Importance sampling and Monte Carlo-based calibration for time-changed Lévy processes

Stefan Kassberger Thomas Liebmann

BFS 2010

Motivation

- Variety of tractable Lévy-models can be represented as time-changed Brownian motion
- Esscher-transform well-established for Lévy-models
- Kassberger and Liebmann (2009) apply independent Esscher transforms to Brownian motion and subordinator in the TCL-context
- Exploit above idea in the context of Monte Carlo simulation:
  - Variance reduction through importance sampling
  - Calculating sensitivities by likelihood-ratio methods
  - Extending sampling algorithms to more general classes of distributions
  - Monte-Carlo based model calibration

Setup

- n-dimensional Brownian motion (B_t), subordinator $(T(t))$.
- Fix τ and denote by $\kappa(u) = \log \mathbb{E}_{\mathbb{Q}}[\exp(uT(\tau))]$ the cumulant generating function of $T(\tau)$.
- Define equivalent measure $\mathbb{S}_{\eta,\gamma}$ via

$$\frac{d\mathbb{S}_{\eta,\gamma}}{d\mathbb{Q}} = \exp\left(\eta' B_{T(\tau)} - \frac{1}{2}|\eta|^2 T(\tau) + \gamma T(\tau) - \kappa(\gamma)\right).$$

- This transform is composed of two Esscher transforms:
  1. One with parameter $\eta \in \mathbb{R}^n$ on the Brownian motion B. Shifts the drift of B from 0 under $\mathbb{Q}$ to η under $\mathbb{S}_{\eta,\gamma}$.
  2. One with parameter γ on the subordinator. Its cgf under $\mathbb{S}_{\eta,\gamma}$ is given by $\kappa^{\mathbb{S}_{\eta,\gamma}}(u) = \kappa(u+\gamma) - \kappa(\gamma)$.

Outline

Importance sampling

- Set $X(t) = B_{T(t)} + mT(t)$.
- Take expectations under the transformed measure:

$$\mathbb{E}_{\mathbb{Q}}\left[f((X(t))_{t\in[0,\tau]})\right]$$
$$= \mathbb{E}_{\mathbb{S}_{\eta,\gamma}}\left[\exp\left(-\eta' B_{T(\tau)} + \frac{1}{2}|\eta|^2 T(\tau) - \gamma T(\tau) + \kappa(\gamma)\right) f((X(t))_{t\in[0,\tau]})\right].$$

- If the subordinator cannot be efficiently sampled under $\mathbb{Q}$ but under $\mathbb{S}_{\eta,\gamma}$ for certain values of γ, we can choose η for variance reduction via importance sampling (compare example for Normal Tempered Stable process).

Importance sampling

- Goal: Estimate

$$\mathbb{E}_{\mathbb{S}_{\eta,\gamma}}\left[\exp\left(-\eta' B_{T(\tau)}+\frac{1}{2}|\eta|^2 T(\tau)-\gamma T(\tau)+\kappa(\gamma)\right) f((X(t))_{t\in[0,\tau]})\right] \quad (1)$$

  by simulation.
- Simulate appropriately discretized version of $(X(t))$: $(X(t_i))_{i=0,\dots,N}$ with $0=t_0<\dots<t_N=\tau$.
- Proceed as follows:
  - Simulate $(T(t_i))_{i=0,\dots,N}$ and iid $\mathcal{N}_n(0, diag(1_n))$ rvs $W_1,\dots,W_N$.
  - Set $Y(t_i)=Y(t_{i-1})+W_i\sqrt{T(t_i)-T(t_{i-1})}+\eta(t_i-t_{i-1})$.
  - Set $X(t_i)=Y(t_i)+mT(t_i)$.
  - Repeat M times to arrive at set of sample paths indexed by k.
- Estimator for (1) is

$$\frac{1}{M}\sum_{k=1}^{M}\exp\left(-\eta' Y_k(\tau)+\frac{1}{2}|\eta|^2 T_k(\tau)-\gamma T_k(\tau)+\kappa(\gamma)\right) f((X_k(t_i))_{i=0,\dots,N}).$$

Importance sampling: Variance of the Estimator

- Let $|f(x)| \leq ce^{\beta' x}$ for some $c \geq 0$ and $\beta \in \mathbb{R}^n$.
- Bound for variance of the estimator

$$Var_{\mathbb{S}_{\eta,\gamma}}\left[\frac{1}{M}\sum_{k=1}^{M}\exp\left(-\eta' Y_k(\tau)+\frac{1}{2}|\eta|^2 T_k(\tau)-\gamma T_k(\tau)+\kappa(\gamma)\right)f(X_k(\tau))\right]$$
$$\leq \frac{c^2}{M}\exp\left(\kappa(|\beta|^2+|\beta-\eta|^2-\gamma+2\beta' m)+\kappa(\gamma)\right)$$

- κ is increasing and convex. For given β, c and γ, the bound is minimal if $\eta = \beta$.
- If γ can be chosen freely, minimum is attained for $\gamma = |\beta|^2/2 + \beta' m$.
- If further $m = 0$, this holds for $\gamma = |\beta|^2/2 = |\eta|^2/2$, i.e. an Esscher transform on $B_{T(\tau)}$ with parameter β.

Outline

Sensitivities with the LR-method

- Sensitivities w.r.t. Esscher transform parameters η and γ can be estimated by the likelihood ratio method
- If the expectations are finite in some neighborhood of γ_0,

$$\frac{d}{d\gamma}\mathbb{E}_{\mathbb{S}_{\eta_0,\gamma}}\left[f((X(t))_{t\in[0,\tau]})\right]\Big|_{\gamma=\gamma_0} = \mathbb{E}_{\mathbb{S}_{\eta,\gamma}}\left[\frac{d\mathbb{S}_{\eta_0,\gamma_0}}{d\mathbb{Q}}\frac{d\mathbb{Q}}{d\mathbb{S}_{\eta,\gamma}}\left(T(\tau)-\kappa'(\gamma_0)\right)f((X(t))_{t\in[0,\tau]})\right]$$

- If the expectations are finite in some neighborhood of η_0,

$$\frac{d}{d\eta}\mathbb{E}_{\mathbb{S}_{\eta,\gamma_0}}\left[f((X(t))_{t\in[0,\tau]})\right]\Big|_{\eta=\eta_0} = \mathbb{E}_{\mathbb{S}_{\eta,\gamma}}\left[\frac{d\mathbb{S}_{\eta_0,\gamma_0}}{d\mathbb{Q}}\frac{d\mathbb{Q}}{d\mathbb{S}_{\eta,\gamma}}\left(B_{T(\tau)}-\eta_0 T(\tau)\right)f((X(t))_{t\in[0,\tau]})\right]$$

Outline

Simulation of a Normal Tempered Stable process

- Let T be a Tempered Stable $\mathrm{TS}(\kappa, a, (2\gamma)^{\kappa})$ random variable with $\kappa \in (0,1)$, $a, \gamma > 0$ and cumulant generating function

$$\kappa_{\kappa,a,\gamma}(u) = a(2\gamma)^{\kappa} - a(2\gamma - 2u)^{\kappa}.$$

- Algorithm that allows direct sampling is unknown
- Idea: Simulate stable random variables and perform Esscher transform
- Algorithm:
  1. Sample n iid stable $\mathrm{TS}(\kappa, a, 0)$ rvs T_k
  2. Set importance weights to

$$w_k = \exp(-\gamma T_k + \kappa_{\kappa,a,\gamma}(\gamma)) = \exp(-\gamma T_k + a(2\gamma)^{\kappa})$$

- In Normal Tempered Stable model, η still available for IS

Outline

MC-based model calibration

- MC-based model calibration time-consuming and unstable
  - Change of parameters often requires re-simulation
  - Noise due to finite sample size often precludes gradient-based methods and slows down convergence of optimizer
- Idea: Use of two Esscher transforms often allows calibration *based on a single set of paths* or at least helps significantly reduce number of simulation runs
- Simulate BM and subordinator, keep paths in memory, and then apply Esscher transforms in conjunction with basic transformation (shifting & scaling)
- Facilitates use of gradient-based optimization algorithms (more exact & stable gradients)

Outline

Example: MC-calibration of NIG model

- Objective: Calibrate model based on NIG$(\alpha, \beta, \delta, \mu)$ paths, i.e. find $\alpha, \beta, \delta, \mu$ such that observed derivatives prices are replicated
- Proceed as follows (easily generalizes to appropriate time-scaling):
  1. Fix c. Simulate IG$(1, c)$ paths (IG_t) and standard Brownian paths (B_t)
  2. Fix $\alpha, \beta, \delta, \mu$
  3. Calculate $b = \delta\sqrt{\alpha^2 - \beta^2}$ and choose γ such that $\sqrt{c^2 - 2\gamma} = b$.
  4. Calculate Esscher weights corresponding to γ.
  5. Under $\mathbb{S}_{0,\gamma}$,
  $$X = \beta\delta^2 IG_1 + \delta B_{IG_1} + \mu$$
  is NIG$(\alpha, \beta, \delta, \mu)$-distributed.
  6. Calculate option prices / evaluate objective function under $\mathbb{S}_{0,\gamma}$,
  7. Repeat steps 2-7 if necessary
- η (unused above) can be employed for importance sampling.

Example: Pricing a binary down-and-out call via MC

- Let the log-return process (X_t) be a Lévy-process with $X_1 \sim NIG(\alpha, 0, \delta, \mu)$ with $\mu = \mu(\alpha, \delta)$ chosen such that $(S_0 \cdot \exp(X_t))$ is a martingale (assume riskless interest rate $r = 0$ and $S_0 = 100$).
- Price a binary down-and-out call (BDOC) in this model via Esscher-based MC.
- Payoff of BDOC maturing in 1 year:

$$\mathbb{I}(\min(S_t, t \in [0,1]) \geq 80) \cdot \mathbb{I}(S_1 \geq 100)$$

- Price BDOC via MC: Discretize using 250 time-steps. Simulate 100000 Brownian and IG paths.
- Use Esscher-based MC:
  - Smooth dependence of BDOC-price on parameters (α, δ).
  - Fast calculation of option price surface as only one simulation run is needed.

BDOC price surface

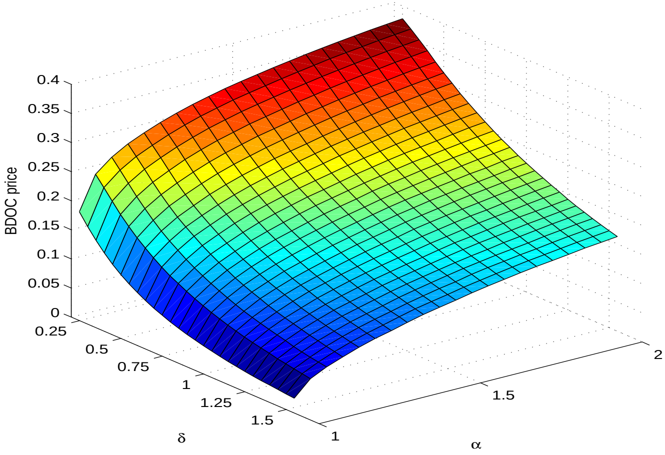

Figure: BDOC price as function of α and δ

Example: Calculating sensitivities w.r.t. model parameters

- Sensitivities based on plain MC notoriously noisy
- Esscher-based MC makes sensitivity estimates more stable and provides faster convergence
- In above setup, calculate sensitivity of BDOC-price w.r.t. model parameter δ for $\alpha_0 = 2$ and $\delta_0 = 1$
- First method: Finite differences with resampling (same seed for random number generator in both runs)
- Second method: Finite differences with Esscher (only one sample plus Esscher transform)

Convergence of parameter sensitivities

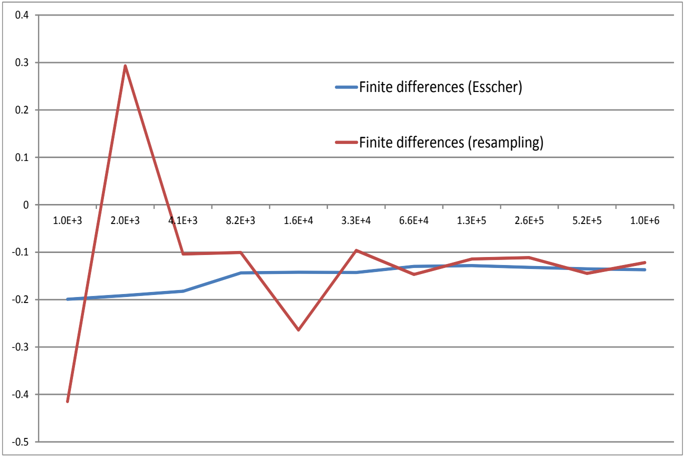

Figure: Sensitivities w.r.t. δ as function of sample size

Summary

- Apply independent Esscher transforms to subordinator and BM
- Gives higher degree of flexibility than Esscher transform applied to TCBM itself
- Subclass of structure preserving transforms (Kassberger and Liebmann (2009))
- Variety of applications:
  1. IS with upper bounds for variance
  2. IS with a mixture of importance distributions
  3. sensitivities w.r.t. Esscher parameters via LR-methods
  4. development of sampling algorithms
  5. model calibration
- Approach also works for subordinated stable processes

Thanks for your attention!