CHARTA LEARNING · LESSON NOTES

Measuring and Monitoring Volatility

FRM Part I

1 Why volatility matters

- Volatility = the typical size of a daily move
- Value at risk needs a volatility estimate
- The bell curve is easy: two numbers say it all
- But real markets break the bell curve's rules

IN THE LESSON

- Imagine you run a small fund and must tell your boss, every morning, how much you could lose on a bad day.
- To answer, you need a measure of how much prices usually wiggle.
- That measure is volatility: the standard deviation of returns, which is simply the typical size of a daily move, up or down.
- For decades the textbook shortcut was to assume returns follow a normal distribution, the familiar symmetric bell curve, with a fixed volatility.
- It is easy, because two numbers describe everything.
- The trouble is that markets do not behave that way.
- This lesson is about how they really behave, and about how we can measure and track volatility honestly, day after day.

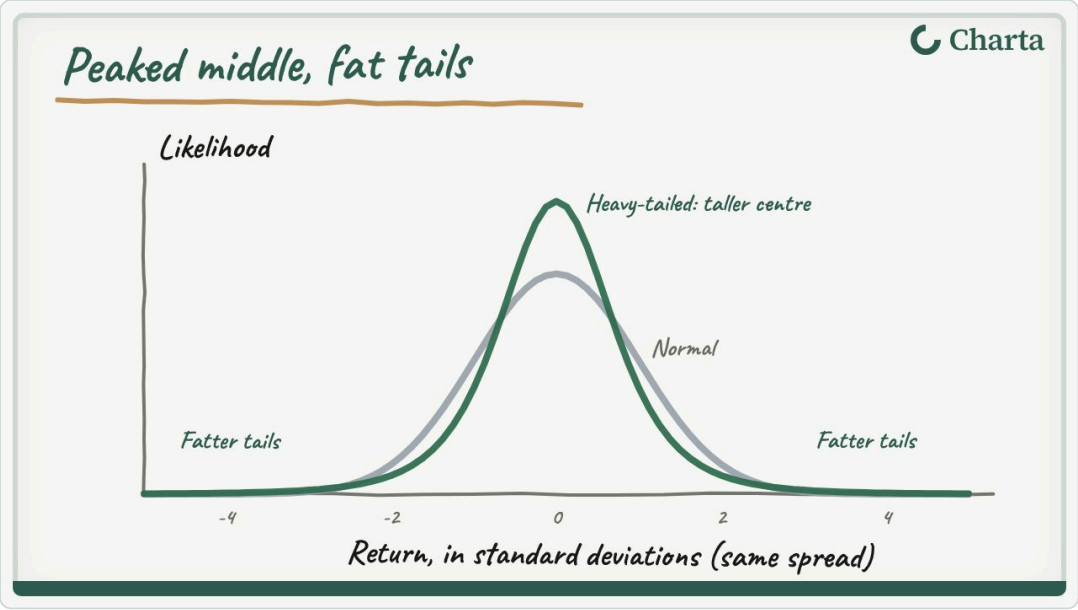
2 Four ways returns depart from normal

- Fat tails: extreme days far too frequent
- Skew: big equity losses likelier than equal gains
- Clustering: calm and stormy spells persist
- Dependence: correlations rise in sell-offs

IN THE LESSON

- Let us name the four main departures.
- First, fat tails.
- Extreme days, moves of three, four or more standard deviations, happen far more often than the bell curve allows, and the middle of the distribution is also more peaked.
- Second, skewness.
- Equity indices tend to lean to the left, so big losses are more likely than equally big gains.
- Third, volatility clustering.
- Calm spells and stormy spells each tend to last, so a big move today suggests another big move tomorrow, although the direction stays unpredictable.
- Fourth, changing dependence.
- In sell-offs, correlations between assets rise, and joint extreme losses happen more often than a normal model with the same correlation would suggest.

3 Peaked middle, fat tails



IN THE LESSON

- Here are two curves with exactly the same standard deviation.
- The grey one is the normal bell curve.
- The green one is a typical heavy-tailed shape, of the kind we see in daily returns.
- Look at the centre first: the green curve is taller and narrower, so more days than expected are very quiet.
- Look at the shoulders: the curves cross, and the green one is lower there, so moderate days are rarer.
- Now look far out at the edges.
- The green curve stays above the grey one, so extreme days are more common than the bell curve says.
- A single measure of this, called kurtosis, equals three for a normal distribution.
- Excess kurtosis is how far above three it sits.
- Do not confuse kurtosis, which measures tail weight, with skewness, which measures asymmetry.

4 How often are big moves?

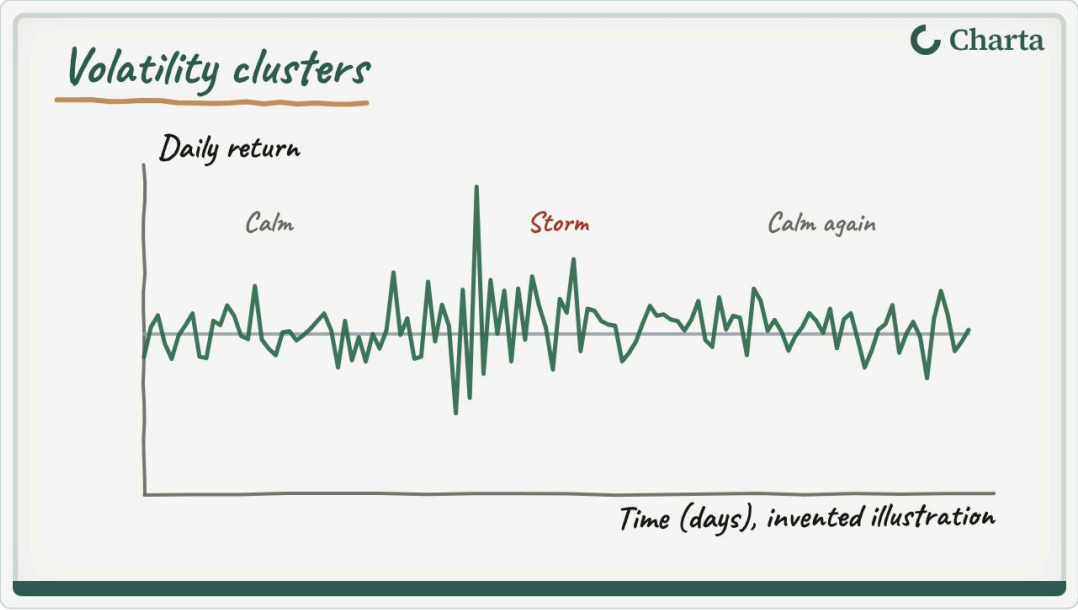
Beyond k standard deviations, either side. Heavy tails: Student t, 4 d.f., illustration

Move size	Normal | Heavy-tailed
3 sd	0.27% | 1.32%
4 sd	0.0063% | 0.48%
5 sd	0.000057% | 0.21%

IN THE LESSON

- Numbers make the point sharper.
- Take a move of at least three standard deviations in either direction.
- A normal distribution gives it a chance of about point two seven per cent.
- A Student t shape with four degrees of freedom, scaled to the same standard deviation, gives about one point three two per cent: roughly five times as often.
- At four standard deviations the normal gives about point zero zero six three per cent, while the heavy-tailed shape gives about point four eight per cent, nearly eighty times more.
- At five, the difference is thousands of times.
- A one-sided four standard deviation fall is, under the normal, a once in a hundred and twenty five years event.
- Real markets deliver such days within a working lifetime, sometimes within a single year.

5 Volatility clusters



IN THE LESSON

- The third departure deserves its own picture.
- Here is an invented series of daily returns, an illustration rather than real data.
- For a while the market is calm and the bars are small.
- Then a shock arrives and big moves of both signs follow each other for days.
- Slowly things settle again.
- Notice that you could not have guessed the direction of any single move, so the returns themselves are hard to predict.
- But the size of the moves is clearly predictable: large follows large, small follows small.
- Statisticians say the squared returns are autocorrelated, even when the returns are not.
- That persistence is exactly what we will exploit when we build volatility models later.

6 Where fat tails come from

- Volatility itself moves: a mixture of bell curves
- Jumps from news and announcements
- Regime shifts, leverage and forced selling
- More data does not make the tails normal

IN THE LESSON

- Why do tails come out fat?
- The main reason is that the conditional distribution keeps changing.
- On any given day, returns may be roughly normal, but the volatility of that day moves between calm and turbulent levels.
- Pool all the days together and you have a mixture of bell curves with different widths, and any such mixture, with a common mean, has excess kurtosis.
- A few other causes pile on top.
- News and announcements create sudden jumps.
- Markets shift between regimes.
- And in a stress, leverage and forced selling push prices further than fundamentals would.
- Finally, even very large samples keep producing extremes, rather than settling down to look normal, which is a warning against complacency.

7 Example 1: mixing two volatilities

Half the days have volatility 1%, half 3%, all with mean zero

Overall variance	$0.5\times0.0001 + 0.5\times0.0009$	**0.0005**
Overall volatility	$\sqrt{0.0005}$	**2.24%**
Average fourth power	$0.5\times3\times0.0001^2 + 0.5\times3\times0.0009^2$	**0.00000123**
Kurtosis	$0.00000123 \div 0.0005^2$	**4.92**

IN THE LESSON

- Let's prove that mixing creates fat tails, with a tiny example.
- Suppose half of all days are calm, with volatility one per cent, and half are stormy, with volatility three per cent, and every day has mean zero.
- The overall variance is the average of the two variances: half of point zero zero zero one plus half of point zero zero zero nine, which is point zero zero zero five, so overall volatility is about two point two four per cent.
- Kurtosis needs the average fourth power.
- For a normal day, the fourth moment is three times variance squared, so we average three times each variance squared.
- That gives point zero zero zero zero zero one two three, and dividing by the squared overall variance, kurtosis comes out at four point nine two, comfortably above three.

8 Example 2: the VaR cost of a normal assumption

Same mixture, overall volatility 2.24%. Daily loss quantiles

Normal, 99%	2.326 × 2.24%	**5.20%**
Mixture, 99%	true 1-in-100 loss	**6.16%**
Normal, 99.9%	3.090 × 2.24%	**6.91%**
Mixture, 99.9%	true 1-in-1000 loss	**8.63%**

IN THE LESSON

- Now see what that does to value at risk.
- Using the same mixture, with overall volatility two point two four per cent, a normal model gives a ninety nine per cent loss of two point three three times volatility, about five point two per cent.
- But the true mixture puts half its days at three per cent volatility.
- Solving for the loss that is exceeded one day in a hundred gives about six point one six per cent.
- The normal understates by nearly a full percentage point.
- Go further out.
- At ninety nine point nine per cent, the normal says about six point nine one per cent, while the mixture says eight point six three.
- The further into the tail we go, the bigger the miss.

9 What fat tails mean for a risk manager

- Normal VaR understates at 99% and beyond
- Stress test, or use heavier tails
- Use historical or simulated tails
- Never scale a quantile by normal z ratios

IN THE LESSON

- So what should a risk manager do with all this?
- First, accept that a normal value at risk calibrated to the standard deviation understates losses at ninety nine per cent and beyond, and that extreme events are far more frequent than normal tables suggest.
- Second, use tools that respect the tails: stress tests that ask what happens in a violent scenario, heavier-tailed distributions, historical or simulated tails, and models of volatility that changes over time, which the rest of this lesson builds.
- Third, avoid a classic error: taking a value at risk at a lower confidence level and scaling it up with ratios of normal z scores.
- If the tails are fat, that shortcut gives far too small an answer.

10 Conditional versus unconditional

- Conditional: tomorrow, given what we know today
- Unconditional: the average over all states
- Unconditional VaR: too low in storms
- Too high in calm spells

IN THE LESSON

- Here is the key distinction.
- The conditional distribution is the distribution of tomorrow's return given what we know today, such as the latest volatility estimate or whether markets are in turmoil.
- The unconditional distribution is the average over all such situations, which is what a long-run histogram shows.
- Because volatility changes, the conditional standard deviation can sit far above or far below the unconditional one.
- A risk figure built from the unconditional number is therefore too low in stormy periods, exactly when you need it most, and too high in calm ones.
- A weather analogy helps.
- The yearly average temperature tells you little about whether to carry an umbrella this afternoon.
- Conditional on today's sky, you can do much better.

11 Regime switching



IN THE LESSON

- One way to model changing volatility is regime switching.
- Imagine the market is always in one of two states.
- In the calm regime, returns come from a distribution with low volatility.
- In the crisis regime, they come from one with high volatility.
- Each day the market may stay where it is or jump to the other state, with given transition probabilities.
- These are usually persistent: the chance of staying put is high, which produces the clustering we saw.
- The unconditional distribution is the probability weighted mixture of the two regime distributions.
- Its variance is the weighted average of the regime variances, plus an extra term if the regime means differ, and, as we showed, the mixture has fat tails.

12 Example 3: a two-regime market

Calm σ 0.8%, crisis σ 2.5%. Calm to crisis 2%, crisis to calm 10%

Share of time in crisis	0.02 ÷ (0.02 + 0.10)	**1/6**
Unconditional variance	5/6×0.000064 + 1/6×0.000625	**0.0001575**
Unconditional volatility	√0.0001575	**1.25%**
Kurtosis	3 × E[σ to the 4th] ÷ var²	**8.29**

IN THE LESSON

- Let's put numbers on it.
- Calm volatility is point eight per cent a day and crisis volatility is two point five per cent.
- Each day, a calm market turns to crisis with probability two per cent, and a crisis ends with probability ten per cent.
- In the long run, the share of time in crisis is the calm to crisis probability divided by the sum of both, so point zero two over point one two, one sixth of the time.
- Averaging variances with those weights, five sixths of point zero zero zero zero six four plus one sixth of point zero zero zero six two five, gives about point zero zero zero one five seven five.
- The unconditional volatility is one point two five per cent, and kurtosis is about eight point three.

13 Four ways to estimate volatility

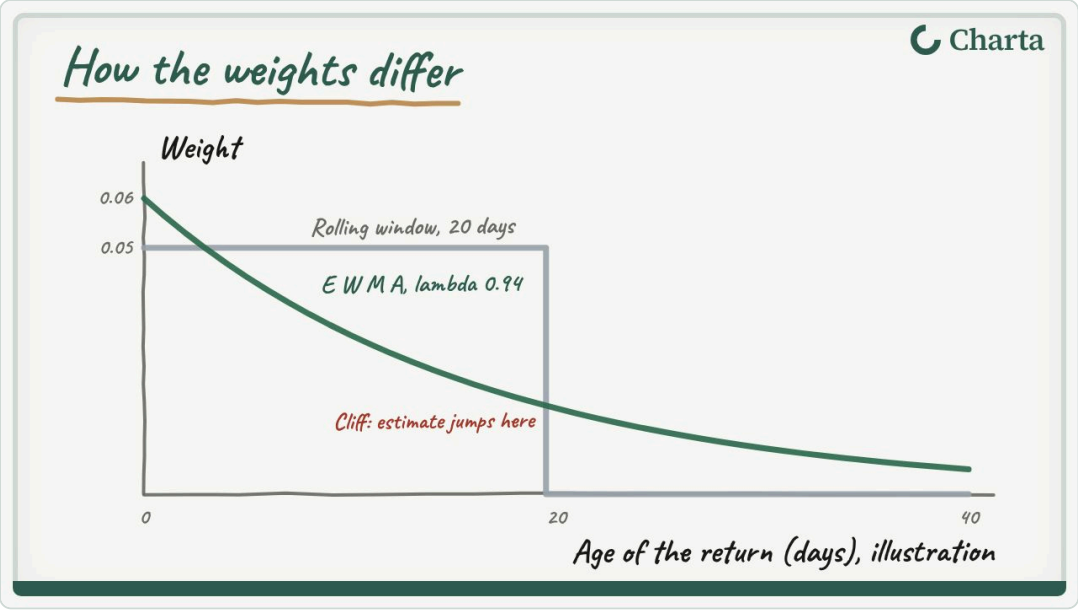
All are weighted averages of past squared returns, except implied

Method	Key feature
Rolling window	Equal weights, then a cliff
E W M A	Smooth fade, no long-run level
G A R C H (1,1)	Fade plus pull to a long-run level
Implied volatility	Forward-looking, from options

IN THE LESSON

- Now to the practical question: how do we estimate conditional volatility from data?
- There are several methods, and the big idea is that they are all alike.
- Each one says that tomorrow's variance is a weighted average of squared past returns.
- They differ in the weights.
- The equally weighted rolling window gives every day in the window the same weight and nothing outside, so estimates jump when a large return enters or leaves.
- The E W M A gives geometrically declining weights, so shocks fade smoothly, but has no long-run level.
- G A R C H also declines but adds a constant, so variance drifts back to a long-run value, at the cost of three parameters.
- Implied volatility is different: it looks forward, using option prices.

14 How the weights differ



IN THE LESSON

- This picture shows the weights each method puts on past returns, by how old they are.
- The grey line is a rolling window of twenty days: each of the latest twenty returns gets weight one twentieth, and anything older gets nothing.
- That cliff is the problem.
- The day a huge return drops out of the window, the estimate suddenly falls, even though nothing happened that day.
- The window length is a trade: short windows react quickly but are noisy, long windows are smooth but slow.
- The green line is the E W M A with a decay factor of point nine four.
- Yesterday gets the most weight, and the weights fade smoothly, with no cliff and no arbitrary cut-off.

15 The E W M A update

$$\sigma^2_n = \lambda \cdot \sigma^2_{n-1} + (1 - \lambda) \cdot u_{n-1}^2$$

Decay factor, lambda	Decay factor, 0.94 for daily
Yesterday's variance	Last estimate of variance
Latest squared return	Newest squared return
Today's variance	Updated variance

IN THE LESSON

- The E W M A model updates variance in one step.
- New variance equals lambda times yesterday's variance, plus one minus lambda times the latest squared return, where we usually assume the mean return is zero.
- Lambda is the decay factor.
- Unrolling the formula, the weight on a return that is k days old is proportional to lambda to the power k.
- A lambda near one, such as the point nine four that RiskMetrics used for daily data, gives slow decay, a smooth estimate and a long memory.
- A smaller lambda reacts faster.
- Another feature: a forecast of future variance equals the current estimate at every horizon.
- And two slips to avoid: swapping lambda with one minus lambda, and updating volatility rather than variance.

16 Example 4: one E W M A step

Old volatility 1% a day. Today's return 2%. Lambda 0.94

Old part	0.94 × 0.0001	**0.000094**
New part	0.06 × 0.0004	**0.000024**
New variance	0.000094 + 0.000024	**0.000118**
New volatility	√0.000118	**1.09%**

IN THE LESSON

- Let's run one update.
- Yesterday's volatility estimate was one per cent a day, so yesterday's variance is point zero zero zero one.
- Today the market moved by two per cent, so the squared return is point zero zero zero four.
- With lambda point nine four, the old variance gets weight point nine four, giving point zero zero zero zero nine four.
- The new squared return gets weight point zero six, giving point zero zero zero zero two four.
- Add them: new variance is point zero zero zero one one eight.
- Only now do we take a square root, getting a new volatility of about one point zero nine per cent.
- A two per cent move, twice the old volatility, nudged the estimate up by under a tenth.
- That is the smoothing at work.

17 Example 5: weighted historical simulation

K = 5 scenarios, λ = 0.9. Latest weight = (1 − λ) ÷ (1 − λ⁵)

Latest weight	0.1 ÷ (1 − 0.9⁵)	**0.2442**
One day older	0.9 × 0.2442	**0.2198**
Two days older	0.9 × 0.2198	**0.1978**
Three days older	0.9 × 0.1978	**0.1780**
Four days older	0.9 × 0.1780	**0.1602**

IN THE LESSON

- The E W M A idea has a cousin for historical simulation.
- Instead of treating every past scenario alike, give recent ones more weight.
- Suppose we have five scenarios and lambda is point nine.
- Each earlier scenario gets lambda times the weight of the next, and the weights must add to one.
- That fixes the latest weight: one minus lambda, divided by one minus lambda to the power five.
- That is point one over point four zero nine five, about point two four four.
- The next is point two two, then point one nine eight, point one seven eight, and point one six.
- Together they sum to one.
- Other schemes include equal weights and a plain window cut-off.
- All of these are just different ways of ageing old data.

18 The G A R C H one one model

$$\sigma^2_n = \omega + \alpha \cdot u_{n-1}^2 + \beta \cdot \sigma^2_{n-1}$$
$$V = \omega \div (1 - \alpha - \beta)$$

Constant, omega	Constant, sets long-run level
Reaction, alpha	Weight on newest squared return
Weight on old variance, beta	Carries old variance forward
Long-run variance, V	Variance it settles towards

IN THE LESSON

- G A R C H one one adds one idea to the E W M A: a pull towards a long-run level.
- Each day the new variance equals omega, plus alpha times the latest squared return, plus beta times the previous variance.
- Alpha is the reaction to news.
- Beta is the weight on the previous variance.
- Alpha plus beta together is the persistence, and for the long-run variance to be finite it must be below one.
- That long-run variance, V, equals omega divided by one minus alpha minus beta.
- So if you know the long-run volatility, you can back out omega.
- The three parameters are chosen by maximum likelihood, which maximises the sum of minus the natural log of the variance, minus the squared return divided by the variance.
- E W M A is the special case with omega zero and persistence exactly one.

19 Example 6: one G A R C H step

$\omega = 0.000002$, $\alpha = 0.08$, $\beta = 0.90$. Old volatility 1.2%, today's return −1.5%

Constant	ω	**0.000002**
Reaction	0.08×0.000225	**0.000018**
Old variance term	0.90×0.000144	**0.0001296**
New variance	sum of the three	**0.0001496**
New volatility	$\sqrt{0.0001496}$	**1.22%**

IN THE LESSON

- Take omega equal to two times ten to the minus six, alpha point zero eight and beta point nine.
- Persistence is point nine eight, so the long-run variance is omega over point zero two, which is point zero zero zero one, a long-run daily volatility of one per cent.
- Yesterday's volatility was one point two per cent, so the old variance is point zero zero zero one four four.
- Today the return was minus one point five per cent, and its square is point zero zero zero two two five; the sign does not matter.
- Add omega, point zero zero zero zero zero two, to alpha times the square, point zero zero zero zero one eight, and beta times the old variance, point zero zero zero one two nine six.
- The new variance is point zero zero zero one four nine six: volatility one point two two per cent.

20 Example 7: backing out omega

Long-run annual volatility 16%; $\alpha = 0.08$, $\beta = 0.90$

Annual variance	0.16^2	**0.0256**
Daily variance, V	$0.0256 \div 252$	**0.0001016**
$1 - \alpha - \beta$	$1 - 0.98$	**0.02**
Omega	0.02×0.0001016	**0.00000203**

IN THE LESSON

- In practice you often start from a view of long-run volatility.
- Say the long-run annual volatility is sixteen per cent.
- Work in variance and in daily terms.
- Square sixteen per cent to get an annual variance of point zero two five six, then divide by two hundred and fifty two trading days: about point zero zero zero one zero one six daily variance.
- That is V.
- Now omega equals one minus alpha minus beta, times V.
- With persistence point nine eight, one minus persistence is point zero two, so omega is about two point zero three times ten to the minus six, close to the value we used.
- Two traps: forgetting to divide by two hundred and fifty two, and mixing volatility with variance.

21 Looking ahead: mean reversion

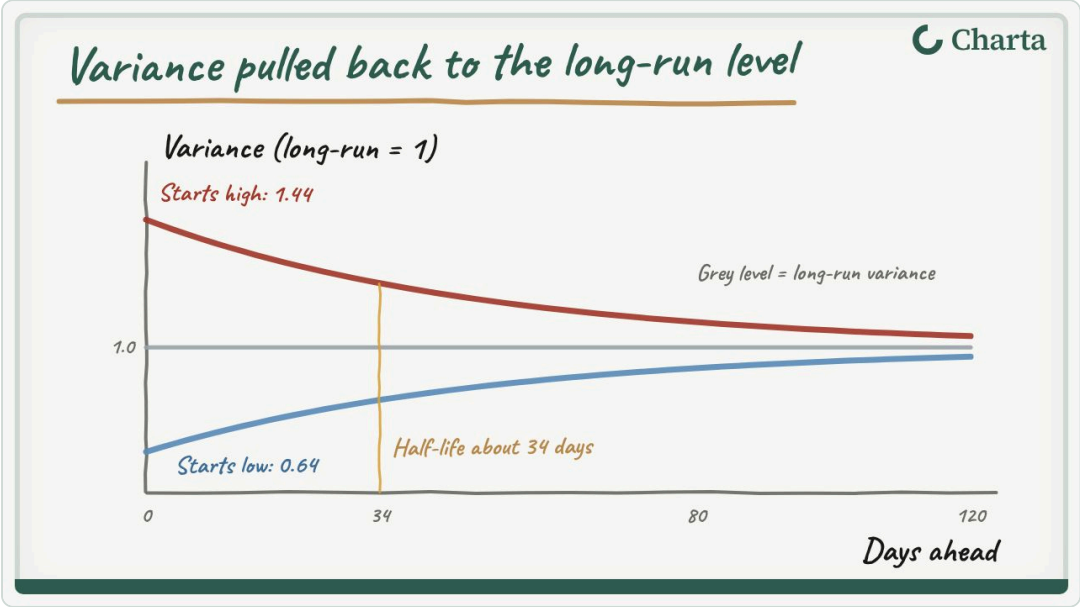
$$E[\sigma^2_{n+k}] = V + (\alpha + \beta)^k \cdot (\sigma^2_n - V)$$
$$\text{half-life} = \ln(0.5) \div \ln(\alpha + \beta)$$

Forecast variance	Expected variance k days out
Long-run variance, V	Long-run variance
Persistence, alpha + beta	Persistence, below one
Current variance	Variance now
Days ahead, k	Number of days ahead

IN THE LESSON

- Now to longer horizons.
- Under G A R C H the expected variance k days ahead is the long-run variance V, plus persistence to the power k, times the gap between current variance and V.
- So the forecast decays geometrically towards V.
- How fast?
- The half-life is the time to close half the gap: the natural log of one half, divided by the natural log of persistence, in days.
- Persistence close to one means shocks fade slowly.
- The T-day variance is the sum of the daily expected variances, and the T-day volatility is the square root of that sum.
- The usual shortcut multiplies daily volatility by the square root of T, and the T day V a R scales the same way under normality with a zero mean.
- That needs independent returns with constant variance, and positive autocorrelation makes it understate.
- When volatility is mean reverting, summing the expected variances is the correct approach.

22 Variance pulled back to the long-run level



IN THE LESSON

- Here is mean reversion drawn.
- The flat grey level is the long-run variance, shown as one.
- One curve starts above it, at one point four four, which is our one point two per cent volatility.
- The other starts below, at point six four, a volatility of point eight per cent.
- Each day the gap shrinks by two per cent, so both curves glide towards the long-run level.
- The marked point shows the half-life: after about thirty four days, half the starting gap has gone.
- Now the practical lesson.
- When current volatility is above the long-run level, scaling today's volatility by the square root of time overstates long-horizon risk.
- When it is below, scaling understates it.

23 Example 8: a ten-day horizon

A fresh case: variance tomorrow 0.000144 (1.2%), V = 0.0001, α + β = 0.98, T = 10

Square-root scaling	1.2% × √10	**3.79%**
Sum of 0.98 powers	$(1 - 0.98^{10}) \div 0.02$	**9.146**
Ten-day variance	10×0.0001 + 9.146×0.000044	**0.001402**
Ten-day volatility	√0.001402	**3.74%**

IN THE LESSON

- Let's measure ten-day risk for a similar case, where tomorrow's variance is point zero zero zero one four four, V is point zero zero zero one, and persistence is point nine eight.
- Naive square root scaling gives one point two per cent times the square root of ten: three point seven nine per cent.
- Now do it properly.
- The ten-day variance is ten times V, point zero zero one, plus the gap, point zero zero zero zero four four, times the sum of persistence to the powers zero through nine, about nine point one four six.
- That adds point zero zero zero four zero two, for a total of point zero zero one four zero two four.
- Its square root is three point seven four per cent.
- Scaling overstated risk slightly, as the theory predicted.

24 Implied volatility

- The volatility that reproduces the option's price
- Forward-looking: reflects what traders expect
- Reacts immediately to news
- Quoted annually: divide by √252 for daily

IN THE LESSON

- There is a completely different source of volatility numbers: the options market.
- An option's price depends on the volatility of the underlying asset, and the Black-Scholes-Merton model turns that dependence into a formula.
- Implied volatility is the volatility that, put into the model, reproduces the option's observed market price.
- Its great strength is that it looks forward.
- It reflects what traders expect over the option's remaining life, and it adjusts at once to news, which no historical estimator can do.
- By convention it is quoted as an annual figure.
- To compare with a daily number, divide by the square root of two hundred and fifty two, the usual count of trading days in a year.

25 Example 9: reading an implied volatility

Index option implied volatility 20% a year

Daily volatility	20% ÷ √252	**1.26%**
Ten-day volatility	1.26% × √10	**3.98%**
Annual variance	0.20^2	**0.04**
Daily variance	0.04 ÷ 252	**0.000159**

IN THE LESSON

- Suppose an index option has an implied volatility of twenty per cent a year.
- What does that say about a typical day?
- Divide by the square root of two hundred and fifty two, which is about fifteen point eight seven.
- Twenty divided by fifteen point eight seven is a daily volatility of about one point two six per cent.
- Over ten trading days, scale by the square root of ten, to get about three point nine eight per cent.
- In variance terms, the annual variance is point zero four, and the daily variance is point zero zero zero one five nine.
- Compare this with a G A R C H estimate: if the model says one point two per cent today and the market implies one point two six, the market implies a little more than the model, which may reflect expected turbulence or simply the premium that implied volatility usually carries.

26 Weaknesses of implied volatility

A useful input, never the whole answer

Weakness	Why it matters
Model dependent	Smile and skew break the model
Limited coverage	Needs liquid options
Biased upwards	Includes a risk premium
Market frictions	Spreads and demand move it

IN THE LESSON

- Implied volatility is useful, but it has real weaknesses.
- It depends on the model: if Black-Scholes-Merton were right and volatility constant, options of all strikes and maturities on one asset would show the same implied volatility, yet in practice they form a smile or skew, which shows the assumption fails.
- It exists only for assets with liquid options, and only at certain maturities.
- On average it usually exceeds the volatility that is later realised, partly because it embeds a premium for bearing volatility risk, so it is a biased forecast.
- And illiquidity, wide bid-ask spreads and heavy demand for protection can all move it.
- Treat it as informative but imperfect.

27 Updating a correlation

$$\text{cov}_n = \lambda \cdot \text{cov}_{n-1} + (1 - \lambda) \cdot x \cdot y$$
$$\rho_n = \text{cov}_n \div (\sigma X_n \cdot \sigma Y_n)$$

Decay factor, lambda	Same decay factor for all three
Covariance	Old covariance: ρ × σX × σY
Return on asset X	Newest return, asset X
Return on asset Y	Newest return, asset Y
Correlation	New covariance ÷ new volatilities

IN THE LESSON

- Our last topic is correlation.
- It is updated the very same way as a variance.
- With E W M A, the new covariance equals lambda times the old covariance, plus one minus lambda times x times y, where x and y are the latest returns on the two assets.
- The two variances are updated with the same lambda.
- The new correlation is the new covariance divided by the product of the two new volatilities.
- To begin, convert the old correlation and volatilities into an old covariance: correlation times volatility of X times volatility of Y.
- A G A R C H version also exists: covariance equals omega plus alpha times x times y plus beta times the old covariance, with omega equal to one minus alpha minus beta, times the long-run covariance.

28 Example 10: updating a correlation

σX 1.0%, σY 1.2%, ρ 0.6. Today X +1.5%, Y −1.0%. λ = 0.94

Old covariance	0.6 × 0.010 × 0.012	**0.000072**
New covariance	0.94×0.000072 − 0.06×0.00015	**0.00005868**
New volatilities	√0.0001075, √0.00014136	**1.04%, 1.19%**
New correlation	0.00005868 ÷ (0.01037×0.01189)	**0.476**

IN THE LESSON

- Time for a full example.
- Volatility of X is one per cent, of Y is one point two per cent, and the correlation is point six.
- Old covariance is point six times point zero one times point zero one two, which is point zero zero zero zero seven two.
- Today X rose one point five per cent while Y fell one per cent.
- The product is minus point zero zero zero one five, with a negative sign.
- The new covariance is point nine four times the old, point zero zero zero zero six seven six eight, plus point zero six times minus point zero zero zero one five, minus point zero zero zero zero zero nine: point zero zero zero zero five eight six eight.
- New variances give volatilities of one point zero four and one point one nine per cent.
- The new correlation is about point four eight.

29 Correlation: the common errors

- Divide by the NEW volatilities
- One lambda for all three quantities
- Keep the sign of x times y
- Opposite signs pull correlation down

IN THE LESSON

- Three errors come up again and again.
- First, dividing by the old volatilities instead of the updated ones.
- The updated correlation needs all three pieces, covariance and both variances, brought up to date.
- Second, using different lambda values for the three quantities; in an E W M A scheme, one lambda should serve all three, otherwise the resulting correlation can fall outside minus one and plus one.
- Third, ignoring the sign of the return product.
- Returns of opposite sign make the product negative, which pulls the covariance and the correlation down, as in our example, where the estimate fell from point six to point four eight.
- Same signs push it up.
- And as with variance, G A R C H parameters should be consistent across the quantities.

30 Recap: the main ideas

- Fat tails: volatility moves, normal VaR too low
- Weighted squared returns: window, E W M A, G A R C H
- G A R C H reverts to V; half-life ln 0.5 ÷ ln(α+β)
- Implied volatility: forward-looking, but biased
- Covariance updates like variance; use new volatilities

IN THE LESSON

- Let us pull it together.
- Real returns have fat tails, skew, clustering and shifting dependence, mainly because volatility moves, so a normal model with matching variance understates extreme risk.
- The conditional distribution, given today's state, differs from the unconditional one, and regime switching models that as a mixture.
- All the estimators are weighted averages of past squared returns: the window with cliff-edge weights, E W M A with geometric weights, and G A R C H with a constant and a pull towards a long-run level.
- Over long horizons, G A R C H forecasts decay to the long-run variance, with a half-life, and square-root scaling can mislead.
- Implied volatility is forward-looking but biased and model dependent.
- Correlations update like variances.

31 Try three quick questions

1. λ 0.94, old variance 0.0001, return 2%. New variance?
   Answer: 0.94×0.0001 + 0.06×0.0004 = 0.000118.
2. G A R C H with α 0.10, β 0.85. Persistence and half-life of the variance gap?
   Answer: 0.95. Half-life = ln 0.5 ÷ ln 0.95 = 13.5 days.
3. Implied volatility 30% a year. Daily?
   Answer: 30% ÷ √252 = 1.89% a day.

IN THE LESSON

- Time for three quick questions.
- Pause the video, think each one through, then listen to the answer.
- Question one.
- With lambda point nine four, yesterday's variance point zero zero zero one and a return today of two per cent, what is the new E W M A variance?
- Point nine four times point zero zero zero one plus point zero six times point zero zero zero four gives point zero zero zero one one eight.
- Question two.
- A G A R C H model has alpha point one and beta point eight five, so what is its persistence, and what is the half-life of the variance gap?
- Persistence is point nine five, and the half-life is natural log of point five over natural log of point nine five, which is thirteen point five days.
- Question three.
- An implied volatility of thirty per cent a year is how much per day?
- Thirty per cent divided by the square root of two hundred and fifty two is one point eight nine per cent a day.