Stochastic-alpha-beta-rho (SABR) Model
Applied Stochastic Processes (FIN 514)

Instructor: Jaehyuk Choi

Peking University HSBC Business School, Shenzhen, China

2020-21 Module 3 (Spring 2021)

The project overview

SABR Model

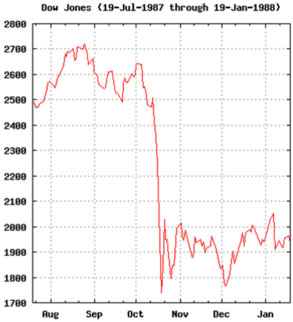
- One of the most popular stochastic volatility (SV) model.
- Heavily used for pricing and risk-managing options in interest rate and FX.
- Explains volatility skew/smile with minimal and intuitive parameters.

Project Goal

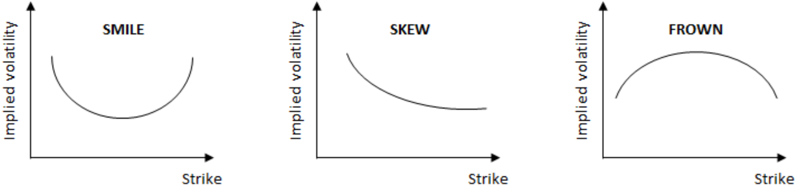
- Implement option pricing with Euler/Milstein scheme
- Implement conditional MC method (and check the variance reduction)
- Compare to the approximation formula by Hagan (code provided)
- Implement a smile calibration routine

Background: volatility skew/smile

- Black Monday crash in 1987:
  DJIA -22.6% in one day!
- Overall 'short gamma' due to the *portfolio insurance* (put on equity index)
- Market values (down-side) tail event higher than before.
- Market sees volatility skew/smile



(From Wikipedia)



Why need model for smile? challenges in risk management

- Option trading desk (market-making/sell-side) usually accumulates option positions with different strikes.
- Under BSM model,
  - Vol σ fixed under spot change $S_0 \to S_0 + \Delta$.
  - Risk-management is easy: delta and vega clearly defined
  - One can hedge delta (with underlying stock) and vega (with ATM option)
  - However, the OTM option prices/risks are not correct!
- BSM model with different σ to each option K?
  - How do we fix the volatilities?
  - Sticky strike rule $\sigma = \sigma(K)$ vs sticky delta rule $\sigma = \sigma(S_0 - K)$.
  - Need to characterize the smile with a few minimal parameters.
- For better risk management, we need models which can capture the volatility smile.

How to model smile? Local volatility (LV)

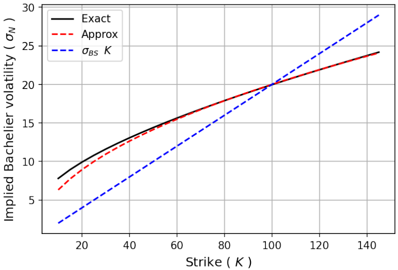
- Volatility depending on the 'current location' of S_t:

$$\text{BSM: } \frac{dS_t}{S_t} = \sigma_\text{BS} f_\text{BS}(S_t)\, dW_t \qquad \text{Normal: } dS_t = \sigma_\text{N} f_\text{N}(S_t)\, dW_t$$

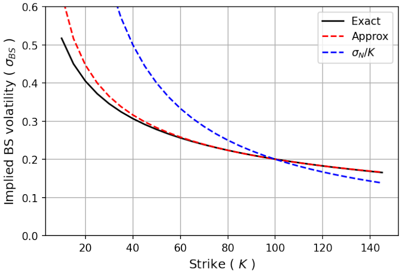
- BSM model: a trivial case with $f_\text{BS}(x) = 1$. However, it is a local vol model under normal volatility ($f_\text{N}(x) = x$).
- Normal model: a trivial case with $f_\text{N}(x) = 1$. However, it is a local vol model under BSM volatility ($f_\text{BS}(x) = 1/x$).
- What is the implied normal volatility of the Black-Scholes price on varying K? What is the relation between the implied volatility and the local vol?
- The implied volatility is the volatility average of the in-the-money paths
- **Exercise 1**: Chart the normal implied vol of the prices under BSM model for typical parameter sets. Measure the slope, $\partial\sigma(K)/\partial K$, at the money.

Parameters: $S_0 = 100, \sigma_{\rm BS} = 20\% (\sigma_{\rm N} = 20), r = q = 0$:

- Implied normal vol for constant BSM vol ($\sigma_{\rm BS} = 20\%$)
- Approx: $\sigma_{\rm N} \approx \sigma_{\rm BS}\sqrt{S_0 K}$
- Local vol: $\sigma_{\rm N} \approx \sigma_{\rm BS} K$



- Implied BSM vol for constant normal vol ($\sigma_{\rm N} = 20$):
- Approx: $\sigma_{\rm BS} \approx \sigma_{\rm N}/\sqrt{S_0 K}$
- Local vol: $\sigma_{\rm BS} \approx \sigma_{\rm N}/K$



Displaced BM (DBS) model

- A simple local vol model with analytic solution (i.e., Black-Scholes formula)
- *Displaced* asset price $D(S_t) = \beta S_t + (1-\beta)S_0$ follows a GBM:

$$\frac{dS_t}{D(S_t)} = \sigma_{\rm D}\, dW_t \quad \text{where} \quad D(S_t) = \beta\, S_t + (1-\beta)S_0.$$

- Calibration of $\sigma_{\rm D}$ (ATM option price on target):

$$\sigma_{\rm N} \approx \sigma_{\rm D} D(S_0) \approx \sigma_{\rm BS}\, S_0 \quad \Rightarrow \quad \sigma_{\rm D} = \sigma_{\rm BS}$$

- Bridges the Bachelier ($\beta = 0$) and BS ($\beta = 1$) models.

Displaced BM (DBS) model

- Final asset price S_T:

$$S_T = \left(S_0 + \frac{1-\beta}{\beta} S_0\right) \exp\left(\beta\sigma_{\mathrm{D}} W_T - \frac{\beta^2\sigma_{\mathrm{D}}^2 T}{2}\right) - \frac{1-\beta}{\beta} S_0,$$

- Option price:

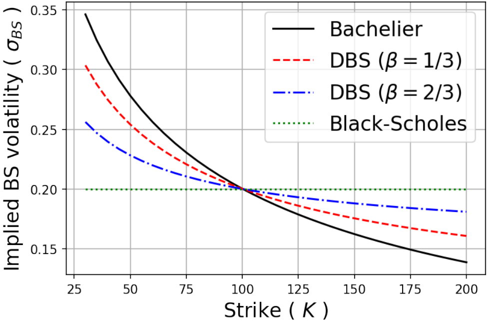
$$C_{\mathrm{D}}(K) = \frac{D(S_0)N(d_{1\mathrm{D}}) - D(K)N(d_{2\mathrm{D}})}{\beta},$$

$$\text{where} \quad d_{1\mathrm{D},2\mathrm{D}} = \frac{\log\left(D(S_0)/D(K)\right)}{\beta\sigma_{\mathrm{D}}\sqrt{T}} \pm \frac{\beta\sigma_{\mathrm{D}}\sqrt{T}}{2}.$$

- We can reuse BS formula with the following substitutions:

$$F_0 \Rightarrow D(F_0), \quad K \Rightarrow D(K), \quad \sigma_{\mathrm{BS}} \Rightarrow \beta\sigma_{\mathrm{D}}, \quad C_{\mathrm{BS}} \Rightarrow \beta C_{\mathrm{D}}$$

Exercise 2: Chart the BS implied volatility of the prices under the DBS model.



How to model smile? Stochastic volatility (SV)

- Volatility changing over time:

$$\text{BSM: } \frac{dS_t}{S_t} = \sigma_t \, dW_t \qquad \text{Normal: } dS_t = \sigma_t \, dW_t$$

- Many models proposed (mostly for BSM). For $dW_t dZ_t = \rho \, dt$,
  - Hull-White and SABR:

$$\frac{d\sigma_t}{\sigma_t} = \nu \, dZ_t$$

  - Heston: $V_t = \sigma_t^2$ follows Cox-Ingersoll-Ross (CIR) process,

$$dV_t = \kappa(V_\infty - V_t)dt + \nu\sqrt{V_t}dZ_t$$

- SV model correctly captures the smile, ν for curvature and ρ for skewness.

SABR model: LV + SV

Stochastic-alpha-beta-rho model SDE:

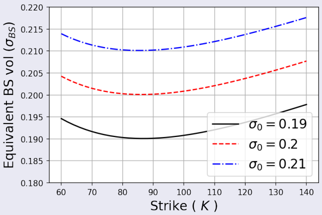
$$\frac{dS_t}{S_t^\beta} = \sigma_t dW_t \quad \text{and} \quad \frac{d\sigma_t}{\sigma_t} = \nu dZ_t \quad (dW_t\, dZ_t = \rho\, dt)$$

- Parameters: σ_0, ν, β, ρ.
- σ_0: overall volatility, calibrated to ATM implied vol
- β: elasticity or 'backbone'. (Normal: $\beta = 0$, BSM: $\beta = 1$)
- ν: volatility of volatility, σ following a GBM
- ρ: correlation between asset price and volatility

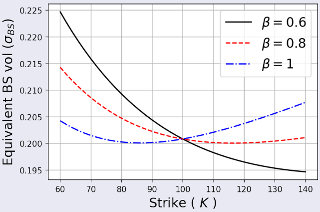
The impact of parameters

Base parameters: $\sigma_0 = 0.2,\ \nu = 0.2,\ \rho = 0.1,\ \beta = 1.$

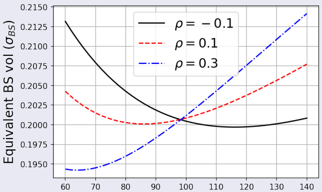
Initial vol σ_0



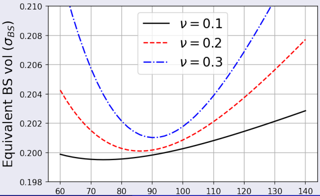
Backbone β



Correlation ρ



Vol-of-vol ν



Equivalent BSM-volatility formula (Hagan et al, 2002)

The leading-order and first-order terms of Taylor's expansion around $\nu\sqrt{T} \approx 0$.

$$\sigma_{\text{BS}}(K) = H(z)\frac{\alpha}{k^{\beta_*/2}}\,\frac{1 + \left(\frac{\beta_*^2}{24\,k^{\beta_*}}\alpha^2 + \frac{\rho\beta}{4\,k^{\beta_*/2}}\alpha\nu + \frac{2-3\rho^2}{24}\nu^2\right)T}{1 + \frac{\beta_*^2}{24}\log^2 k + \frac{\beta_*^4}{1920}\log^4 k},$$

where,

$$\beta_* = 1-\beta,\ \alpha = \frac{\sigma_0}{F_0^{\beta_*}},\ k = \frac{K}{F_0},\ z = \frac{\nu}{\alpha}k^{\beta_*/2}\log k,$$

$$H(z) = \frac{z}{x(z)},\ x(z) = \log\left(\frac{V(z)+z+\rho}{1+\rho}\right),\ \text{and } V(z) = \sqrt{1+2\rho z+z^2}.$$

The option price can be obtained by plugging $\sigma_{\text{BS}}(K)$ in to the Black-Scholes formula.

$$C_{\text{SABR}} = C_{\text{BS}}(K, F_0, \sigma_{\text{BS}}(K), T)$$

Success of the SABR model

- Volatility smile information encoded into three parameters σ_0, ν, ρ.
- These three parameters are parsimonious (minimal) and intuitive.
- Equivalent BSM volatility is available although not accurate for wide parameter range.
- Vega (volatility) risk managed by the three parameters rather than each individual vol.
- Three implied vols (or option prices) on the smile can calibrate the parameters. $\rightarrow$ An effective interpolation method for implied volatility (or option price)

Limitation of Hagan's formula

- Arbitrage is equivalent to some event happening with negative probability. The price of a derivative paying $1 on that event is negative (should be free at most)!
- Probability (digital call option price) from the call spread:

$$\begin{aligned}\mathbb{P}(S_T > K) &= D(K, \sigma(K))\\ &= \frac{C_{\text{BS}}(K, \sigma(K)) - C_{\text{BS}}(K+\Delta K, \sigma(K+\Delta K))}{\Delta K} \quad = -\frac{\partial C_{\text{BS}}(K, \sigma(K))}{\partial K}\end{aligned}$$

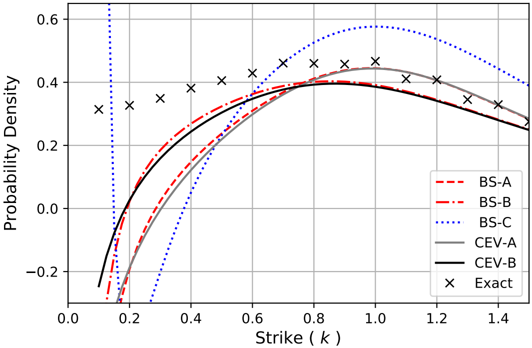
- Probability density from the second derivative:

$$f_{SABR}(K) = \frac{\partial^2 C_{\text{BS}}(K, \sigma(K))}{\partial K^2} \geq 0.$$

- The PDF from the exact SABR solution should be positive. When $\nu\sqrt{T} \gg 1$, however, Hagan's approximation formula often implies negative PDF.

$$D(K, \sigma(K)) < D(K+\Delta K, \sigma(K+\Delta K)).$$

  The volatility effect $\sigma(K+\Delta K)$ dominates (should NOT!) the moneyness effect $K+\Delta K$.



- Parameters: $\sigma_0 = 0.25,\ \nu = 0.3,\ \rho = -0.2,\ \beta = 0.6,\ T = 20$
- Many volatility approximation methods imply negative PDF at low K.
- Reference: Choi, J., & Wu, L. (2019). The equivalent constant elasticity of variance (CEV) volatility of the stochastic-alpha-beta-rho (SABR) model. ArXiv:1911.13123 [q-Fin]. http://arxiv.org/abs/1911.13123

Euler method (MC with time-discretization)

- Unlike normal or BSM model (as in spread/basket option project), we can not jump the simulation directly from $t = 0$ to T.
- Divide the interval $[0, T]$ into N small steps, $t_k = (k/N)T$ and $\Delta t_k = T/N$ and simulate each time step with

$$S_t : \begin{cases} \beta = 0 : \;\; S_{t_{k+1}} = S_{t_k} + \sigma_{t_k} W_1 \sqrt{\Delta t_k} \\ \beta = 1 : \;\; \log S_{t_{k+1}} = \log S_{t_k} + \sigma_{t_k} \sqrt{\Delta t_k}\, W_1 - \frac{1}{2}\sigma_{t_k}^2 \Delta t_k, \end{cases}$$

$$\sigma_t : \sigma_{t_{k+1}} = \sigma_{t_k} \exp\left(\nu \sqrt{\Delta t_k}\, Z_1 - \frac{1}{2}\nu^2 \Delta t_k\right),$$

  where $W_1,\ Z_1 \sim N(0,1)$ with correlation ρ.
- Typically, $\Delta t_k \approx 0.25$. For $T = 30$, $N = 120$, quite time-consuming.
- Any good control variate?

$$C(K) = \frac{1}{N} \sum_{i=1}^{N} (S_T^{(i)} - K)^+$$

Euler method vs Milstein method

For a stochastic process,

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,$$

the Euler scheme is given as:

$$X_{t+\Delta t} = X_t + \mu(X_t)\Delta t + \sigma(X_t)\, W_1\sqrt{\Delta t} \quad \text{for} \quad W_1 \sim N(0,1).$$

In Milstein scheme, an higher-order correction is added:

$$\begin{aligned} X_{t+\Delta t} &= X_t + \mu(X_t)\Delta t + \sigma(X_t)\Delta W_t + \frac{\sigma(X_t)\sigma'(X_t)}{2}\left((\Delta W_t)^2 - \Delta t\right), \\ &= X_t + \mu(X_t)\Delta t + \sigma(X_t)W_1\sqrt{\Delta t} + \frac{\sigma(X_t)\sigma'(X_t)}{2}\Delta t\left(W_1^2 - 1\right). \end{aligned}$$

The idea is from the well-known stochastic integral

$$\int_0^{\Delta t} W_t dW_t = \frac{1}{2}((\Delta W_t)^2 - \Delta t) = \frac{\Delta t}{2}(W_1^2 - 1).$$

Milstein Scheme (continued)

For the time s, $t \le s \le t + \Delta t$, the dynamics of $\sigma(X_s)$ is

$$d\sigma(X_s) = \sigma'(X_s)dX_s + O(\Delta t) = \sigma'(X_s)\sigma(X_s)dW_s + O(\Delta t).$$

Applying the Euler scheme, we get

$$\sigma(X_s) = \sigma(X_t) + \sigma'(X_t)\sigma(X_t)(W_s - W_t) + O(\Delta t).$$

The Milstein scheme is derived as

$$\begin{aligned}
X_{t+\Delta t} - X_t &= \mu(X_t)\int_{s=t}^{t+\Delta t} ds + \int_{s=t}^{t+\Delta t} \sigma(X_s)\, dW_s \\
&= \mu(X_t)\Delta t + \int_{s=t}^{t+\Delta t} \big(\sigma(X_t) + \sigma'(X_t)\sigma(X_t)(W_s - W_t)\big)\, dW_s \\
&= \mu(X_t)\Delta t + \sigma(X_t)\Delta W_t + \sigma'(X_t)\sigma(X_t)\int_{s=t}^{t+\Delta t} (W_s - W_t)\, dW_s \\
&= \mu(X_t)\Delta t + \sigma(X_t)\Delta W_t + \frac{\sigma'(X_t)\sigma(X_t)}{2}((\Delta W_t)^2 - \Delta t) \\
&= \mu(X_t)\Delta t + \sigma(X_t)W_1\sqrt{\Delta t} + \frac{\sigma(X_t)\sigma'(X_t)}{2}\Delta t\big(W_1^2 - 1\big).
\end{aligned}$$

Stochastic integral of σ_t

From Itô's lemma,

$$\frac{d\sigma_t}{\sigma_t} = \nu\, dZ_t \quad \Rightarrow \quad d\log\sigma_t = -\frac{1}{2}\nu^2 dt + \nu dZ_t$$

we can solve the volatility process:

$$\sigma_t = \sigma_0 \exp\left(\nu Z_t - \frac{1}{2}\nu^2 t\right).$$

We also know

$$\nu \int_0^T \sigma_t dZ_t = \sigma_T - \sigma_0 = \sigma_0 \exp\left(-\frac{1}{2}\nu^2 T + \nu Z_T\right) - \sigma_0,$$

which will be useful for the integration of S_t.

Stochastic integral of S_t (normal: $\beta = 0$)

Writing the SDE in a de-correlated form,

$$dS_t = \sigma_t \left(\rho dZ_t + \sqrt{1-\rho^2} dX_t\right) \quad \text{with} \quad dX_t\, dZ_t = 0.$$

Integrating S_t, we get so far as

$$\begin{aligned} S_T - S_0 &= \rho \int_0^T \sigma_t dZ_t + \sqrt{1-\rho^2} \int_0^T \sigma_t dX_t \\ &= \frac{\rho}{\nu}(\sigma_T - \sigma_0) + \sqrt{1-\rho^2} \int_0^T \sigma_t dX_t \end{aligned}$$

From Itô's Isometry, the integration in blue is equivalent to

$$\int_0^T \sigma_t dX_t = X_1 \sqrt{V_T} \quad \text{where} \quad X_1 \sim N(0,1), \quad V_T := \int_0^T \sigma_t^2 dt.$$

Here, the random variable X_1 is independent from V_T and σ_T.

Normalization of V_T

- Note that $V_T = \sigma_0^2 T$ if $\nu = 0$ (i.e., volatility is not stochastic).
- We normalize by $I_T = I_T/(\sigma_0^2 T)$:

$$
\begin{aligned}
I_T = \frac{V_T}{\sigma_0^2 T} &= \frac{1}{\sigma_0^2 T}\int_0^T \sigma_t^2 dt \\
&= \frac{1}{\sigma_0^2 T}\int_0^T \sigma^2 \exp\left(2\nu Z_t - \nu^2 t\right) dt \\
&= \frac{1}{T}\int_0^T \exp\left(2\nu Z_t - \nu^2 t\right) dt \\
&= \int_0^1 \exp\left(2\hat{\nu}\, Z_s - \hat{\nu}^2 s\right) ds, \quad (\hat{\nu} = \nu\sqrt{T})
\end{aligned}
$$

- We don't need to simulate σ_t even if σ_0 changes.

Conditional MC method (normal $\beta = 0$)

Conditional on (σ_T, I_T), S_T can be sampled from

$$S_T = S_0 + \frac{\rho}{\nu}(\sigma_T - \sigma_0) + \sigma_0\sqrt{(1-\rho^2) I_T\, T}\; X_1$$

and the option price is from the normal model:

$$C_\mathrm{N}\left(K, S_0 := S_0 + \frac{\rho}{\nu}(\sigma_T - \sigma_0),\ \sigma_\mathrm{N} := \sigma_0\sqrt{(1-\rho^2) I_T}\right)$$

Then, the price is obtained as an expectation over (σ_T, I_T):

$$C_{\beta=0} = E\left(C_\mathrm{N}(\sigma_T, I_T)\right), \quad \text{where} \quad I_T = \frac{1}{N}\sum_k \sigma_{t_k}^2 \quad (\sigma_0 = 1)$$

For I_T, we can use higher-order numerical integration methods (trapezoidal rule or Simpson's rule)

$$I_T = \frac{1}{2N}\sum_{k=0}^{N-1}(\sigma_{t_k}^2 + \sigma_{t_{k+1}}^2) = \frac{1}{2N}\left(\sigma_{t_0}^2 + 2\sigma_{t_1}^2 + \cdots + 2\sigma_{t_{N-1}}^2 + \sigma_{t_N}^2\right)$$

Conditional MC method (BSM $\beta = 1$)

Conditional on (σ_T, I_T), S_T can be sampled from

$$\log\left(\frac{S_T}{S_0}\right) = \frac{\rho}{\nu}(\sigma_T - \sigma_0) - \frac{\sigma_0^2 T}{2} I_T + \sigma_0\sqrt{(1-\rho^2) I_T\, T}\; X_1$$

and the option price is from the BSM formula:

$$C_{\rm BS}\left(K, S_0 := S_0 e^{\frac{\rho}{\nu}\left(\sigma_T - \sigma_0\right) - \frac{\rho^2\sigma_0^2 T}{2} I_T}, \sigma_{\rm BS} := \sigma_0\sqrt{(1-\rho^2) I_T}\right)$$

Then, the price is obtained as an expectation over (σ_T, I_T):

$$C_{\beta=1} = E\left(C_{\rm BS}(\sigma_T, I_T)\right), \quad \text{where} \quad I_T = \frac{1}{N}\sum_k \sigma_{t_k}^2 \quad (\sigma_0 = 1)$$

For I_T, we can use higher-order numerical integration methods (trapezoidal rule or Simpson's rule)

$$I_T = \frac{1}{3N}\left(\sigma_{t_0}^2 + 4\sigma_{t_1}^2 + 2\sigma_{t_2}^2 + \cdots + 4\sigma_{t_{N-2}}^2 + 2\sigma_{t_{N-1}}^2 + \sigma_{t_N}^2\right) \quad \text{for even } N$$

Advantages of conditional MC method

- No need to simulate S_t: less computation, less memory use.
- Given (σ_T, I_T), the option price is exact. Therefore, MC variance is much smaller than that of the MC simulating both σ_t and S_t.
- Can obtain correct option value for extreme strike values: If we have so simulate S_T, no simulation path arrives at $S_T > K$ for very big or small K, option value from MC is zero. The conditional MC method result in very small (correct) option value because the price comes from (analytic) BSM formula.

Smile Calibration

- When β is given (0 or 1), three parameters, σ_0, ρ and ν, can be calibrated to three option prices (or implied volatilities), typically at $K = S_0$ (ATM), $S_0 - \Delta$ and $S_0 + \Delta$.

$$\text{SABR}(\sigma_0, \rho, \nu) \rightarrow \sigma(S_0), \sigma(S_0 - \Delta), \sigma(S_0 + \Delta)$$

- Write a calibration routine in R to solve σ_0, ρ and ν in homework.