A Comparison Between Approximations of Option Pricing Models and Risk-Neutral Densities using Hermite Polynomials

Nathaniel Ahy

June, 2020

Outline

- Introduction
- Theory
- Option Pricing
- Model Fitting Results
- Option Data
- Data Fitting Results
- Conclusions

Hermite Polynomials

- Anders Hald gives a rich description of the history behind Hermite polynomials (2000)[1]
- Originated by Laplace (1810), Poisson (1829), Bessel (1838)
- Chebyshev (1859) looked for ways of polynomials orthogonal to a weight function for the sake of regression
- Hermite (1864) defined said polynomials as:

Definition 1

The physicist's Hermite polynomial $H_n(x)$ takes on the following form:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} = \left(2x - \frac{d}{dx}\right)^n \cdot 1,$$

where it can easily be verified that $H_0(x) = 1$, $H_1(x) = 2x$.

- Thiele and Gram (1870s) introduced these polynomials in finance

Financial Derivatives

- Black and Scholes (BS) [2] paper played a major role and popularizing option pricing models
- They assumed that the asset price, $S(t)$ followed a Geometric Brownian motion

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t) \tag{1}$$

- $\mu \in \mathbb{R}$ - the log mean return
- $\sigma \in \mathbb{R}^+$ - the log standard deviation
- W - a standard Wiener process

Heston Stochastic Volatility

- Many alternative asset pricing models have emerged to capture various characteristics their predecessors failed to do
- Among others, the Heston [3] model describes volatility as stochastic

$$d\nu_t = \kappa[\theta - \nu_t]dt + \sigma\sqrt{\nu_t}dB_t.$$

$$dS_t = \mu S_t dt + \sqrt{\nu_t}S_t dW_t.$$

Hermite Polynomials in Finance I

- Cheng [4] displays a method of using Gram-Charlier (GC) expansions to price financial derivatives
- The main application of Hermite polynomials in Cheng's paper is in representing the probability distribution of an asset in terms of its moments

Definition 2

The probabilist's Hermite polynomial $He_n(x)$ is defined as follows

$$He_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2} = \left(x - \frac{d}{dx}\right)^n \cdot 1.$$

- Jarrow and Rudd also used such series to extend the Black–Scholes model

Hermite Polynomials in Finance II

- Necula, Drimus, and Farkas (NDF) [5] use Hermite polynomials for the underlying asset's risk-neutral density
- They refer to this density as a Gauss-Hermite (GH) density
- The trio then constructs their own option pricing formula which, in a special case, resembles the Black–Scholes formula
- We will occasionally refer to the NDF and 'GH' approach synonymously and the Gram-Charlier approach will be referred to as the 'GC' approach

Problem Formulation

- Our goal is to see how well the GC and GH method can fit models and data in different scenarios
- The purpose of this is to illustrate the flexibility of Hermite polynomials

Fundamental Theorem of Finance

- Recall that in the Heston setting

$$\begin{cases} dS(t) = \mu S(t)dt + \sqrt{\nu(t)}S(t)dB(t) \\ d\nu(t) = \kappa[\theta - \nu(t)]dt + \sigma\sqrt{\nu(t)}dW(t) \end{cases}$$

- By using no-arbitrage arguments similar to those used in Black–Scholes, defining C as the price of a European call option, we get the Heston PDE

$$\begin{aligned} \frac{1}{2}\nu S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma\nu S \frac{\partial^2 C}{\partial S \partial \nu} + \frac{1}{2}\sigma^2 \nu \frac{\partial^2 C}{\partial \nu^2} + rS\frac{\partial C}{\partial S} \\ + \{\kappa[\theta - \nu(t)] - \lambda\}\frac{\partial C}{\partial \nu} - rC + \frac{\partial C}{\partial t} = 0. \\ C(S(T), T) = \max(S(T) - K, 0) \end{aligned} \tag{2}$$

  - ρ - the correlation between B and W
  - λ - the market price of volatility risk

Methodology

- Price call options given a set of parameters $\vec{p}$ under the Heston model using a closed form solution $C_H(K_i; \vec{p})$ for strike K_i where $i \in \{1, 2, ..., N\}$.
- Price options using the Gram-Charlier expansion approximation of the Heston model based on the first m moments of the asset price with the same set of parameters as Heston $C^m_{GC}(K_i; \vec{p})$.
- Given a set of parameters in the NDF setting, $\vec{p}_{NDF}$, calibrate the coefficients $\vec{a} \stackrel{def}{=} (a_j)^n_{j=0}$ for the NDF asset return density $P_{NDF}(x; \vec{a}, \vec{p}_{NDF})$ that minimizes the difference between the NDF option price, $C_{NDF}(K_i; \vec{a}, \vec{p}_{NDF})$, and $C_H(K_i; \vec{p})$.
- Equipped with $\vec{a}$ find $\vec{p}_{NDF}$ that best fits the NDF prices to the Heston prices.
- Compute the corresponding call options and their relative differences, repeat everything for different n and m

Gram-Charlier expansions

Definition 3

The *n*th cumulant, c_n of a random variable X is

$$c_n = \frac{d^n}{du^n} \ln\left(E[e^{uX}]\right)\Big|_{u=0}.$$

Theorem 4

A random variable X with distribution f has the Gram-Charlier series

$$f(x) = \sum_{n=0}^{\infty} \frac{q_n}{\sqrt{c_2}} He_n\left(\frac{x-c_1}{\sqrt{c_2}}\right) \phi\left(\frac{x-c_1}{\sqrt{c_2}}\right), \tag{3}$$

where ϕ is normal, $q_0 = 1$, $q_1 = q_2 = 0$, and

$$q_n = \sum_{m=1}^{\lfloor \frac{n}{3} \rfloor} \sum_{\substack{k_1, k_2, \ldots, k_m \\ k_1 + k_2 + \ldots + k_m = n}} \frac{c_{k_1} \ldots c_{k_m}}{m! k_1! \ldots k_m! \sqrt{c_2}^n}, \quad n \geq 3. \tag{4}$$

Gram-Charlier Expansion Option Pricing formula I

- We define $e^{X_T} = S_0 e^{-rT}$, $e^k = Ke^{-rT}$ where T is the time to maturity of an option and r is the risk-free interest rate, K is the strike price
- Then the price for a European call C is given by

$$C = E^Q[(e^{X_T} - e^k)^+] = E^Q[e^{X_T} \mathbb{I}\{X_T \geq k\}] - e^k E^Q[\mathbb{I}\{X_T \geq k\}] \tag{5}$$

  - $(\cdot)^+ = max(\cdot, 0)$

- Using Theorem 4 we display the Gram-Charlier option price for the Heston model in the following slide

Gram-Charlier Expansion Option Pricing formula

Theorem 5

A European call, C_H, in the Heston setting priced with a Gram-Charlier A series representation of the asset price density, with strike K, and asset cumulants $(c_i)_{i=1}^{\infty}$ can be written in the form of the following infinite series:

$$\begin{aligned} C_H = {} & e^{c_1} \sum_{n=0}^{\infty} q_n J_n \left(\frac{k - c_1}{\sqrt{c_2}}, \sqrt{c_2} \right) - e^k N \left(\frac{c_1 - k}{\sqrt{c_2}} \right) \\ & - e^k \sum_{n=3}^{\infty} (-1)^{n-1} q_n He_{n-1} \left(\frac{c_1 - k}{\sqrt{c_2}} \right) \phi \left(\frac{c_1 - k}{\sqrt{c_2}} \right) \end{aligned} \tag{6}$$

$$J_n(x, a) = \begin{cases} e^{a^2/2} N(a - x), & n = 0, \\ a J_{n-1}(x, a) + He_{n-1}(x) \phi(x) e^{ax}, & n = \{1, 2, \ldots\}. \end{cases}$$

Necula, Drimus, and Farkas' Risk-Neutral Measure

- Necula, Drimus and Farkas defined the log return probability density of an underlying asset S with log-mean and standard deviation μ and σ respectively as:

$$p(x) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\sum_{n=0}^{\infty} a_n H_n\left(\frac{x-\mu}{\sigma}\right) \tag{7}$$

- The function $\phi(x)$ is the standard Gauss density and H_n is the nth order physicist's Hermite polynomial
- By the above points the aforementioned trio get the option pricing theorem in the following slide by performing the following integration to find $c(S_0, K, \mu, \sigma, T, r, q) \stackrel{def}{=} c$:

$$\int_{-\infty}^{\infty} (S_T - K)^+ p_T(S_T) dS_T = \int_{-\infty}^{\infty} (S_0 e^{\mu T + \sigma\sqrt{T}x} - K)^+ p(x) dx$$

Necula, Drimus, and Farkas' Option Pricing Formula

Theorem 6

The price of a European call, c, with strike K and maturity T is given by:

$$c = S_0 e^{-qT}\Pi_1 - Ke^{-rT}\Pi_2, \tag{8}$$

where q is the dividend and

$$\begin{cases} \Pi_1 = \exp\left\{\left(\mu - (r-q) + \frac{\sigma^2}{2}\right)T\right\} \sum_{n=0}^{\infty} a_n I_n \\ \Pi_2 = \sum_{n=0}^{\infty} a_n J_n. \end{cases}$$

The I_n and J_n terms satisfy recursive formulas:

$$I_{n+1} = 2z(-d_1)H_n(-d_2) + 2\sigma\sqrt{T}I_n + 2nI_{n-1}, \qquad n \in \{1, 2, ...\}$$
$$J_{n+1} = 2z(-d_2)H_n(-d_2) + 2nJ_{n-1}, \qquad n \in \{1, 2, ...\}.$$

Necula, Drimus, and Farkas' Option Pricing Formula (continued)

- We have that $I_0 = N(d_1)$, $I_1 = 2z(-d_1) + 2\sigma\sqrt{T}N(d_1)$, $J_0 = N(d_2)$, $J_1 = 2z(-d_2)$, finally

$$\begin{cases} d_1 = \frac{log(S_0/K)+(\mu+\sigma^2)T}{\sigma\sqrt{T}} \\ d_2 = d_1 - \sigma\sqrt{T}. \end{cases}$$

- The NDF paper also includes an explicit formula for inferring the coefficients $(a_j)_{j=0}^{\infty}$ from option prices, although our implementation of this method yielded a relative error of 470% between the actual option prices and NDF's model!

Estimating Coefficients through Quadratic Programming

- We want to find matrices L, f, A, and A_{eq} as well as vectors b and b_{eq} so that the problem of finding the risk-neutral density coefficients, a, that best fit the targeted option price $C(K_k)$ for strikes $(K_k)_{k=0}^{N}$ corresponds to:

$$\min_a \frac{1}{2} a^\top L a + f^\top a \text{ subject to } \begin{cases} A \cdot a \leq b \\ A_{eq} \cdot a = b_{eq}. \end{cases} \tag{9}$$

- Denoting the NDF call price as $C_{NDF}(K_k; a, \mu, \sigma)$ our optimization problem is:

$$\min_a \frac{1}{2} a^\top L a + f^\top a = \min_a \sum_{k=0}^{N} (C_{NDF}(K_k) - C(K_k))^2. \tag{10}$$

- By applying Theorem 6 to (10) and observing the quantity of a-coefficients involved in each term we are able to derive matrices L and f

Quadratic Programming Constraints

- The constraint matrices of (9) are intended to ensure that we select $(a_j)_{j=0}^n$ such that the risk-neutral density in (7), $p(x)$, is a probability density:
  - $p(x) \geq 0$ for all x
  - $\int_{-\infty}^{\infty} p(x)dx = 1$
  - By ensuring $p(x_j) \geq 0$ for $-3m = x_0 < x_1 < ... < x_M = 3m$ where m is the degree of the Hermite polynomial of the risk-neutral density we roughly satisfy the first constraint
  - Necula, Drimus and Farkas claim the identity: $\sum_{j=0}^{\lfloor \frac{n-1}{2} \rfloor} a_{2j} \frac{(2j)!}{j!} = 1$ ensures the second constraint

Optimizing model parameters

- With all the necessary matrices of (9) found we solve the quadratic programming problem using the active-set algorithm described by Nocedal and Wright (2006) [6].
- We define $(\mu, \sigma) = \vec{p}$ and having found the *a*-coefficients now want to optimize:

$$\min_{\vec{p}} \sum_{i=0}^{n} (C_{NDF}(K_i; \vec{p}) - C_{Data}(K_i))^2.$$

- This problem is solved using the trust-region reflective non-linear least squares algorithm
- We do the same to find the parameters v_0, κ, θ, ρ, and σ in the Heston and GC models as well as σ in the Black–Scholes model

Parameter selection

- Using the same set of parameter values as those used in Cheng's paper we price Heston call option prices $C_H(\vec{p}_H)$ where $\vec{p}_H = (\nu_0, \kappa, \sigma, \rho)$
  - initial volatility - $\nu_0 = 0.03$
  - mean reversion - $\kappa = 0.15$
  - volatility of volatility - $\sigma = 0.05$
  - correlation between variance and asset's Wiener processes - $\rho = -0.55$
  - time to expiration - $T = 1$
  - annual log return of risk-free asset - $r = 0.04$
  - initial asset price - $S_0 = 100$
  - strike price for a European call option - $K \in \{40, 40.001, ..., 170\}$,

Cumulant Computation

- Recall the definition of cumulants; Definition 3, i.e. the log moments of our asset price in the Heston setting described in the previous slide.
- Using Mathematica's online platform and MATLAB as well as an explicit identity for the moment generating function from Utzet et. al. [7] we compute the cumulants to be:

i	c_i
1	4.589456320893087
2	0.031838851938017
3	-0.001261815533270
4	0.000115992103852
5	-0.000011185767429
6	0.000001414417922
7	-0.000000200913900
8	0.000000027458226

Model Terminology

- We consider $m \in \{4, 6, ..., 14\}$ degree polynomial for the Gauss–Hermite risk-neutral density behind NDF's option pricing method and $n \in \{3, 4, ..., 8\}$ cumulants (recall the q_n-terms derived in (4) for the Gram-Charlier option pricing formula).
- We make the following defintions:

  - $C^m_{NDF}(K)$ - the option price in the NDF setting with an mth degree Gauss-Hermite polynomial as the risk-neutral density (7).

  - $C^n_{GC}(K)$ - the Gram-Charlier option price using only n-terms from its infinite summation.

  - $C_H(K)$ - option price under Heston's model.

Relative Error

- We evaluate the performance of a model using the relative error metric

Definition 7

The adjusted relative error R for an option pricing method $C^j_{method}(K_k)$ targeting the price $C_{target}(K_k)$ is computed as:

$$R(C^j_{method}(K_k), C_{target}(K_k)) = \frac{|C^j_{method}(K_k) - C_{target}(K_k)|}{max(C_{target}(K_k), 0.01)}$$

Model Errors

- Upon pricing we realized that in-sample and out-of sample errors hardly differed, we therefore only display the out-of sample (test) errors

NDF Test Error	GC Test Error	m	n
0.495763409124678	0.108409844389963	4	3
0.0728483878146307	0.0298546397931812	6	4
0.0780088696974434	0.0350818888644778	8	5
0.0388685528271606	0.0810631455081564	10	6
0.151286360809425	0.0591659254461921	12	7
0.157431617652327	0.0762545916270493	14	8

Table: Our Gram-Charlier approximations for the Heston model using the first n cumulants

Model Errors

- In spite of the relative errors being considerably large our methods still do a visually satisfying job of approximating the Heston option prices:

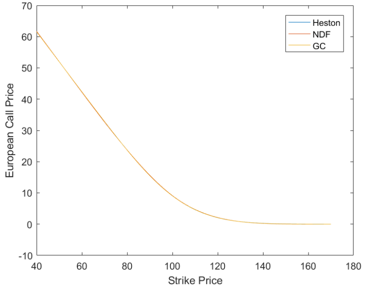

Figure: Option Prices for Various models, m=14, n=8

Option Data

- We consider call and put options written on Amazon stock on 2/11/2018 expiring at 21/6/2019 with parameters:
  - $S_t = 1665.53$ - Amazon asset price at birth of option
  - $T = 231/365$ - Lifespan of option[1]
  - $r = 0.027$ - The risk-neutral rate
  - $m \in \{2, 4, ..., 16\}$ - Degree of polynomial in GH distribution
  - $n \in \{3, 4, 5, 6\}$ - Number of cumulants used in GC option price
- This time we will be comparing the NDF model, Gram-Charlier approximation of the Heston model, the pure Heston model, and the Black–Scholes model
- We fill the gaps between observations of data for different strike price using cubic splines

[1] We are considering a full year as opposed to a trading year in this case.

Necula, Drimus, and Farkas' Model Relative Errors, I

- Using the active-set and trust-region reflective methods we achieved the following relative errors for *m*-degree Hermite polynomials.

NDF Test Error	m
0.0713333716082011	2
0.0693943514826509	4
0.00909296454191811	6
0.00914694076555854	8
0.00728423120119476	10
0.00644105230738876	12
0.0064934167710657	14
0.00646823680508747	16

Necula, Drimus, and Farkas Model Relative Errors

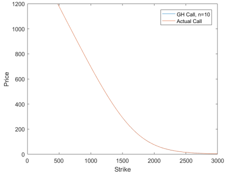

Figure: NDF option prices plotted with actual option data, $n = 16$

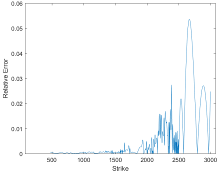

Figure: NDF relative error, $n = 16$

Black–Scholes Model Relative Errors

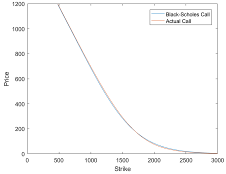

Figure: Black–Scholes option prices plotted with actual option data.

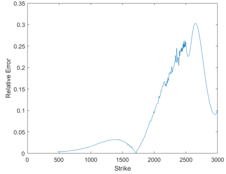

Figure: Black–Scholes model relative error (averaging at 0.09)

Pure Heston Model Relative Errors

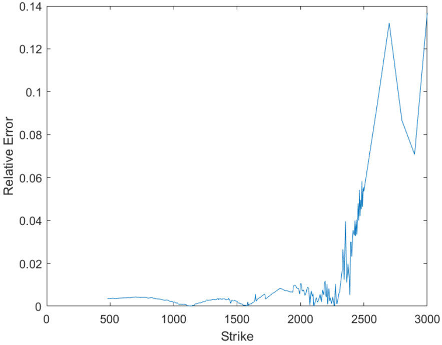

Figure: Heston model relative error (averaging at 0.0098)

- As the Heston model is nearly as accurate as the NDF model its comparison with actual option prices looks exactly like NDF's

GC Heston Model Approximation Relative Errors

- The Gram-Charlier approximations for the Heston model were computed using $n \in \{3, 4, 5, 6\}$ cumulants producing the table below

GC Test Errors	n
0.2198345379137871	3
0.259549772655078	4
0.451540429338985	5
0.682847026715799	6

GC Heston Model Approximation Relative Errors

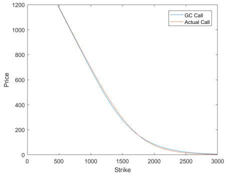

Figure: GC option prices plotted with actual option data, $n = 3$

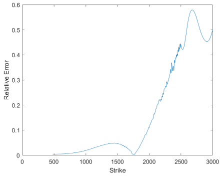

Figure: GC relative errors, $n = 3$

Gram-Charlier Approximation Algorithm Speed

GC Time (seconds)	n
11.5413138000000	3
17.6311419900000	4
39.4671034100000	5
112.610722900000	6

Cumulant time (seconds)	n
0.172281400000000	1
0.488967600000000	2
1.66210270000000	3
6.06789310000000	4
21.6428043000002	5
73.2235639000002	6

NDF, BS, and Pure Heston speed

- BS time:0.630270000000000
- Heston time: 12.7383890999999999

NDF time (seconds)	n
0.325923466666667	2
0.9274233	4
0.4394605	6
0.812202066666667	8
0.570558	10
0.702097066666667	12
0.9264461	14
1.23152356666667	16

Conclusion

- We described and applied Hermite polynomials for approximating a probability density (Gram-Charlier) and to construct a risk-neutral density (Gauss-Hermite)
- The Gram-Charlier approximation outperformed the Gauss-Hermite fit to the Heston model with a small margin
- We thereafter applied these models along with their more well-known counterparts to data-fitting
- The NDF model performed the best followed by the Heston, Black–Scholes, and the Gram-Charlier models in that order
- The NDF and Black–Scholes were the fastest due to the duration of the non-linear least squares algorithm and the respective models smaller parameter space in comparison to the alternatives applied to this algorithm

Further Studies

1. There is a wide range of alternative option pricing and volatility models which could have been explored
2. It would be interesting to see how well these models would fit to various other options
3. An interesting extension to the NDF model would be one similar to that Jarrow and Rudd applied to the Black–Scholes model

Bibliography

[1] A. Hald, "The early history of the cumulants and the gram-charlier series," *The Review of Financial Studies*, vol. 68, no. 2, pp. 137–153, 2000.

[2] F. Black and M. Scholes, "The pricing of option and corporate liabilities," *The Journal of Political Economy*, vol. 81, no. 3, p. 637–654, 1973.

[3] S. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The Review of Financial Studies*, vol. 6, no. 2, pp. 327–343, 1993.

[4] Y. H. Cheng, "Pricing derivatives by gram-charlier expansions," 2013.

[5] C. Necula, G. Drimus, and W. Farkas, "A general closed form option pricing formula," *Review of Derivatives Research*, vol. 22, p. 1–40, 2017.

[6] J. Nocedal and S. Wright, *Numerical Optimization*. Springer, 2006.

[7] S. Del Baño Rollin, A. Ferreiro-Castilla, and F. Utzet, "A new look at the heston characteristic function," *Mathematics Subject Classification*, 2009.

Active-set Algorithm I

- A basic description of this algorithm is we consider a working region $\mathcal{W}_\ell$ for some iteration ℓ, where all equality constraints and a subset of the inequality constraints are considered as equalities
- Find the optimal step, x, for iteration ℓ

$$\min_x \frac{1}{2} x^\top L x + g_\ell^\top x \tag{11}$$

$$\text{subject to } A_i^\top x = 0, \quad i \in \mathcal{W}_\ell. \tag{12}$$

Active-set Algorithm II

- Letting λ be a vector of Lagrange multipliers we can represent the equation as follows

$$\begin{pmatrix} L & -A_i^\top \\ A_i & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}, \tag{13}$$

- We solve the above matrix equation and update our working region until we have reached an optima according to our preferences
- We now want to use the coefficients $(a_j)_{j=0}^n$ to find the underlying asset's log-mean return and standard deviation, μ and σ from Theorem 6 respectively

Non-linear least squares I

- The non-linear least squares algorithm works by taking a function $q(\vec{p})$ that approximates the objective function above by conducting a second order Taylor approximation reducing our problem to

$$\min_{s} \frac{1}{2} s^\top H s + s^\top g \quad \text{subject to } ||Ds||_2^2 \leq \Delta, \tag{14}$$

- H is the Hessian, g is the gradient, D is a diagonal scaling matrix and s is our parameter set within the current 'trust region'

Non-linear least squares II

- To solve (14) the space of s-values is reduced to the two-dimensional space $(s_1, s_2) \in \mathcal{S}$.
- The value s_1 is the direction of the gradient and s_2 satisfies one of the two relations:

$$\begin{cases} Hs_2 = -g. \\ s_2^\top Hs_2 < 0, \end{cases}$$

- We proceed by either finding the next step that reduces the value of our objective function or shrinking the trust region, Δ and repeating.
- The same method is used to find the parameters of the Gram-Charlier approximation of the Heston model, the pure Heston model's parameters, as well as the only free parameter of the Black–Scholes model; the log standard deviation, σ