Modeling Joint Probability Distributions (GS Quantify 2016 Problem 3)

Tushant Jha, Ameya Prabhu and Sneh Vyas
Team - Category Killers

October 16, 2016

1 Introduction

The posed problem requires estimation of the price for certain exotic derivatives, given the option prices for a sequence of strike prices. The given sequence of prices of European-style options against respective strike prices carries information about the underlying probability distribution, that encapsulates the a priori knowledge about the value of underlier at the expiration date. Insofar as we could interpret, the central challenge posed is to use this encapsulated information in the option prices to evaluate prices of given exotics.

The values are given for the derivatives where the payoff is simply the difference of the index price and strike price, if the transaction is indeed exercised, otherwise providing zero returns. The payoff in the first task, ie. Digital Options, provide unit returns if the transaction is exercised, otherwise zero. And the second task deals with quadratic payoff functions, where the square of the difference of index and strike denotes the price when the transaction is exercised.

The third task, however, is further challenging that it requires the estimation of a composite financial derivative with two underlying indices. The information about the relevant joint distribution is encoded both in the independent option prices given for those indices, as well as two other sparse sequences of prices for composite payoffs.

In Section 2, we study some mathematical properties of the pricing functions, that can be directly deduced from their mathematical definitions, without any further assumptions on the shape of distributions. We show how to effectively gain knowledge of the distribution from the given sequence of prices; and also show, at the same time, how given sequence is sufficient to directly compute the values required in first two tasks.

We use Excel computations to perform some rectangular integration and first-order finite differences to gain the output for first two tasks. In the document we outline the broader approach towards third task.

2 Some Mathematical Properties of the Univariate Derivatives

First we note that for purposes of estimation, the expiration time is irrelevant, since the problem does not deal with the temporal aspects of the value evolution of the underlying index. Also, the present value of the index is not visible, insofar as the data is provided in the problem.

For all practical purposes, we are making statements about the future value of some random variable, for which there exists some probability distribution that is reflected in the options prices provided. The cost functions, therefore, of the different financial derivatives, are functions of the strike prices. Borrowing the notation from problem statement, let that future value be denoted by X_T, and the probability distribution function (pdf) by f_X. The option price, as a function of strike price, can be simply denoted by $C(K)$, and let the cumulative distribution function be denoted by F_X.

We observe (as in Eqn 2), that the Cumulative Distribution function can be extracted by differentiation of the given option price function against the strike prices. We can further breakdown the computation digital option price and quadratic payoff option price simply in terms of the provided option price function, **agnostic of any structural assumptions on the distribution** for X_T, as demonstrated in Eqn 3 and Eqn 6.

2.1 'Usual' Options Price

From the given definition for the option price, we find

$$\begin{aligned} C(K) &= \int_0^\infty max(X_T - K, 0) f_X(X_T)\,\mathrm{d}X_T \\ &= \int_K^\infty (X_T - K) f_X(X_T)\,\mathrm{d}X_T \end{aligned} \tag{1}$$

Applying Leibniz's integral rule, we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}K} C(K) &= \int_K^\infty -f_X(X_T)\,\mathrm{d}X_T \\ &= -(1 - F_X(K)) = F_X(K) - 1 \end{aligned} \tag{2}$$

2.2 Digital Options

In similar terms to what has been given, the pricing for Digital Options can be given by:

$$\begin{aligned} D(K) &= \int_0^\infty \mathbb{I}(X_T > K) f_X(X_T)\,\mathrm{d}X_T \\ &= \int_K^\infty f_X(X_T)\,\mathrm{d}X_T \\ &= 1 - F_X(K) = -\frac{\mathrm{d}C}{\mathrm{d}K}(K) \end{aligned} \tag{3}$$

This also implies that an estimation of $D(K)$ can be made by approximating differentiation of $C(K)$, for instance by **method of finite differences**.

2.3 Quadratic Payoff

The pricing for quadratic payoffs, as earlier, can be defined in the following manner.

$$\begin{aligned} P(K) &= \int_0^\infty (max(X_T - K, 0))^2 f_X(X_T)\,\mathrm{d}X_T \\ &= \int_K^\infty (X_T - K)^2 f_X(X_T)\,\mathrm{d}X_T \end{aligned} \tag{4}$$

Again, by Leibniz's integral rule

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}K} P(K) &= \int_K^\infty f_X(X_T)\,\frac{\partial}{\partial K}(X_T - K)^2\,\mathrm{d}X_T \\ &= \int_K^\infty -2(X_T - K) f_X(X_T)\,\mathrm{d}X_T \\ &= -2C(K) \end{aligned} \tag{5}$$

We can move further, by making a very small, and natural, structural assumption that $\lim_{x\to\infty} f_X(X_T) = 0$, ie. the probability of X_T being arbitrarily large vanishes.

$$\int_{K_0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}K} P(K)\,\mathrm{d}K = \lim_{x\to\infty} P(x) - P(K_0)$$
$$P(K_0) = 2\int_{K_0}^{\infty} C(K)\,\mathrm{d}K \tag{6}$$

As with $D(K)$, similarly from here we see that for $P(K)$ too, we can make an estimation by **numerical integration, by some quadrature rule**, of $C(K)$.

3 Towards some Structural Assumptions

By methods established earlier, we can plot the estimates derived from first order differences.

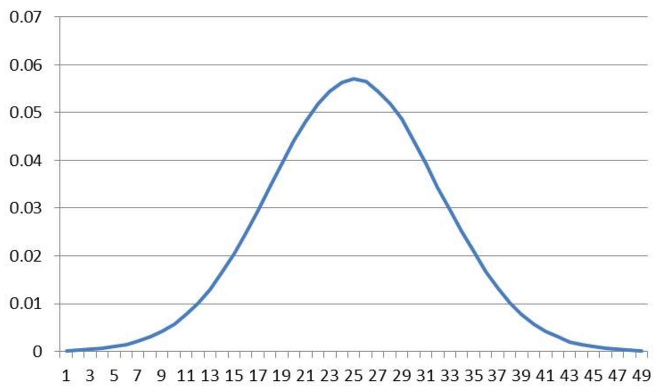

Figure 1: Graph of the PDF for Oil index, as computed from given option prices

We can claim, therefore by observation, that the probability distribution can be modeled by a normal distribution.

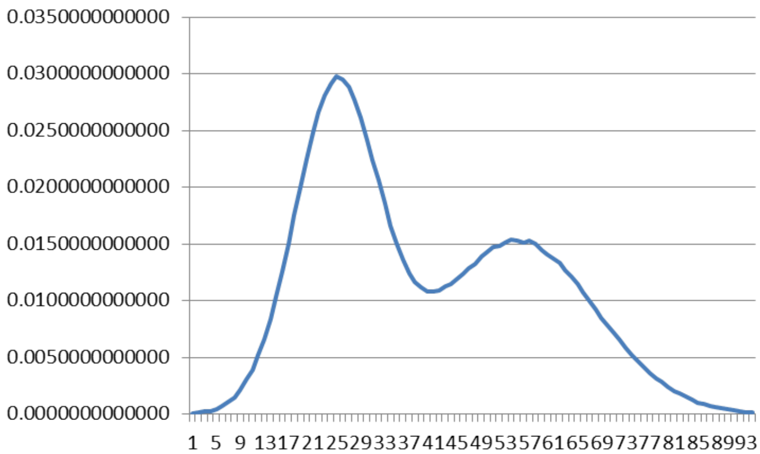

Figure 2: Graph of the PDF for Exchange Rate index, as computed from given option prices

Similarly, we claim that the distribution of Exchange Rates can be modeled by a mixture of two normal distributions.

Consequently we can also assume, though with some loss of generalization, that the joint distribution is also a mixture of two bivariate normal distributions.

4 Methodology for Modeling Joint Distributions

While, unlike earlier, we cannot compute the exotic price directly in analytic terms of the given prices, as a consequence of absent knowledge of covariances, we can still use numerical approximation and simulation. Having extracted the self-variances and means of the component bivariate distributions, we can now use Markov Chain Monte Carlo to compute the relevant prices by of Q_1 and Q_2, as given in the data, for given covariances of in the components.

The error in the simulated values of Q_1 and Q_2, as a function of the covariance parameters in the gausssians, against the given values, will form a convex search space. And therefore, we can perform any local search, like Tabu search, to reach a reasonable approximation for the covariances. This can be further used to compute the desired exotic.

However, due to our failures with time and delays, we have not done that.