (e.g. by a rolling mean) to reduce noise. This is the case also for our code, where we use a five-day average (see `realized_variance(S,dt)` in `utils.py`).

4.3.2 Discrete approximation of the variance dynamics

From (4.18), using Euler discretization, the variance update is given by

$$v_{t+\Delta t} = v_t + \kappa(\theta - v_t)\Delta t + \xi \sqrt{v_t \Delta t}\, Z_t^{(2)}, \tag{4.21}$$

where $Z_t^{(2)} \sim N(0,1)$.

We now replace the unobserved v_t by its proxy $\hat{v}_t$ given by (4.20). Dropping the hat notation for clarity (writing v_t for the proxy), we consider the regression

$$v_{t+\Delta t} - v_t = \kappa\theta\Delta t - \kappa v_t \Delta t + \xi \sqrt{v_t \Delta t}\, Z_t^{(2)}. \tag{4.22}$$

For additional clarity we can rewrite (4.22) as

$$Y_t = A + B X_t + \epsilon_t, \tag{4.23}$$

where we have defined

$$Y_t = v_{t+\Delta t} - v_t, \qquad X_t = v_t$$

and

$$A = \kappa\theta\Delta t, \tag{4.24}$$

$$B = -\kappa\Delta t, \tag{4.25}$$

$$\epsilon_t = \xi \sqrt{v_t \Delta t}\, Z_t^{(2)}. \tag{4.26}$$

Note that ϵ_t is heteroskedastic due to the factor $\sqrt{v_t}$, but we still use OLS as a simple first-order approximation.

4.3.3 Regression-based estimators for κ and θ

Given a time series $\{v_t\}_{t=0}^{T}$, we regard (4.23) as a simple linear regression of Y_t on X_t. OLS provides estimates $\hat{A}$ and $\hat{B}$ via

$$\hat{B} = \frac{\mathrm{Cov}(X_t, Y_t)}{\mathrm{Var}(X_t)}, \tag{4.27}$$

$$\hat{A} = \mathbb{E}[Y_t] - \hat{B}\,\mathbb{E}[X_t]. \tag{4.28}$$

From the identification

$$\hat{B} \approx -\kappa\Delta t, \qquad \hat{A} \approx \kappa\theta\Delta t,$$

we obtain

$$\hat{\kappa} = -\frac{\hat{B}}{\Delta t}, \qquad \hat{\theta} = \frac{\hat{A}}{\hat{\kappa}\Delta t}. \tag{4.29}$$

4.3.4 Estimating ξ from residual variance

The regression residuals are

$$\hat{\epsilon}_t = Y_t - (\hat{A} + \hat{B}X_t) \approx \xi\sqrt{v_t\Delta t}\,\varepsilon_t.$$

Conditional on v_t, we have

$$\mathrm{Var}(\epsilon_t \mid v_t) \approx \xi^2 v_t \Delta t.$$

If we ignore the dependence on v_t and approximate by using the average variance $\bar{v}$, then

$$\mathrm{Var}(\epsilon_t) \approx \xi^2 \bar{v}\Delta t.$$

Let $\widehat{\mathrm{Var}}(\epsilon_t)$ be the sample variance of the residuals

$$\widehat{\mathrm{Var}}(\epsilon_t) = \frac{1}{T-1}\sum_{t=0}^{T-1}(\hat{\epsilon}_t - \bar{\hat{\epsilon}})^2.$$

Then the vol-of vol is approximately

$$\xi^2 \approx \frac{\widehat{\mathrm{Var}}(\epsilon_t)}{\bar{v}\Delta t}.$$

In the code implementation, we use a simpler variant, by ignoring the $\bar{v}$ factor and defining

$$\hat{\xi} = \sqrt{\frac{\widehat{\mathrm{Var}}(\epsilon_t)}{\Delta t}},$$

which effectively absorbs the average scale of v_t into the estimate.

4.3.5 Estimating the correlation ρ

We recall from (3.5) and (3.6) that $\mathrm{d}W_t^{(1)}$ drives the price and $\mathrm{d}W_t^{(2)}$ drives the variance, with the instantaneous correlation ρ. From the log-return approximation (4.19),

$$r_t \approx \sqrt{v_t\Delta t}\,Z_t^{(1)},$$

and the variance error term from (4.26),

$$\epsilon_t \approx \xi \sqrt{v_t \Delta t}\, Z_t^{(2)},$$

where $Z_t^{(1)}$ and $Z_t^{(2)}$ are correlated standard normals, we have

$$\rho \approx \mathrm{Corr}(r_t, \epsilon_t).$$

Hence one can estimate ρ via the empirical correlation between the series of log-returns $\{r_t\}$ and the residuals from the variance regression $\{\hat{\epsilon}_t\}$.

4.4 Merton jump-diffusion model calibrated using moment matching

Under the Merton model, the log-return over Δt can be written as (c.f. (3.10))

$$r_t = \left(\mu - \tfrac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}\, Z_t + \sum_{k=1}^{N_{\Delta t}} Y_k,$$

where again μ and σ is the diffusion drift and volatility (per year), $Z_t \sim N(0,1)$, $N_{\Delta t} \sim \mathrm{Poisson}(\lambda_j \Delta t)$, and $Y \sim N(\mu_j, \sigma_j^2)$. For calibration we use a method-of-moments approach, starting with finding theoretical expressions for the mean, variance, skewness and kurtosis.

4.4.1 Central moments and cumulants

We start by listing the moments of the jump term Y, given by:

$$\begin{aligned}
\mathbb{E}[Y] &= \mu_j, \\
\mathbb{E}[Y^2] &= \mu_j^2 + \sigma^2, \\
\mathbb{E}[Y^3] &= \mu_j^3 + 3\mu_j\sigma_j^2, \\
\mathbb{E}[Y^4] &= \mu_j^4 + 6\mu_j^2\sigma_j^2 + 3\sigma_j^4,
\end{aligned}$$

where μ_j represents the mean jump-size in log-returns and σ_j the jump volatility. For a compound Poisson sum $J = \sum_k Y_k$, the cumulants satisfy

$$\kappa_n(J_t) = \lambda \Delta t\, \mathbb{E}[Y^n], \qquad n \geq 1, \tag{4.30}$$

where λ parametrizes the jump intensity (expected jumps per year). From (4.4), we see that the log-return r_t is the sum of three independent components; a deterministic drift, a Gaussion diffusion term, and a compound Poisson jump term. The full cumulants are then, explicitly:

$$\begin{aligned}
\kappa_1(J) &= (\mu - \sigma^2/2) + \lambda\Delta t \mu_j, \\
\kappa_1(J) &= \sigma^2\Delta t + \lambda\Delta t(\mu_j^2 + \sigma_j^2), \\
\kappa_1(J) &= \lambda\Delta t(\mu_j^3 + 3\mu_j\sigma_j^2), \\
\kappa_1(J) &= \lambda\Delta t(\mu_j^4 + 6\mu_j^2\sigma_j^2 + 3\sigma_j^4).
\end{aligned}$$

For a univariate distribution, the central moments are related to the cumulants via

$$\begin{aligned}
\mu_2 &= \kappa_2, \\
\mu_3 &= \kappa_3, \\
\mu_4 &= \kappa_4 + 3\kappa_2^2,
\end{aligned}$$

and so

$$\mu_2 = \sigma^2 \Delta t + \lambda \Delta t(\mu_j^2 + \sigma_j^2),$$
$$\mu_3 = \lambda \Delta t(\mu_j^3 + 3\mu_j \sigma_j^2),$$
$$\mu_4 = \lambda \Delta t(\mu_j^4 + 6\mu_j^2 \sigma_j^2 + 3\sigma_j^4) + 3\mu_2^2.$$

Finally, the (Pearson) skewness and kurtosis are defined as

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}, \qquad \gamma_2 = \frac{\mu_4}{\mu_2^2},$$

where the skewness measures the asymmetry of the distribution and kurtosis measures tail thickness.

In the context of return distributions, skewness and kurtosis inform us about directional tail risk and the frequency of extreme events, respectively. It is also common to consider *excess kurtosis*, which measures the kurtosis of a distribution relative to that of a normal distribution. It is defined as the Pearson kurtosis minus three, so that a normal distribution has zero excess kurtosis. Financial return distributions typically display positive excess kurtosis, reflecting fat tails and a higher probability of extreme gains and losses than implied by Gaussian models.

4.4.2 Method-of moments calibration

Given observed log-returns $\{r_t\}$, we extract the *empirical* estimates:

- mean $\hat{m}_1$,
- variance $\hat{m}_2$,
- skewness $\hat{\gamma}_1$,
- kurtosis $\hat{\gamma}_2$.

We remark that in the context of method-of-moments analysis, it is customary to denote the mean and variance by $\hat{m}_1$ and $\hat{m}_2$, respectively, even though they are equivalent to the sample estimate of the theoretical expectation variance used in previous sections, i.e., $\hat{m}_1 \equiv \hat{\mathbb{E}}[r]$ and $\hat{m}_2 \equiv \widehat{\mathrm{Var}}(r)$.

In order to calibrate the Merton parameters $\mu, \sigma, \lambda, \mu_j$ and σ_j, we solve for

$$\hat{m}_1 \approx \kappa_1(r), \tag{4.31}$$
$$\hat{m}_2 \approx \kappa_2(r), \tag{4.32}$$
$$\hat{\gamma}_1 \approx \gamma_1(r), \tag{4.33}$$
$$\hat{\gamma}_2 \approx \gamma_2(r). \tag{4.34}$$

In practice, however, the above equation system is often underdetermined (or numerically unstable). A common approach is therefore to fix some parameters and solve for the remaining ones, which is also adopted in our implementation. Specifically, if the skewness is significant ($|\hat{\gamma}_1| > 0.1$), we set $\mu_j = \mathrm{sign}(\hat{\gamma}_1) \times 0.02$, and otherwise zero. This corresponds to 2% jumps in log-price. In addition, we impose a small jump-volatility, $\sigma_j = 0.05$, corresponding to a 5% jump dispersion in log-space.

5 Risk measures

In this section we define the risk measures value-at-risk (VaR) and expected shortfall (ES). For daily (log-)returns r_t, we define losses $L_t := -r_t$.

5.1 Value-at-risk

For a confidence level $\alpha \in (0, 1)$ (e.g. $\alpha = 0.95$ or 0.99), the one-day VaR is defined as

$$\mathrm{VaR}_\alpha = \inf\{\ell \in \mathbb{R} : \mathbb{P}(L_t \leq \ell) \geq \alpha\}, \tag{5.1}$$

where l represents the candidate loss level (percentage/fraction), inf is the infimum (the smallest value that is greater than, or equal to, all the values in the set), and $\mathbb{P}(\,\cdot\,)$ is the probability measure. In plain terms, the VaR is defined as the smallest percentage loss l such that, with probability at least α, the loss will not exceed l. As an example, if $\mathrm{VaR}_{0.99} = 0.025$, the 99% VaR is 2.5%. This means that: on 99% of days, losses will not exceed 2.5%. On the worst 1% of days, losses may be larger.

For continuous distributions, the VaR usually corresponds to a quantile. In terms of returns, if $q_p(r)$ denotes the p-quantile of r_t, then the VaR can be defined as

$$\mathrm{VaR}_\alpha = -q_{1-\alpha}(r_t),$$

which is the version that we implement in the code.

In summary, the VaR is a positive number, expressed in the same units as returns (typically percentages), such that with probability α the loss over the given time horizon does not exceed this level. Equivalently, losses larger than the VaR occur only in the worst $1 - \alpha$ fraction of outcomes.

Note that the VaR does *not* tell us how bad the losses in the worst $1 - \alpha$ fraction tend to be. That is what expected shortfall informs us about, to which we turn to next.

5.2 Expected shortfall

Expected shortfall, also called conditional VaR (CVaR), is defined as the conditional expectation of losses given that they exceed VaR:

$$\mathrm{ES}_\alpha = \mathbb{E}[L_t \mid L_t \geq \mathrm{VaR}_\alpha]. \tag{5.2}$$

In terms of returns,

$$\mathrm{ES}_\alpha = -\mathbb{E}[r_t \mid r_t \leq q_{1-\alpha}(r_t)].$$

The ES measures the average loss in the worst $1 - \alpha$ fraction of outcomes. That is, the expected loss given that the loss exceeds the VaR.

In final, in the code implementation, VaR and ES are estimated empirically by:

- computing the quantile $q_{1-\alpha}(r_t)$ from the sample,
- taking the negative of this quantile for VaR,
- averaging the subset of returns r_t below this quantile to obtain ES.

6 Conclusion

We have derived the key formulas and approximations underlying the stochastic price models used in the project, along with practical calibration procedures based on moments, regression, and realized variance. These derivations establish the analytical foundation of the estimators implemented in the Python code and the subsequent risk assessment via VaR and ES.