Training Data Generation

Current Idea

1. Given a set of parameters for GARCH (in Physical measure)
2. Given the initial asset price S_0, use Monte Carlo method to simulate a path of asset prices, $S_1, S_2, ...S_N$, with say $N = 500$ (Under **P** measure)
3. Select last 30-50 days on the path, for each day, use the selected asset price (under **Q**) as the initial price to generate American option prices with various strike prices (11-17) and maturities (7 days to 1 year). **Pay attention to the transformation from the physical measure to the risk-neutral measure**.

Pseudo Code

```
# Pseudo Code of Pricing American Options By the Willow Tree and Monte Carlo Method
# 1. Initialize option parameters
T = ... # Time to maturity
K = ... # Strike price
r = ... # Risk-free rate
S_0 = ... # Initial asset price
N = ... # Number of paths
delta = T/N # Time step
h0 = ... # Initial volatility
# 2. Initialize HN-GARCH parameters under P Measure
alpha = ...
beta = ...
omega = ...
gamma = ...
lambda = ...
# 3. Initialize Willow Tree parameters
m_h = ...
m_ht = ...
m_x = ...
gamma_h = ...
gamma_x = ...
# 4. Construct the willow tree for ht
hd, qhd = genhDelta(h0, beta, alpha, gamma, omega, m_h, gamma_h)
nodes_ht = TreeNodes_ht_HN(m_ht, hd, qhd, gamma_h, alpha, beta, gamma, omega, N + 1)
P_ht_N, P_ht = Prob_ht(nodes_ht, h0, alpha, beta, gamma, omega)
# 5. Construct the willow tree for Xt
nodes_Xt, mu, var, k3, k4 = TreeNodes_logSt_HN(m_x, gamma_x, r, hd, qhd, S0, alpha,
beta, gamma, omega, N)
q_Xt, P_Xt, tmpHt = Prob_Xt(nodes_ht, qhd, nodes_Xt, S0, r, alpha, beta, gamma,
omega)
nodes_S = np.exp(nodes_Xt)
# 6. Generate Data (can possibly be done in parallel or concurrently)
days = ...
A_prices = np.zeros(days) # American option prices
A_sig = np.zeros(days) # Implied volatility
A0_prices = np.zeros(days) # Price of option using model parameters

for i in range(days): # Days on the path
  CorP = -1 # Call or Put
  A_sig[i], A_prices[i], A0_prices[i] = impVol_HN(r, lambda, omega, beta, alpha,
                                        ...gamma, h0, S0, K, T, N, m_h, m_x, cor_p)
```

Functions to be aware of:

- `genhDelta`: Generates the discrete values and probabilities of a std normal distribution that are used to construct a Willow tree for the conditional variance in the HN model.
- `TreeNodes_ht_HN`: Constructs the Willow tree for the conditional variance in the HN model.
- `Prob_ht`: Calculates the transition probabilities of the nodes in the Willow tree for the conditional variance in the HN model.
- `TreeNodes_logSt_HN`: Constructs the nodes of the Willow tree for the log asset price in the HN model, as well as the first four moments.
- `Prob_Xt`: Calculates the transition probabilities of the nodes in the Willow tree for the log asset price in the HN model.
- `impVol_HN`: Calculates the American option price, the implied volatility, and the option price using the model parameters in the HN model.

Link to Code

Once Python code is able to produce similar results to Matlab, it will have an initial Github release, with this link. It will be ZIP file under the name `JC_WT_DataGen`

Alpha 1 Release