FX Final Preperation

October 13, 2019

1 FX final preparation

Author: Chenyu Zhao

Date: Oct 13, 2019

1.1 Spot

1.1.1 OTC voice business model & roles

Traditional OTC market structure:

1. Clients: humans who are not dealers who want to trade FX
   - Corporates, hedge funds, pension and sovereign wealth funds ("real money"), smaller banks without global FX desks, retail channels
2. Salespeople: humans who talk to human clients on the phone (or over chat channels like Bloomberg) and take client requests for trades
3. Traders: humans who talk to human salespeople in response to a client trade request
   - Make markets to clients, based on where they can hedge, what their risk position is, and who the client is
   - Manage market risk that comes from taking the other side of client trades

Inter-dealer market for traders to trade with each other

1. Dealers have a privileged position in the market

Traders are market makers

1. They execute at a pre-agreed price with a client and take the market risk on the other side of the trade
2. Trader looks at the
   - inter-dealer market;
     - Market where a dealer can hedge (if desired)
   - Current risk position;
     - Already long, bias down prices; already short, bias up

- Market views
- Client behavior
  - Does the client typically buy vs sell? Is the client typically right about market direction?
  - Does the client tend to ask more dealer and take the best price
  - Ability to predict (short-term) direction of the market V.S. Price sensitivity

Clients are market takers

1. Can leave orders but those are seen only by the dealer holding the order, not the broader market
2. ABC trader calls Dealers B and C and checks the pricing and deals on the lowest offer price across the dealers

steps

1. Client calls a dealer and asks for bid and ask prices for a specific currency in a specific notional.
2. Salesperson takes the client call, and yells over to the spot trader for that currency, asking for the two-way price in the size the client wants, and tells the trade the client name as well.
3. Trader looks at the inter-dealer market, his risk position, his views on the short-term market direction, and historical trading behavior of the client, and shouts bid and ask prices back to the salesperson.
4. The salesperson relays the prices back to the client, who decides whether she wants to buy at the offer or sell to the bid. If so, she tells the salesperson so, and the trade is done.

1.1.2 OTC electronic business model & roles

Single-dealer platforms: electronify voice dealing

1. Client requests a price through an app, can click to deal
2. Still a price from just one dealer

Multi-dealer platforms: electronic auctions

1. Client requests a price from multiple dealers simultaneously through an app, who know who the client is when they quote
2. Client sees prices from many dealers at once, by name, and can click to deal with whomever they like

Inter-dealer trading has also moved electronic

1. ECNs (Electronic Communication Networks) are the machines they use to trade
2. Inter-dealer bid/ask spread is often wider than the bid/ask spread shown to clients to be competitive

steps

1. Client opens a trading application for a specific dealer on their computer, finds the screen for the currency she wants to trade and enters the notional she wants to trade.
2. The price-making engine at the dealer recognizes the request for quote and looks at the inter-dealer electronic market, the risk position of the electronic book, includes any short-term market direction bias (either manually-specified or determined algorithmically), and includes any bias due to historical trading behavior of the client. It calculates bid and ask prices and streams them back to the client, ticking them as those four inputs change through time.
3. The client sees the streaming prices in their application and, if she wants to trade, clicks to buy from the offer or sell to the bid.
4. The dealer receives the trade request and executes the trade (retaining the option to reject the trade request if the market has moved sufficiently far).

1.1.3 Basics of electronic price-making algos & hedging algos

Computers at dealers now make prices automatically to clients via single dealer and multi-dealer platforms major concerns are same with real trader 1. inter-dealer market: aggregated across ECNs - aggregated across ECNs - ECNs show resting bids and offers which define the inter-dealer marketplace - Dealers have electronic connections to ECNs and stream in the market data 2. Current risk position - How do you know how much to "skew" your bid and/or offer when you have a net risk position? - Real answer depends on market microstructure - Usually dealers do not bother with microstructure models, Ad hoc algorithm for moving bid and/or offer as a function of net risk position 3. Market views: manually specified or algorithmic - The time scale of a view on the market is approximately the average risk holding time - Economics and fundamentals don't matter on those time scales, just technical signals - Momentum signals; Mean reversion signals; Directional signals based on order book structure 4. Client behavior: historical profitability of trades against the client over various interesting time intervals - Two types of client behavior matter - Ability to predict (short-term) direction of the market - Not always because the client is good – sometimes just because they are big and move the market - Sometimes due to latency arbitrage - Price sensitivity - Some clients deal with many dealers and are always putting prices in competition on every trade - Others trade with only a small number of dealers and try to spread out their trades between dealers

Electronic Hedging market evolution 1. Voice trading market 2. Clients can click to trade, but humans set prices and manage portfolio risk 3. Price-making is automated, but humans manage portfolio risk 4. Hedging is automated as well; humans monitor the machine but do not get involved in individual trades

Two extremes of risk management 1. Don't hedge at all - Take market risk - Get paid bid/ask spread on each client trade - Client trades hopefully net against each other so get paid the spread to close out risk 2. Hedge every trade - No material market risk - Get paid bid/ask spread on each client trade - Pay inter-dealer bid/ask spread to hedge each client trade

Inter-dealer bid/ask spread is often wider than the bid/ask spread shown to clients to be competitive

Typical approach taken by modern dealers: break up positions into two buckets 1. Bucket 1: trades against clients who are generally right on the direction of market moves - Hedge their trades aggressively - what does it mean - Time scale is determined by average risk holding period

- Clients who hold risk positions over much longer periods (eg global macro traders) are not relevant here - Clients who are right on short timescales are relevant - High frequency traders, - big names whose trades move the market

2. Bucket 2: trades against everyone else
   - Do not hedge each trade, and hope for netting
   - Monitor net risk position and hedge when a risk limit is exceeded

   Use inter-dealer market to hedge, pay wider spread

1.2 Forwards

1.2.1 Spot vs forward arb and how it works in practice

FX is a financial market where the two requirements for spot/forward arbitrage hold

1. You can store currencies (and receive an interest rate for them)
2. You can borrow & short currencies (and pay an interest rate)

$$V(t) = S(t)e^{-Q(T-t)} - Ke^{-R(T-t)}$$

Q: asset, R: denominated, USD

$$F(t,T) = S(t)e^{(R-Q)(T-t)}$$

example

1. forward settling in 1y, receiving 1M EUR, paying 1.13M USD
2. Equivalent to being long 1 unit of a zero coupon bond in EUR with a 1y settlement date and being short 1.13 units of a zero coupon bond in USD

1.2.2 Why FX forwards businesses are really short-term rates businesses

1. An FX forwards portfolio might contain forwards settling on many different dates

2. Usually represent risk to benchmark dates only, since those are the easiest points to trade in the inter-dealer market on the hedge

3. What is the risk to?

   - Forward points: the difference between all-in forward prices and the spot price
   - Or equivalently, non-USD interest rates, but risk displayed in notional of FX swap equivalents
   - Forwards traders also always monitor delta (risk to spot) but hedge that with their friends on the spot desk

there are three dimensions of risk in a forward point: spot S, and the two interest rates (of which one is the USD rate). Other risk reports already report risk to spot and the USD interest rates. So the one dimension of risk that's special to FX forward points is the risk to the non-USD interest rate curve.

1.2.3 Ways of reducing dimensionality of risk

Want a way to efficiently reduce the dimensionality of the market risk(currency rate minus USD rate)

Two common approaches:

1. Principal component analysis
2. Parametric factor models

PCA

1. Look for most important (non-parametric) shocks that tend to drive moves in the whole curve
2. Use historical daily interest rate spread moves, particularly covariances between them, to figure out the factors
3. In practice people do not use this very much
   - Non-parametric shocks are hard to understand properly
     - Can have unusual shapes due to specific data points in the history you're using
   - Non-parametric shock shapes change over time
     - Do you fix the shock shape based on a particular historical run?
     - Do you recalculate the shock shapes every day based on rolling historical runs?
   - Together they make traders unsure what their risk numbers mean

Factor models We'll focus on forward curve models

$$dS(t,T) = \sum_{i=1}^{N} \sigma_i(t,T)dz_i(t)$$

S is term interest rate spread for tenor T as seen at time t

Can compare the model covariances with historical covariances to calibrate

The two factor model can give shocks that look a lot like the first two principal component shocks

compare this two methods Principal component analysis and factor models are two different techniques for reducing dimensionality in a high-dimensional system. In this example we were looking at these techniques as ways of reducing the dimensionality of non-USD interest rate curve hedging for a book of FX forwards.

Principal component analysis is a non-parametric method for discovering the main factors driving the moves in the interest rate curve. It involves calculating the historical covariances of day-on-day changes in interest rates of different tenors (assuming that each day's sample for a given tenor is drawn from the same distribution on each day of the historical period), then orthogonalizing that covariance matrix. The eigenvectors represent the orthogonal shocks to the curve, and the eigenvalues represent the relative sizes of those shocks. Principal component analysis involves picking off the top N largest eigenvalues and approximating all moves in the curves as being linear combinations of the respective eigenvectors. The shapes of the shocks are an output of the calculation and are non-parametric.

A factor model assumes a stochastic differential equation driving interest rates of different tenors, where all tenors are shocked by a small number of Brownian motion factors. The shocks in a factor model are parametric: their functionality forms are determined by the model parameter (for example, exponential shocks).

The main difference between the two approaches is that PCA results in non-parametric shocks, whereas a factor model has parametric shocks. Non-parametric shocks give more flexibility in terms of shape of the shocks, but also less robustness (historical data artifacts can lead to odd shapes). Parametric shocks are more robust and easier for traders to understand, but have less flexibility that PCA's non-parametric shocks.

1.3 Options

1.3.1 Conventions (delta, risk reversal, butterfly)

delta

1. Black-Scholes delta using the implied volatility for the strike
2. Usually "spot delta" – delta to spot in the BS delta, including discounting
3. Rarely "forward delta" – delta to forward in the Black delta, not including discounting (notional of forward to use on hedge)
    - Sometimes used for long-dated options

risk reversal Implied volatility skew is traded as a separate asset in FX options markets, and is called the "risk reversal"

1. 25-delta risk reversal means two things:
    - An implied volatility spread: the 25-delta call implied volatility minus the 25-delta put implied volatility
        - A measure of skew
    - An option spread: long a 25-delta call option and short the same amount of a 25-delta put option
        - A position in the option market, not a skew measure
2. Similarly for 10-delta risk reversal
    - 10-delta and 25-delta are the two liquid benchmarks

butterfly Implied volatility smile is traded as a separate asset in FX options markets, and is called the "butterfly" Sometimes "smile margin"

1. 25-delta butterfly means two things:
    - An implied volatility spread: the average of the 25-delta call and put implied volatilities less the ATM volatility
        - A measure of smile
    - An option spread: long 25-delta call and put options and short some notional (2x, vega-neutral are common variations) of the ATM option
        - A position in the option market, not a smile measure
2. Similarly for 10-delta butterfly

1.3.2 Risk reversal beta

1. Moves in risk reversals have a relatively high correlation with moves in spot
   - This matters a lot for barrier option pricing, as we will see when we talk about exotic derivatives
2. This is often quantified as the "risk reversal beta", or the slope of a linear regression of day-on-day risk reversal change against spot log return
   - A number like 0.2 means "risk reversal gets more positive by 0.2 vols for every 1% move up in spot"

The risk reversal beta is a measure of the covariance between moves in risk reversal and spot returns. A number like "0.1" means that if spot moves up 1%, the expected move in the risk reversal is 0.1 vols.

The risk reversal beta can be determined by running a historical regression of daily moves in risk reversal (in vols) against daily spot returns (in %).

Note that the risk reversal beta is a curve – you need to specify what tenor of risk reversal you are regressing against spot. Typically risk reversal betas are larger for shorter tenors.

1.3.3 Vol interpolation in the strike direction

The inter-dealer market is the source for implied volatility market data 1. It gives you implied volatilities for benchmark expiration tenors - 1d, 1w, 2w, 1m, 2m, 3m, 6m, 9m, 1y, 2y, maybe 3y-10y 2. And for five benchmark deltas on each expiration - 10d put, 25d put, ATM, 25d call, 10d call

we can interpolate imp-vol to different tenors and strikes

What are the requirements

1. $dC/dK < 0$: call prices decrease as strike increases
2. $d^{2C/dK}2 > 0$: call price curvature wrt strike is positive

Two main classes of volatility interpolation:

1. Model-based fits
   - like SABR, Heston
2. Non-model fits
   - like cubic spline

1.3.4 Vol interpolation in the time direction & trading time/theta

Volatility interpolation in the time direction assumes some form for the instantaneous volatility and fits it to the market implied volatilities

1. Piecewise-constant instantaneous volatility is a common choice
2. A mean-reverting form is sometimes used

$$\sigma_I^2(T)T = \sum_{i=0}^{N_T} \sigma_i^2(t_i - t_{i-1})$$

σ_I is implied volatility to time T, σ_i is the ith piece of the piecewise-constant instantaneous volatility

This formula enable us to calculate σ_N

requirement The arbitrage in the time direction is weaker: it is hard to define an explicit portfolio that is a true arbitrage when there is skew and smile in the market (you can do it in a pure Black-Scholes world, however).

There you must guarantee that there is no "negative forward variance" – that the implied variance as a function of time to expiration T, $\sigma^2(T)T$,must be always increasing with T. Here, is the implied volatility to time T.

Trading Time/Calendar Time

1. Calendar time is used to quote implied volatilities
   - ie you get the market option price by passing the implied volatility into the Black-Scholes formula along with a time measure in calendar time
   - However, variance does not increase smoothly in calendar time
     - weekend, holiday
     - event
2. Best way to think about this is to translate from calendar time to "trading time"
   - Assume variance increases smoothly in trading time
   - Do all time-based vol interpolation in trading time
   - Then convert back to calendar time when displaying implied volatilities, since that is market convention
3. Trading time is a monotonically increasing function of calendar time
   - Or could be flat: eg trading time stops over a weekend if we assume weekends have no variance
   - Can have jumps: eg across the Non-Farm Payrolls release we might assume trading time steps up

trading time and theta

1. "Theta" means "how much does the price of my portfolio change when I roll the clock ahead to tomorrow"
   - Should assume that vols move to "forward vols" when you roll the clock ahead
   - "Forward vol" means the implied volatility you expect to see the next day, given that some variance has rolled off as time moves forward
2. In a Black-Scholes world you can do this pretty easily
   - Calculate "forward variance" at future time t for an expiration T == implied volatility(T)ˆ2 – implied volatility(t)ˆ2

- Forward volatility = sqrt(forward variance / (T-t))

3. Unclear what to do when there is implied volatility skew/smile!

    - Mostly people keep RR/BF the same and move ATM to forward vols; or run the forward vol calculation on fixed vol-by-delta for all five benchmark deltas

1.3.5 Why vega gamma and vega dspot lead to skew & smile

smile Smile comes from volatility of volatility, paired with the symmetric vega gamma profile of vanilla options. If you buy a high-strike or a low-strike vanilla, you have a position that is long vega gamma and long vega. You can sell enough ATM vanilla to vega hedge the position; ATM options have zero vega gamma, so the position is still long vega gamma. Now, whichever way vol moves, you make money. That means traders tend to buy up high- and low-strike options and sell ATM options, which increases implied volatilities for high- and low-strike options vs the ATM implied volatility. That market pressure creates the implied volatility smile (symmetric because vega gamma is positive for both high- and low-strike options).

skew Skew comes from spot/vol correlation, paired with the asymmetric vega dspot profile of vanilla options. Assume spot/vol correlation is positive: if you buy a high-strike option, you have a position with long vega dspot and long vega. ATM options have zero vega dspot, so if you sell enough ATM vanilla to hedge vega, your position is long vega dspot and flat vega. Then if spot goes up, your vega turns positive; and you expect vol to go up because of the positive correlation, so you expect to make money. Similarly if spot goes down, vega turns negative right as you expect vol to go down because of the positive correlation, so you make money.

If you had bought a low-strike option instead, and put on a negative vega dspot position, you would lose money if spot goes up or down. This PNL behavior causes traders to buy high-strike options and sell low-strike options, creating a positive implied volatility smile.

If spot/vol correlation is negative, all the signs change, and traders want to sell high-strike options and buy low-strike options, creating a negative implied volatility smile.

So the sign of spot/vol correlation, along with the volatility of volatility (which determines the magnitude of correlated vol moves), paired with the asymmetric vega dspot profile, generates skew.

1.3.6 Definition of delta/volatility market models

implied volatility as a function of strike does not stay fixed as spot moves. Need some "market model" that defines how implied volatility moves when spot moves

Delta and market model: "Sticky delta": vol-by-delta stays fixed as spot moves 1. Most common in FX markets since vol-by-delta is quoting convention - Sticky delta: a function of the slope of implied vol vs strike 1. Most desks use sticky delta - Mostly because they mark vol by delta and want to think about PNL that way

1. Not the most efficient way to delta hedge!

    - There really is some correlation between moves in spot and moves in volatility

- In general: if an option market is liquid, people don't spend a lot of time thinking about how to hedge vol moves most efficiently with spot

volatility market model:

1. What is the best market model to use to have the most effective delta?
   - One that includes spot/vol correlation and that includes spot/RR correlation
   - We already saw that spot/RR correlation is significant (the risk reversal beta)
2. Spot/vol correlation is not that stable
   - But the risk reversal is proportional to spot/vol correlation, as we saw earlier
   - Can use that in a regression

$$\Delta\sigma(T) = A\,\mathrm{RR}(T)\,\Delta\ln(S)$$

Why do most FX shops use a "sticky delta" volatility market model when defining delta for hedging purposes Risk managers need to define "axes" for their risk calculations: that is, which market data inputs they will treat as their main risk variables.

In the FX markets, implied volatility trades in the inter-dealer market in delta terms, so ATM vol, 25d risk reversal, 25d butterfly, 10d risk reversal, and 10d butterfly are the traded market variables.

Because of that, traders tend to look at portfolio risks that move on market input at a time: spot, keeping vol-by-delta constant; ATM vol, keeping spot and RR/BF constant; and RR/BF, keeping ATM vol and spot constant.

1.3.7 Regarding cross pair volatility in terms of correlation

consider two USD pair, and cross pair

$$dS_1 = \sigma_1 S_1 dz_1$$

$$dS_2 = \sigma_2 S_2 dz_2$$

$$dS_x = (\sigma_1 dz_1 - \sigma_2 dz_2) S_x = \sigma_x S_x dz_x$$

where

$$\sigma_x = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$$

1. Can think of cross-pair vega then as vega to the two USD-pair volatilities plus risk to a correlation
2. That correlation is observable in the market (via the cross-pair options market implied volatilities)

1.3.8 Volatility relative value

1. Options risk management is complex and has a lot of moving parts
2. That complexity also leads to opportunities for market inefficiencies
3. One relative value signal: ATM curve relative value
4. If one point on the ATM volatility curve seems out of whack with the rest, buy/sell that point and sell/buy the rest

- In practice you do not run a separate strategy for this and pay spreads on all the legs
- Instead you use relative value to decide which is the best hedge to do
  - eg imagine long 6m vega
  - Could hedge by selling 6m vega: low residual risk
  - Or could hedge by selling 3m vega if 3m vol is too high relative to 6m
    * Higher residual risk but perhaps good return

1.4 Exotic

1.4.1 Pricing European payoffs through replication

European digit

1. Replication with vanillas, so no exotic model required, just volatility interpolation
2. Replicate a digital call (strike K) with a call option spread
   - Buy N units of a call struck at K-1/(2N)
   - Sell N units of a call struck at K+1/(2N)
   - Limit N->infinity, gives the digital payoff

$$V = \frac{1}{N}\left[C(K - \frac{1}{2N}) - C(K + \frac{1}{2N})\right] = -\frac{dC(K)}{dK} = -\frac{\partial C_{BS}}{\partial K} - \frac{\partial C_{BS}}{\partial \sigma}\frac{\partial \sigma}{\partial K}$$

1. In practice you cannot trade N->infinity!
   - Some maximum practical notional of vanillas, or minimum strike interval, determines the limit
2. This leaves a digital market maker with fundamentally unhedgeable risk
3. Constructing super-replicating portfolios is how we determine bid/ask spreads for digitals
   - In addition to charging bid/ask on implied volatility itself

1.4.2 Why RR beta drives barrier option pricing

OTM knockout

1. OTM (out-of-the-money) knockouts are ones where the option is OTM when the barrier is hit
   - Down and out call/up and out put
2. Very tight market for OTM knockouts: spreads=vanilla spreads
3. There is no static replication for an OTM knockout, so a model is required to estimate dynamic hedging costs
4. What market dynamics are OTM knockouts sensitive to that a model should be designed to pick up?
   - Risk reversal beta is the most important dynamic for knockouts!

hedging Consider a down-and-out call option: Strike K, barrier B<K, expiration T

1. Look at an approximate replication:
   - Long 1 unit of a call with same expiration T, same strike K
   - Short 1 unit of put with same expiration T, strike $K = B^2/K$
2. If spot never touches the barrier by expiration, the vanilla call replicates the knockout payoff
3. If spot touches the barrier, the knockout price goes to zero
   - Replication price is close to zero too as call & put offset
   - If spot touches the barrier, the knockout price really is zero, but the replication portfolio's price is only approximately zero
     - Need to do some dynamic hedging in that case!
     - Need to **unwind the replication portfolio** if the barrier is touched
   - The cost of unwinding the replication if spot touches the barrier depends on implied volatilities in that state
   - But the replication at that point looks like a risk reversal
     - Vega (to ATM vol) is very low
     - BF risk (to moves in smile) is very low
     - RR risk is the only real risk
4. The cost of unwinding the replication depends mostly on the level of the risk reversal (skew) when spot has moved to the barrier
   - This is basically the risk reversal beta: how much risk reversal moves as spot moves from the current spot down to the barrier
   - Not exactly, because it's not an instantaneous move in spot: you care what happens to risk reversal as spot moves over some extended time
   - But pretty close!

   Don't care about what happens to ATM vol or smile

Knockout pricing is most sensitive to the risk reversal beta For knockouts, as we saw in class, you can construct a hedge portfolio of vanilla options that is a "semi-static" hedge for the knockout. If I'm long a down-and-out knockout call, I can sell the regular call (same strike as the knockout) and buy a put whose strike is the knockout strike "reflected" through the barrier: B2/K. That two-vanilla portfolio hedges almost all my market risk unless the barrier is hit – then I need to unwind it. The **cost of unwinding** that two-vanilla portfolio when spot is at the barrier is mostly a function of risk reversal, because at that point the call and put are (roughly) equally out-of-the-money. That means their net sensitivity to ATM vol is small, as is their net sensitivity to butterfly. That portfolio is sensitive only to the level of risk reversal. The expected level of risk reversal is determined by how much we think risk reversal will have moved while spot moved down to the barrier, which is basically our measure of risk reversal beta: the regression coefficient of moves in risk reversal with moves in spot.

1.4.3 LV/SV mixture model definitions

More common approach due to computational efficiency: local vol/stochastic vol mixture models

1. Local vol limit: lots of risk reversal beta
2. Stochastic vol limit: very little risk reversal beta
3. Introduce a parameter that interpolates somehow between those two limits
   - That parameter controls the model's risk reversal beta

LV/SV mixture model

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma(S,t)\sqrt{v}dz_s$$

$$dv = \beta(1 - v(t))dt + \alpha\sqrt{v}dz_v$$

$E[dz_s, dz_v] = \rho dt$
$v(t)$: stochastic bit of volatility
$\sigma(S,t)$: local volatility bit of volatility, calibrated to vanillas
α is effectively the mixture parameter

1.4.4 The local vol 3-state model

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma(S,t)\sqrt{v}dz_s$$

$$v(t) = e^{Y(t)\epsilon}$$

$Y(t)$ can be -1, 0, or +1
ϵ is something like a volatility of volatility

1.4.5 Pricing vanillas using characteristic functions

let $x = \log S$, characteristic function is

$$f = E[e^{i\theta x_T}|t]$$

apply ito's Lemma to f and set dt term equal to 0, initial condition is $f(T = t) = e^{i\theta x_t}$
give an ansatz to f, to solve.
Use a complicate integration to get vanilla option's price

1.4.6 Gaussian copulas

example Consider a dual digital option that pays 1 if EURUSD is above a strike K1 and GBPUSD is above a strike K2. All discount rates are zero. The price of the EURUSD European digital option (paying 1 if EURUSD is above K1) is 65% and the price of the GBPUSD European digital option (paying 1 if GBPUSD is above K2) is 30%.

Plot the price of the dual digital option priced under a Gaussian copula model, for correlation parameter ranging from -100% to +100%. Qualitatively explain the behavior of the price sensitivity to correlation.

answer We'll use a Gaussian copula, which means that the first step is translating both the two strikes into standard normal variables. This transformation is governed by matching the cumulative distributions:

$$N(x) = F(K)$$

$N(x)$ is the standard cumulative normal distribution function: the probability that a standard normal variable is less than or equal to some level x. $F(K)$ is the (risk neutral) probability that the spot is less than or equal to some level K.

For the EURUSD piece, we know the price of a digital call, which is $1 - F(K)$(since we stated that all the interest rates are zero). So we can calculate

$$x_1 = N^{-1}(1 - 0.65) = -0.3853$$

Similarly for the GBPUSD piece we can calculate

$$x_2 = N^{-1}(1 - 0.3) = -0.5244$$

Then we can calculate the joint probability of the first standard normal being less than or equal to x_1 AND the second standard normal being less than or equal to x_2 by using the bivariate standard normal cumulative distribution function.

1.4.7 Why realized vol of implied volatility drives vol swap pricing

variance swap A variance swap contract pays out against realized volatility squared

$$\sigma^2 = \frac{N_d}{N} \sum_{i=1}^{N} \log^2\left(\frac{S_i}{S_{i-1}}\right)$$

σ^2 is the realized vol squared swapped against a fixed strike
N_d is the number of trading days/year: specified in contract
N is the number of daily spot returns in the contract period
S_i is the spot fixing for fixing date i

volatility swap

1. Volatility swaps pay off against realized volatility
2. You can think about pricing these as a square-root payoff on an asset that is the variance swap
    - The average of that asset is the **fair strike** for the variance swap
    - Need to model volatility of variance swap fair strikes

hedging

1. Buy the volatility swap, a contract with a square root payoff against the "asset", the variance swap

2. Sell an appropriate amount of the variance swap against it

    - Notional 1/2/sqrt(var swap fair strike), from derivative of square root

- When the market moves and the variance fair strike moves, the variance swap notional needs to be adjusted
  - Negative gamma to moves in the variance swap fair strike
- Because of the negative convexity, the volatility swap fair strike is less than the square root of the variance swap fair strike
  - Buy vol swap, dynamically hedge with variance swap, lose money due to negative gamma
- The spread there is a function of the realized volatility of the variance swap fair strike
  - Pretty close to the realized volatility of ATM volatility
- This is the dynamic that you need to model for vol swaps
  - Very different to what barrier derivatives care about!

volatility swap pricing is most sensitive to realized volatility of implied volatility For vol swaps, the natural hedge for a long vol swap is a short variance swap, since we can replicate the variance swap with vanilla options in a **model-free** way (usual caveats apply: doesn't work if there are jumps, and gives exposure to implied vol extrapolation behavior). The vol swap then looks like a square-root payoff against the variance swap, which is non-linear dependence. That means **you need to keep rebalancing** the notional of the short variance swap, much like **you'd need to keep rebalancing the notional of a delta hedge** against a short option position. And much the same way, you expect to **lose money** doing so, as it's a **short gamma position** (in the case of a vol swap it's short gamma to the variance swap fair strike; in the case of a vanilla option it's short gamma to the underlying spot). You then expect to lose money over time running a long vol swap position, and you get compensated for that by being able to enter the vol swap at a fair strike that's less than the square root of the variance swap fair strike. The discount in fair strike depends on the realized volatility of the variance swap fair strike, which is roughly the same as realized volatility of implied volatility.