CHARTA LEARNING · LESSON NOTES

Measuring Returns, Volatility and Correlation

FRM Part I

1 A story: up fifty, down fifty

- 100 → +50% → 150 → −50% → 75
- Adding simple returns says 0%; the loss was 25%
- Simple returns do not add across periods
- Fix: use log returns, which do add

IN THE LESSON

- Here is a puzzle to start.
- A fund starts at one hundred.
- In the first year it gains fifty per cent.
- In the second year it loses fifty per cent.
- The simple returns are plus fifty and minus fifty, and if you add them you get zero, so you might say the fund broke even.
- But look at the money.
- One hundred grows to one hundred and fifty, then falls by half to seventy five.
- The investor has lost a quarter of the stake.
- Simple returns cannot just be added over time.
- This lesson starts by fixing that problem, with a second way to measure returns that does add up, and then builds the other measuring tools we need.

2 Simple and log returns

$$R = (P_1 + D - P_0) \div P_0$$
$$r = \ln((P_1 + D) \div P_0) = \ln(1 + R)$$

Simple return, R	Percent change, with income
Gross return	Ending value over starting price
Natural log	Natural log of the gross return
Log return, r	Equals ln(1 + R); back: e^{r} − 1

IN THE LESSON

- There are two standard ways to measure a return.
- The simple return is the percentage change in value over one period, including any income.
- It equals the end price plus income, minus the start price, all divided by the start price.
- The continuously compounded return, or log return, is the natural logarithm of the gross return, where the gross return is end value plus income divided by the start price.
- The link between them is neat.
- The log return equals the natural log of one plus the simple return.
- To go back, the simple return equals the exponential of the log return, minus one.
- For small moves the two are close.
- A log return is below the simple return when both are positive.

3 Example 1: converting between the two

Illustration: bought at 100, dividend 2, ends at 108

Simple return R	(108 + 2 − 100) ÷ 100	**10%**
Log return r	ln(1.10)	**9.53%**
Log is lower when positive	9.53% < 10%	**yes**
Convert back: e^{r} − 1	e^{0.0953} − 1	**10%**

IN THE LESSON

- Let's convert with made up numbers.
- A share is bought at one hundred, pays a dividend of two, and is later worth one hundred and eight.
- The gain including income is eight plus two, which is ten, so the simple return is ten per cent.
- The gross return is one point one.
- Take the natural log of one point one and you get point zero nine five three, so the log return is nine point five three per cent.
- It is below the simple return, as expected.
- Now go back.
- The exponential of point zero nine five three is one point one, and subtracting one gives ten per cent again.
- So you can always move from one measure to the other without losing anything.

4 Example 2: adding across two years

Illustration: year one +20%, year two −10%

Compounded simple return	1.20 × 0.90 − 1	**8%**
Log returns, year 1 and 2	ln 1.20 and ln 0.90	**18.23%, −10.54%**
Sum of the logs	18.23% − 10.54%	**7.70%**
Back to simple	e^{0.0770} − 1	**8%**

IN THE LESSON

- Now the practical difference, aggregation.
- Take a fund that gains twenty per cent in the first year and loses ten per cent in the second.
- Compound the simple returns: one point two times point nine is one point oh eight, so the two year simple return is eight per cent.
- Now use logs.
- The first year log return is eighteen point two three per cent.
- The second year is minus ten point five four.
- Simply add them and you get seven point seven per cent.
- Convert that back with the exponential, minus one, and you get eight per cent, exactly the same answer.
- Adding the simple returns would have given ten per cent, which is wrong.
- Logs add, simple returns must be compounded.

5 How far apart are the two measures?

Log return r = ln(1 + R) for different simple returns R

Simple return R	Log return r
+1%	0.995%
+10%	9.53%
+50%	40.55%
−50%	−69.31%
+50% then −50%	−28.77% (= −25%)

IN THE LESSON

- Let's see how the gap behaves.
- For a one per cent simple return, the log return is point nine nine five per cent, almost identical.
- At ten per cent the log return is nine point five three.
- At fifty per cent the log return has dropped to forty point five five.
- So the gap widens as the move grows.
- Losses behave the other way round.
- A fall of fifty per cent is a log return of minus sixty nine point three one, which is bigger in size than the simple figure.
- The fund from our story now makes sense: forty point five five plus minus sixty nine point three one is minus twenty eight point seven seven, which converts back to a loss of twenty five per cent.
- Remember the habit: use logs to model and aggregate, and convert back to simple returns to report the gain or loss on a position.

6 Volatility, variance rate and implied volatility

- Volatility = standard deviation of returns, per unit of time
- Variance rate = variance per unit time = vol²
- Independent variances add: variance grows with time
- Volatility grows with the square root of time
- Implied vol = the value that matches an option price

IN THE LESSON

- Now volatility.
- Volatility is the standard deviation of an asset's returns, quoted per unit of time, usually per year.
- It tells you how widely the returns scatter around their average.
- The variance rate is the variance per unit of time, so it is simply the volatility multiplied by itself.
- These two have a very useful property.
- The variances of independent returns add up, so the variance rate grows in a straight line with the horizon, and the volatility grows with the square root of time.
- A third idea, implied volatility, is different in kind.
- It is not measured from past returns at all.
- It is the volatility you must put into an option pricing model so that the model matches the option's price in the market.

7 Scaling volatility with time

$\sigma_T = \sigma_{daily} \times \sqrt{T}$
$\sigma_{annual} = \sigma_{daily} \times \sqrt{252}$

Daily volatility	Volatility of one day
Number of days, T	Days in the horizon (252 a year)
Square root of time	Variance adds, so volatility × √T
Horizon volatility	Volatility over the whole horizon

IN THE LESSON

- Here is the scaling rule.
- If the daily volatility is known, the volatility over a horizon of T days equals the daily volatility multiplied by the square root of T.
- The reason is that variance, not volatility, adds across independent days.
- Over T days the variance is T times the daily variance, and taking the square root gives the square root of T.
- To annualize, we use two hundred and fifty two trading days in a year, so the annual volatility equals the daily volatility times the square root of two hundred and fifty two.
- That square root is about fifteen point nine.
- The common mistake is to multiply by the number of days instead of its square root.
- That overstates risk badly.

8 Example 3: from daily to annual

Illustration: daily volatility 1.2%, 252 trading days a year

Daily variance rate	1.2² (percent squared)	**1.44**
Annual volatility	1.2 × √252 = √362.88	**19.05%**
Ten-day volatility	1.2 × √10	**3.79%**
Wrong: multiply by days	1.2 × 252	**302.4%**

IN THE LESSON

- Let's scale a daily figure.
- Suppose the daily volatility of a stock is one point two per cent.
- That is a variance rate of one point four four per day, in percentage points squared.
- Over a year of two hundred and fifty two days, the variance is two hundred and fifty two times one point four four, which is three hundred and sixty two point eight eight.
- Take the square root and the annual volatility is nineteen point oh five per cent.
- Equivalently, one point two times the square root of two hundred and fifty two.
- For a ten day horizon the volatility is one point two times the square root of ten, about three point seven nine per cent.
- The wrong way, multiplying by two hundred and fifty two, would give three hundred and two per cent.

9 Example 4: from annual to short horizons

Illustration: annual volatility 24%, 21 trading days in a month

Daily volatility	24% ÷ √252	**1.51%**
Share of a year	21 ÷ 252	**1/12**
Monthly volatility	24% × √(1/12)	**6.93%**
Not 24% ÷ 12	2% would be wrong	**6.93% ≠ 2%**

IN THE LESSON

- The scaling also works in reverse.
- Imagine an index with an annual volatility of twenty four per cent.
- To find the daily volatility, divide by the square root of two hundred and fifty two, which is about fifteen point eight seven.
- The answer is one point five one per cent a day.
- For a one month horizon, take twenty one trading days.
- The volatility is twenty four times the square root of twenty one over two hundred and fifty two.
- That fraction is one twelfth, and its square root is about point two eight eight nine, so the monthly volatility is six point nine three per cent.
- Notice that a month is a twelfth of a year, yet its volatility is not a twelfth of twenty four.
- It is the annual figure divided by the square root of twelve.

10 Historical versus implied volatility

Two different ways to arrive at a volatility number

Feature	Historical | Implied
Source	Past returns | Option prices
Direction	Looks back | Looks forward
How found	Std deviation | Invert model
Reflects	What happened | Supply, demand
Model needed	No | Yes (e.g. BSM)

IN THE LESSON

- Now set implied volatility against historical volatility.
- Historical volatility is calculated from realized past returns, so it looks backward, and it depends on the window you pick.
- Implied volatility looks forward.
- Take an option on the asset, observe its price in the market, and ask which volatility, when fed into the Black Scholes Merton model, reproduces that price.
- Because the price comes from supply and demand, the implied volatility reflects what the market expects and also what it is willing to pay for protection.
- The two usually differ.
- Neither is simply right.
- Historical is a fact about the past, implied is a market view of the future, so risk managers watch both.

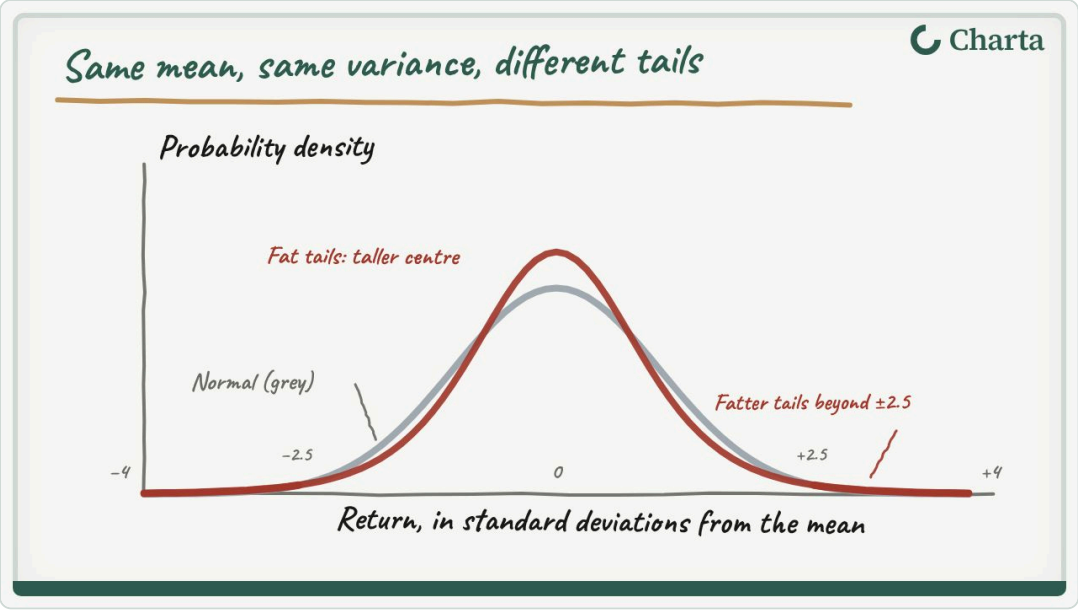
11 Mean and variance: not always enough

- Normal: mean and variance describe everything
- Skewness (3rd moment): 0 if normal; negative = left tail
- Kurtosis (4th moment): 3 for normal
- Excess kurtosis = kurtosis − 3; positive = fat tails
- Same mean and variance can hide very different tail risk

IN THE LESSON

- Mean and variance are the first two moments of a distribution.
- For a normal distribution they say everything: knowing the mean and variance, you know the probability of any loss.
- Real returns are not normal, and two more moments matter.
- Skewness is the standardized third moment and measures asymmetry.
- A negative skew means a longer left tail, so large losses are more likely than equally large gains.
- Kurtosis is the standardized fourth moment.
- It equals three for a normal distribution, and the amount above three is the excess kurtosis.
- Positive excess kurtosis means fat tails: extreme moves happen far more often than the normal model predicts.
- Two series can share the same mean and variance yet carry very different tail risk.

12 Same mean, same variance, different tails



IN THE LESSON

- Here are two distributions drawn with exactly the same mean and the same variance.
- The grey bell is the normal.
- The red curve is a fat tailed one, with positive excess kurtosis.
- Look at the middle first.
- The fat tailed curve is taller and more peaked, so more of its mass sits close to the centre.
- Now look at the far ends of the axis, beyond about two and a half standard deviations.
- There the red curve lies above the grey one, so extreme moves, both gains and losses, are more common than the normal model allows.
- This is why a value at risk number based on the normal can understate losses at high confidence levels, and why it can overstate risk at lower confidence levels, where the taller centre matters.

13 The Jarque-Bera test

$JB = (n \div 6) \times [\, S^2 + (K - 3)^2 \div 4 \,]$
Reject normality if JB > 5.99 (5%) or 9.21 (1%)

Observations, n	Number of return observations
Skewness, S	Sample skewness (0 if normal)
Kurtosis, K	Plain kurtosis (3 if normal)
Jarque-Bera statistic	Chi-square, 2 degrees of freedom

IN THE LESSON

- To judge whether returns are normal we can use the Jarque Bera statistic.
- It tests whether a sample's skewness and kurtosis are consistent with a normal distribution.
- With n observations, sample skewness S and sample kurtosis K, the statistic is n over six, times skewness squared plus a quarter of the squared excess kurtosis.
- Note that K here is the plain kurtosis, and we subtract three inside the formula.
- The null hypothesis is that returns are normal, so skewness is zero and excess kurtosis is zero.
- Under that null, the statistic follows a chi square distribution with two degrees of freedom.
- The critical values are about five point nine nine at the five per cent level and nine point two one at the one per cent level.
- Above the critical value, reject normality.

14 Example 5: testing a return series

Illustration: n = 250, S = −0.4, K = 4.5

Skewness part	$(-0.4)^2$	**0.16**
Kurtosis part	$(4.5 - 3)^2 \div 4$	**0.5625**
JB = (250 ÷ 6) × 0.7225	41.67 × 0.7225	**30.10**
Compare with 9.21	30.10 > 9.21	**Reject**

IN THE LESSON

- Let's run the test on made up data.
- A fund has two hundred and fifty daily returns.
- The sample skewness is minus point four, and the sample kurtosis is four point five.
- First square the skewness: point four squared is point one six.
- Next, kurtosis minus three is one point five, squared is two point two five, divided by four gives point five six two five.
- Add the two pieces and the bracket is point seven two two five.
- Now multiply by n over six, which is two hundred and fifty over six, about forty one point six seven.
- The statistic is thirty point one.
- That is far above nine point two one, the one per cent critical value, so we reject the idea that these returns are normal.

15 Example 6: the role of sample size

Illustration: same shape (bracket = 0.7225), different n

n = 20	20 ÷ 6 × 0.7225	**2.41**
n = 250	250 ÷ 6 × 0.7225	**30.10**
n = 1,000	1000 ÷ 6 × 0.7225	**120.42**
Critical values	5% and 1%	**5.99, 9.21**

IN THE LESSON

- The statistic has a built in feature.
- Because it is proportional to n, the sample size matters a great deal.
- Keep exactly the same shape, skewness minus point four and kurtosis four point five, so the bracket stays at point seven two two five.
- With only twenty observations, n over six is three point three three, and the statistic is two point four one.
- That is below five point nine nine, so we fail to reject normality, even though the shape is clearly not normal.
- With two hundred and fifty it is thirty point one.
- With a thousand observations it jumps to one hundred and twenty point four two.
- So a huge sample can reject normality for a small departure, and a small sample can miss a clearly non normal series.

16 The power law for tails

- P(loss > x) = k × x^{−α}, where α is the tail index
- Logs: ln P = ln k − α × ln x, a straight line
- Smaller α means a fatter tail
- Power law falls polynomially; normal like e^{−x²/2}
- Risk managers use it to estimate extreme tail probabilities

IN THE LESSON

- When returns have fat tails, we need a model for the tail itself.
- The power law is one.
- It says the probability that a loss exceeds some level x equals a constant k times x raised to the power of minus alpha.
- Alpha is the tail index.
- Take logs of both sides and you get a straight line: the log of the probability equals the log of k minus alpha times the log of x.
- A smaller alpha means a fatter tail.
- The key contrast is with the normal.
- Power law probabilities fall polynomially, far more slowly than the normal tail, which falls roughly like an exponential of minus x squared over two.
- So the power law gives much more weight to extreme losses.

17 The scaling factor and finite moments

$P(\text{loss} > x) = k \times x^{-\alpha}$
Threshold × a ⇒ probability × $a^{-\alpha}$
Moment of order m is finite only if $m < \alpha$

Constant, k	Constant that sets the scale
Loss threshold, x	Loss level being exceeded
Tail index, alpha	Tail index: smaller = fatter
Tail probability	Chance the loss exceeds x

IN THE LESSON

- The power law has a handy consequence.
- If you multiply the loss threshold by a factor a, the tail probability is multiplied by a raised to the power of minus alpha.
- So doubling the threshold cuts the probability by a factor of two to the power alpha.
- That lets a risk manager extrapolate from a level where there is data to a level beyond it, where there is not.
- There is also a warning about moments.
- Only moments of order below alpha are finite.
- If alpha is below two, the variance is infinite.
- If alpha is below four, the kurtosis is infinite.
- That is why heavy tailed series can make sample variances unstable.

18 Example 7: extrapolating a tail

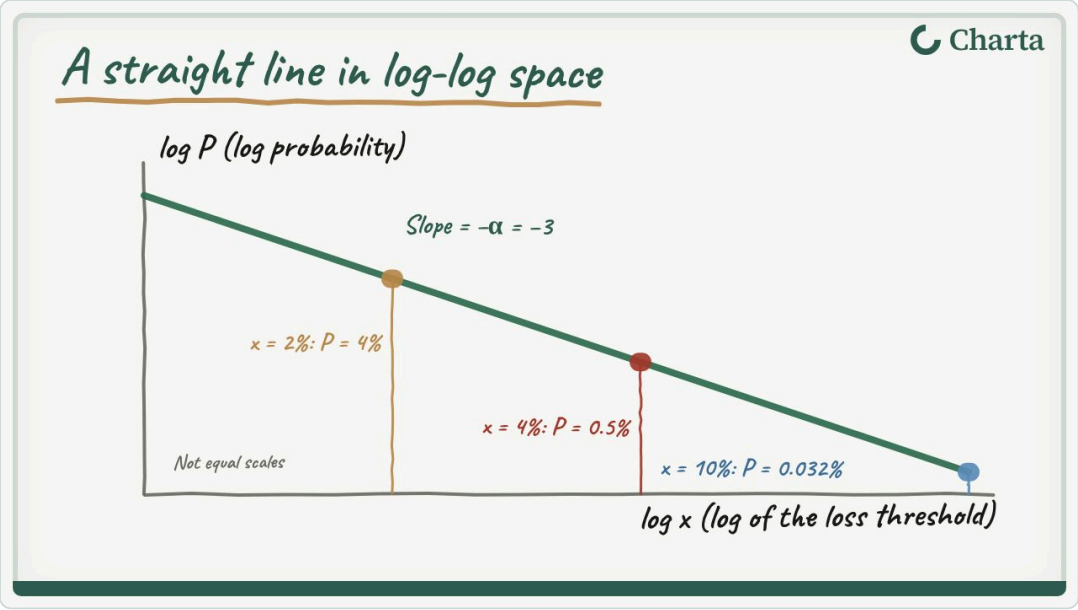
Illustration: alpha = 3, P(loss > 2%) = 4%

Double the threshold	$4\% \div 2^3$	**0.5%**
Loss above 10%	$4\% \times 5^{-3} = 4\% \div 125$	**0.032%**
Recover alpha from two points	$\ln(4\% \div 0.5\%) \div \ln 2$	**3**
Moments: alpha = 3 < 4	kurtosis	**infinite**

IN THE LESSON

- Let's extrapolate with made up numbers.
- Suppose the tail index is three, and the probability of a loss above two per cent in a day is four per cent.
- Double the threshold to four per cent.
- The probability falls by a factor of two cubed, which is eight, so it becomes half of one per cent.
- Now go further, to a threshold of ten per cent, five times the original.
- The probability is four per cent times five to the power of minus three, which is one over one hundred and twenty five, giving point zero three two per cent.
- We can also recover the index from two points.
- The natural log of eight divided by the natural log of two is three, the slope of the log line.
- With alpha equal to three, below four, the kurtosis is infinite.

19 A straight line in log-log space



IN THE LESSON

- Here is the power law drawn on logarithmic axes.
- Along the bottom is the log of the loss threshold, and up the side is the log of the tail probability.
- Because the logs turn the formula into a straight line, we get a straight downward sloping line, and its slope is minus alpha, here minus three.
- The three marked points match our example: a threshold of two per cent, four per cent, and ten per cent, each with the probability we found.
- Every doubling of the threshold moves you the same distance to the right, and drops the probability by the same factor of eight.
- If the data really follow a power law, this plot shows a line, and the slope gives the tail index.
- A normal tail would curve steeply downward instead.

20 Covariance, correlation and dependence

- Covariance = average of (X − mean) × (Y − mean)
- Units are awkward: units of X times units of Y
- Correlation = Cov ÷ (σX × σY), from −1 to +1
- Correlation sees only linear association
- Independent ⇒ zero correlation, not the reverse

IN THE LESSON

- Now we move from one asset to two.
- Covariance measures how two variables move together.
- It is the average product of their deviations from their own means.
- If both tend to be above average together, the covariance is positive.
- Its size is hard to read, because its units are the units of the first variable times the units of the second.
- Correlation fixes that by dividing the covariance by the product of the two standard deviations.
- The result has no units and lies between minus one and plus one.
- But correlation captures only linear association.
- Variables can be perfectly dependent and still have zero correlation.
- Independence implies zero correlation, but zero correlation does not imply independence, except in special cases such as jointly normal variables.

21 Covariance and correlation

Cov(X,Y) = E[(X − μ_X)(Y − μ_Y)]
Corr = Cov ÷ (σ_X × σ_Y) so Cov = Corr × σ_X × σ_Y

Covariance	Average co-movement, awkward units
Volatility of X	Standard deviation of X
Volatility of Y	Standard deviation of Y
Correlation	Unit-free, from −1 to +1

IN THE LESSON

- Here are the two formulas side by side.
- Covariance is the expected value of the deviation of X from its mean, multiplied by the deviation of Y from its mean.
- Correlation is that covariance divided by the volatility of X times the volatility of Y.
- Rearranged, which is how questions often use it, the covariance equals the correlation times the two volatilities multiplied together.
- If you are given the standard deviations and the correlation and asked for the covariance, just multiply all three.
- Correlation of plus one means a perfect upward line, minus one a perfect downward line, and zero means no linear relationship, which is not the same as no relationship at all.

22 Example 8: covariance from correlation

Illustration: σX = 10%, σY = 20%, correlation 0.6

Product of the volatilities	0.10 × 0.20	**0.02**
Covariance	0.6 × 0.02	**0.012**
Check: covariance ÷ product	0.012 ÷ 0.02	**0.6**

IN THE LESSON

- Let's use the rearranged formula.
- Two assets have volatilities of ten per cent and twenty per cent, and a correlation of point six.
- What is the covariance?
- Multiply the two volatilities first: point one times point two is point zero two.
- Then multiply by the correlation: point zero two times point six is point zero one two.
- That is the covariance.
- To check, run it backwards.
- Divide the covariance by the product of the volatilities: point zero one two over point zero two is point six, which brings back the correlation.
- Notice what the covariance tells us on its own.
- Its size, point zero one two, means little without those volatilities, while the correlation of point six is directly readable.

23 Example 9: zero correlation, full dependence

Illustration: X = −2, −1, 0, 1, 2 and Y = X²

Means of X and of Y	0 and (4+1+0+1+4)÷5	**0 and 2**
Products (x)(y − 2)	−4, 1, 0, −1, 4	**sum = 0**
Covariance, correlation	0 ÷ 5, then 0 ÷ (σXσY)	**0 and 0**
Is Y determined by X?	Y = X², exactly	**Yes, fully**

IN THE LESSON

- Here is the classic trap.
- Let X take the five values minus two, minus one, zero, one and two, and let Y be X squared, so Y is four, one, zero, one, four.
- Y is completely determined by X, yet look at the correlation.
- The mean of X is zero.
- The mean of Y is two.
- The products of the deviations are minus two times two, which is minus four; minus one times minus one, which is one; zero; minus one; and plus four.
- They add up to zero.
- So the covariance is zero, and so is the correlation.
- The two variables are perfectly dependent, but not in a straight line.
- The correlation of zero tells us there is no linear relationship, not that there is no relationship.

24 The one-factor model

$$X_i = a_i \times F + \sqrt{(1 - a_i^2)} \times Z_i$$
$$\text{Corr}(X_i, X_j) = a_i \times a_j$$

Variable i	Standard normal variable i
Loading on the factor	Loading; a^2 = variance explained
Common factor	One shared standard normal factor
Own independent noise	Independent noise; drives only i

IN THE LESSON

- Many variables, such as the defaults of many firms, are linked through a single common driver.
- The one factor model captures this.
- Each variable equals its loading, a, times a common factor F, plus the square root of one minus the loading squared, times its own independent noise term Z.
- The factor and every noise term are standard normal and independent of each other.
- Then each variable is itself standard normal.
- The loading squared is the share of the variable's variance explained by the factor.
- The key result: the correlation between any two variables is the product of their loadings.
- Conditional on the factor, the variables are independent, so all the correlation flows through the common factor.

25 Example 10: correlations from loadings

Illustration: loadings $a_1 = 0.8$, $a_2 = 0.6$

Variance explained, firm 1	0.8^2	**64%**
Weight on own noise	$\sqrt{(1 - 0.64)}$	**0.60**
Correlation of firms 1 and 2	0.8×0.6	**0.48**
Inputs for 100 variables	100 vs $100 \times 99 \div 2$	**100 vs 4,950**

IN THE LESSON

- Take two firms with loadings on the factor of point eight and point six.
- The first firm's variance explained by the factor is point eight squared, which is sixty four per cent.
- The rest, thirty six per cent, comes from its own noise, so the weight on its noise term is the square root of point three six, which is point six.
- The correlation between the two firms is the product of the loadings, point eight times point six, which is point four eight.
- Do not add the loadings, and do not square their product.
- The big payoff is economy.
- For one hundred variables there are four thousand nine hundred and fifty pairwise correlations, but the model needs only one hundred loadings, and the resulting matrix is always consistent.

26 Pearson, Spearman and Kendall

Alike: range −1 to +1, zero if independent.
Different: what they use

Measure	What it uses
Pearson	Values; linear dependence only
Spearman's rho	Ranks of the observations
Kendall's tau	Concordant vs discordant pairs
Outliers, monotone maps	Ranks robust; Pearson less so
Bivariate normal link	$\tau = (2/\pi) \times \arcsin(\rho)$

IN THE LESSON

- Correlation has several versions.
- Pearson correlation measures linear dependence.
- It uses the actual values, so it is sensitive to outliers and to the shape of the distributions.
- Rank correlations use only the ordering of the observations.
- Spearman's rho is the Pearson correlation of the ranks.
- Kendall's tau compares pairs of observations, counting those that move in the same direction, called concordant, against those that move in opposite directions, called discordant.
- Both rank measures are unchanged by any strictly increasing transformation of either variable, and they are more robust to outliers.
- All three lie between minus one and plus one, and equal zero for independent variables, but none captures the whole dependence structure.
- For bivariate normal data, Kendall's tau equals two over pi times the arcsine of the Pearson correlation.

27 Example 11: Spearman and Kendall by hand

Illustration: ranks of X: 1,2,3,4,5; ranks of Y: 2,1,4,3,5

Squared rank gaps	1 + 1 + 1 + 1 + 0	**$\Sigma d^2 = 4$**
Spearman	$1 - 6{\times}4 \div (5{\times}24)$	**0.80**
Number of pairs	$5 \times 4 \div 2$	**10**
Concordant, discordant	8 and 2	**C=8, D=2**
Kendall	$(8 - 2) \div 10$	**0.60**

IN THE LESSON

- Let's compute both rank measures on five made up observations.
- Rank the first variable one to five.
- The second variable's ranks come out as two, one, four, three, five.
- The rank differences are one, one, one, one and zero, so the sum of squared differences is four.
- Spearman equals one minus six times four, over five times twenty four, which is one minus twenty four over one hundred and twenty, so point eight.
- For Kendall, there are five times four over two, ten pairs.
- Eight pairs move in the same direction and two do not: the swapped neighbours.
- Kendall is eight minus two, over ten, which is point six.
- Both are positive and below one, and Kendall is smaller.

28 Example 12: monotone but not linear

Illustration: X = 1,2,3,4,5 and Y = X³ = 1, 8, 27, 64, 125

Spearman and Kendall	ranks identical, all pairs agree	**1 and 1**
Pearson (values)	not a straight line	**0.94**
Normal link, ρ = 0.5	(2/π) × arcsin(0.5)	**1/3**

IN THE LESSON

- One last example shows why rank measures are useful.
- Take X equal to one, two, three, four, five, and let Y be X cubed, so one, eight, twenty seven, sixty four, one hundred and twenty five.
- Y always rises when X rises, an exact monotone relation, but it is curved.
- The ranks of both variables are identical, one to five, so every rank difference is zero.
- Spearman is exactly one, and every pair is concordant, so Kendall is exactly one.
- The Pearson correlation, using the values, is only about point nine four, below one, because the relation is not a straight line.
- Finally, check the bivariate normal link.
- A Pearson correlation of point five gives Kendall of two over pi times arcsine of point five, which is one third.

29 Four mistakes to avoid

- Adding simple returns: compound, or add logs
- Daily vol × days: use × √days instead
- JB with raw K: use K − 3; watch sample size
- Zero correlation ≠ independence (linear only)

IN THE LESSON

- Let's collect the four traps from this lesson.
- First, never add simple returns across time as though they were log returns.
- Compound them, or add the logs and convert back.
- Second, do not multiply a daily volatility by the number of days.
- Multiply by the square root of the number of days, because variance is what adds.
- Third, in the Jarque Bera formula, use kurtosis minus three, not the raw kurtosis, and remember that the statistic grows with the sample size.
- Fourth, do not read a zero correlation as the absence of any relationship, because correlation sees only linear association.
- Keep these four fixes in mind when a question looks easy.

30 Recap

- Logs add; simple returns compound; convert back
- Volatility × √T; implied volatility looks forward
- Skewness, excess kurtosis and JB test for non-normality
- Power law: P = k × x^{−α}, a straight line in logs
- Corr = Cov ÷ σσ; one factor a_i×a_j; rank measures

IN THE LESSON

- Let's recap the whole module.
- A simple return is the percentage change including income, and the log return is the natural log of the gross return.
- Log returns add over time, simple returns compound, and you convert back to report gains.
- Volatility is the standard deviation per unit of time, the variance rate is its square, and volatility scales with the square root of time.
- Implied volatility is forward looking and comes from option prices.
- For non normal returns, skewness and excess kurtosis matter, and the Jarque Bera statistic tests for normality.
- The power law describes fat tails with a straight line in logs.
- Correlation is covariance scaled, the one factor model gives correlations as products of loadings, and rank measures such as Spearman and Kendall capture monotone dependence.

31 Quick check

1. Log returns of 5% and 4% over two years: total?
   Answer: They add: 9% log return
2. Daily volatility 1%: annual volatility?
   Answer: 1% × √252 ≈ 15.87%
3. Tail index 2: double the loss threshold?
   Answer: Probability ÷ 2^2 = ÷ 4

IN THE LESSON

- Time for a quick check.
- Three questions, with the answers on the board.
- Question one.
- A log return for the year is five per cent, then four per cent the next year.
- What is the two year log return?
- Question two.
- A stock has a daily volatility of one per cent.
- What is the approximate annual volatility, using two hundred and fifty two days?
- Question three.
- In a power law with a tail index of two, you double the loss threshold.
- By what factor does the tail probability change?
- Pause the video, think about each one, and then compare your answers with what appears.
- Remember the habits from this lesson: logs add, volatility scales with the square root of time, and doubling the threshold multiplies the probability by two to the power of minus alpha.