Non-Parametric Approaches

FRM Part II

1 A story: one number, how sure?

- Halden Bank's equity desk: 500 days of daily losses
- 99% VaR from history: $5.6m. How precise is it?
- Can we read 99.2%? Should an old crash count fully?
- Non-parametric: let the data speak, assume no shape

IN THE LESSON

- Meet Halden Bank's equity desk.
- Its risk manager keeps five hundred days of daily profit and loss, replayed on today's positions.
- From that history she reports a ninety nine per cent V a R of five point six million dollars.
- The head of trading asks three awkward questions.
- How precise is that number, given that it rests on a handful of bad days?
- Can we read the V a R at ninety nine point two per cent, when no single day sits exactly there?
- And should a crash from almost two years ago count as much as yesterday?
- This lesson answers all three.
- The common thread is that we let the data speak, and we never assume a shape for the distribution.
- That is what non parametric means.

2 Basic historical simulation, and our convention

- Apply each past day's moves to today's portfolio
- Rank the 500 simulated losses, worst first
- Tail at 99%: 1% × 500 = 5 losses
- VaR: the 5th worst loss. ES: average of those 5
- Every day weighs the same: 1 ÷ 500 = 0.2%

IN THE LESSON

- First, a reminder of basic historical simulation.
- We take each past day's market moves and apply them to the portfolio we hold today, which gives five hundred simulated losses.
- We rank them from worst to best.
- At ninety nine per cent confidence, the tail holds one per cent of five hundred, which is five losses.
- Conventions differ slightly between texts, so here is ours.
- V a R is the fifth worst loss, the smallest loss in the tail, and expected shortfall, E S, is the average of the five tail losses.
- Notice the weights.
- Every day gets the same probability, one over five hundred, which is point two per cent, and every return is used exactly as it happened.

3 Example 1: Halden's VaR and expected shortfall

Halden's five worst daily losses of 500 days, $m: 9.4, 7.8, 6.9, 6.1, 5.6

Tail size at 99%	1% × 500	**5 losses**
99% VaR	the 5th worst loss	**$5.6m**
Sum of the five	9.4+7.8+6.9+6.1+5.6	**$35.8m**
99% ES	35.8 ÷ 5	**$7.16m**

IN THE LESSON

- Here are Halden's five worst days, in millions of dollars: nine point four, seven point eight, six point nine, six point one and five point six.
- The tail size at ninety nine per cent is one per cent of five hundred, so five losses.
- The ninety nine per cent V a R is the fifth worst loss, five point six million dollars.
- For expected shortfall, add the five: the sum is thirty five point eight million.
- Divide by five, and E S is seven point one six million dollars.
- Keep these two numbers in mind.
- Every method in the lesson either measures how precise they are, or changes them by smoothing or reweighting the same history.

4 Coherent risk measures

- Monotonic: a portfolio that always loses more is riskier
- Subadditive: risk(A + B) ≤ risk(A) + risk(B)
- Homogeneous: double the position, double the risk
- Translation invariant: added cash cuts risk by that cash
- VaR can fail subadditivity, so it is not coherent

IN THE LESSON

- Before we resample anything, we need to know which risk measures are worth estimating.
- A coherent risk measure satisfies four common sense properties.
- Monotonic: if one portfolio always loses more than another, it is riskier.
- Subadditive: the risk of two positions combined is never more than the sum of their separate risks, so diversification never hurts.
- Homogeneous: double the position and you double the risk.
- Translation invariant: adding cash to a portfolio cuts its risk by exactly that cash.
- V a R can fail subadditivity, so it is not coherent.
- Expected shortfall passes all four tests.

5 Example 2: VaR breaks subadditivity

Two independent loans, each loses $100 with a 4% chance. 95% VaR?

Loan A alone: chance of loss	4%, below 5%	**VaR $0**
Loan B alone	the same	**VaR $0**
Both: at least one loses	1 − 0.96^2 = 7.84%	**above 5%**
Both together: 95% VaR	7.84% > 5%	**$100**
Sum of stand-alone VaRs	0 + 0	**$0**

IN THE LESSON

- Here is how V a R fails.
- Take two independent loans, each losing one hundred dollars with a four per cent chance.
- Loan A alone: the chance of a loss is four per cent, below five, so its ninety five per cent V a R is zero.
- Loan B alone is the same, V a R zero.
- Now hold both.
- The chance that at least one loses is one minus point nine six squared, which is seven point eight four per cent, above five per cent.
- So the combined ninety five per cent V a R is one hundred dollars.
- Yet the sum of the stand alone V a Rs is zero plus zero, which is zero.
- Combining the loans seems to create risk, which is exactly what subadditivity forbids.

6 Spectral risk measures

$M = \Sigma\, w_{i} \times L_{(i)}$
$w \geq 0$, $\Sigma\, w = 1$, w never falls as losses worsen

Weight on each ranked loss	Non-negative, sum to 1
Losses ranked, best to worst	Best to worst, $L_{(1)}$ to $L_{(n)}$
Expected shortfall	Equal weights on the tail
Value at risk	All weight on one loss: fails

IN THE LESSON

- Expected shortfall belongs to a wider coherent class, the spectral risk measures.
- Line the losses up from best to worst, and take a weighted average of them.
- The weights must follow three rules.
- They are never negative, they add up to one, and they never fall as the losses get worse.
- So a worse loss always counts at least as much as a milder one.
- Expected shortfall is the simplest case: an equal weight on each tail loss, here one fifth on each of the five worst, and zero elsewhere.
- V a R puts all its weight on one loss, the fifth worst, and zero on the four worse ones.
- Its weights fall as losses worsen, which breaks the third rule, and that is why V a R is not spectral.

7 Example 3: a spectral measure for Halden

Halden's five worst losses; weights rise as losses worsen; every other loss gets 0

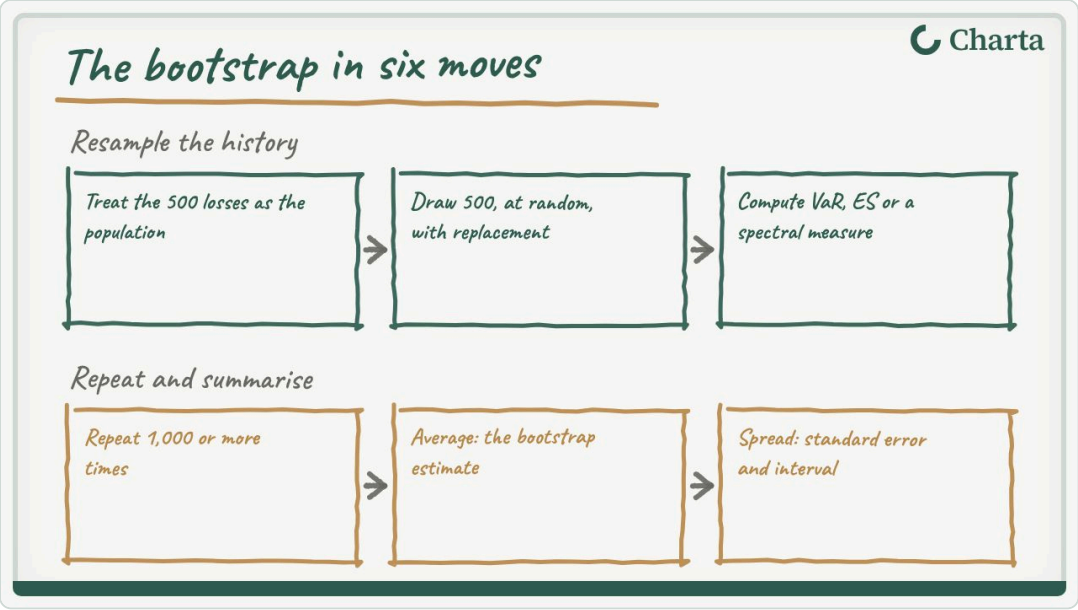
Worst: 9.4	0.30 × 9.4	**2.82**
2nd worst: 7.8	0.25 × 7.8	**1.95**
3rd worst: 6.9	0.20 × 6.9	**1.38**
4th and 5th: 6.1, 5.6	0.15 × 6.1 + 0.10 × 5.6	**0.915 + 0.56**
Spectral measure	2.82+1.95+1.38+0.915+0.56	**≈ $7.63m**

Above the ES of $7.16m, which weights each tail loss 0.2

IN THE LESSON

- Halden's risk committee wants a measure that is more averse to the very worst days, so it puts weights on the five worst losses that rise as the losses worsen, and zero on all the others.
- The worst loss, nine point four, gets point three, giving two point eight two.
- The second, seven point eight, gets point two five, giving one point nine five.
- The third, six point nine, gets point two, giving one point three eight.
- The fourth, six point one, gets point one five, giving point nine one five.
- The fifth, five point six, gets point one, giving point five six.
- The weights add to one and never fall as losses worsen.
- The spectral measure is the sum of the five products, about seven point six three million dollars.
- That is above the expected shortfall of seven point one six, which weights each tail loss point two.

8 The bootstrap in six moves



IN THE LESSON

- How precise are these numbers?
- The bootstrap answers that.
- Look at the first row.
- We treat the five hundred historical losses as if they were the whole population.
- We then draw five hundred losses at random, with replacement, so some days appear several times and others not at all.
- On that resample we compute the risk measure, the V a R, the E S or any spectral measure.
- Now the second row.
- We repeat the draw a thousand times or more.
- The average of the resampled estimates is the bootstrap estimate.
- And the spread of the estimates gives a standard error and a confidence interval.

9 Example 4: four resamples of Halden's history

Each resample draws 500 days with replacement. Its five worst draws, $m

Resample 1	9.4, 9.4, 7.8, 6.9, 6.1	**ES 7.92**
Resample 2	7.8, 6.9, 6.1, 6.1, 5.6	**ES 6.50**
Resample 3	9.4, 7.8, 6.9, 5.6, 5.6	**ES 7.06**
Resample 4	9.4, 6.9, 6.1, 5.6, 5.2	**ES 6.64**
Bootstrap ES	(7.92+6.50+7.06+6.64) ÷ 4	**$7.03m**

IN THE LESSON

- To keep the sums short we show just four resamples; a real run would use a thousand or more.
- In resample one the worst day was drawn twice, and the five worst draws are nine point four, nine point four, seven point eight, six point nine and six point one, an E S of seven point nine two.
- In resample two the worst day was never drawn, and six point one appears twice, an E S of six point five zero.
- Resample three draws nine point four, seven point eight, six point nine, and five point six twice, an E S of seven point zero six.
- Resample four draws nine point four, six point nine, six point one, five point six and five point two, an E S of six point six four.
- The bootstrap E S is the average of the four: seven point zero three million dollars.
- Notice that every draw is one of Halden's real losses.

10 Example 5: standard error and confidence interval

The spread of the four resampled ES values around their average, 7.03

Deviations from 7.03	0.89, −0.53, 0.03, −0.39	**sum 0**
Sum of squares	0.792+0.281+0.001+0.152	**1.226**
Standard error	√(1.226 ÷ 3)	**$0.64m**
90% interval	7.03 ± 1.645×0.64 = 7.03 ± 1.05	**5.98 to 8.08**
Percentile version	sort 1,000 estimates	**50th and 950th**

IN THE LESSON

- Now the spread.
- The deviations of the four estimates from their average of seven point zero three are point eight nine, minus point five three, point zero three and minus point three nine, which sum to zero.
- Square and add them: one point two two six.
- Divide by three, one less than the number of resamples, and take the square root.
- The standard error is about point six four million dollars.
- A ninety per cent interval, using the normal multiplier of one point six four five, is seven point zero three plus or minus one point six four five times point six four.
- That is seven point zero three plus or minus one point zero five, which runs from five point nine eight to eight point zero eight.
- With a thousand resamples there is a simpler percentile version: sort the estimates and read the fiftieth and the nine hundred and fiftieth.

11 What the bootstrap can and cannot do

- It measures precision: a standard error and an interval
- Averaging many resamples improves precision
- It adds no information: every draw is an old loss
- No resample can hold a loss worse than $9.4m
- Trap: a bootstrap does not extrapolate the tail

IN THE LESSON

- Be clear about what the bootstrap buys.
- It measures precision.
- It gives a standard error and a confidence interval for any risk measure, simply and without extra theory.
- Averaging many resamples also improves precision, because the noise in any single estimate is averaged away.
- But it adds no information.
- Every draw is one of the losses we already had.
- So no resample of Halden's history can ever contain a loss worse than nine point four million dollars.
- The exam trap is to believe that bootstrapping extrapolates into the tail.
- It cannot.
- It only tells you how much to trust what the sample already says.

12 The problem with discrete points

- Basic HS treats the sample as separate points
- VaR can only be one of the observed losses
- It jumps from loss to loss as the level changes
- It jumps when one day enters or leaves the window
- Fix: a smooth density, with no parametric shape

IN THE LESSON

- Now the second question: can we read the V a R at ninety nine point two per cent?
- Basic historical simulation treats the sample as a set of separate points.
- So the V a R can only ever be one of the observed losses.
- As the confidence level changes, the answer jumps from one loss to the next, and nothing in between is possible.
- The V a R also jumps when a single day enters or leaves the window.
- The fix is to fit a smooth, continuous density through the data, without assuming a parametric form such as the normal.
- Then a quantile can be read at any confidence level we like.

13 Each loss as the middle of a bin

Quantile of the i-th smallest loss = (i − 0.5) ÷ n
Between two losses: interpolate in a straight line

Rank i, from the smallest	1 = mildest, n = worst
Number of losses, n	Here 500 days
Quantile level	Slightly below i ÷ n

IN THE LESSON

- The simplest smoothing treats each observation as the midpoint of a bin.
- Rank the losses from the smallest, so i equals one is the mildest and i equals n is the worst.
- The i th smallest loss is taken to sit at the quantile level i minus one half, divided by n.
- That is a little below its plain rank share, i over n, because the loss sits in the middle of its slice of probability, not at the top.
- Between two neighbouring losses we interpolate in a straight line.
- So every confidence level now has an answer, even one that falls between two observations.

14 Example 6: Halden's smoothed quantiles

n = 500. Worst losses, $m: 9.4, 7.8, 6.9, 6.1, 5.6, 5.2 (i = 500 down to 495)

5th worst, 5.6 (i = 496)	(496 − 0.5) ÷ 500	**99.1%**
6th worst, 5.2 (i = 495)	(495 − 0.5) ÷ 500	**98.9%**
99% VaR: halfway	(5.6 + 5.2) ÷ 2	**$5.4m**
4th worst, 6.1 (i = 497)	(497 − 0.5) ÷ 500	**99.3%**
99.2% VaR: halfway	(6.1 + 5.6) ÷ 2	**$5.85m**

Basic HS gave $5.6m at 99%: the convention moves the answer

IN THE LESSON

- Let us apply it to Halden.
- With five hundred losses, the worst has i equal to five hundred, and the sixth worst has i equal to four hundred and ninety five.
- The fifth worst, five point six, sits at four hundred and ninety six minus one half, over five hundred, which is ninety nine point one per cent.
- The sixth worst, five point two, sits at ninety eight point nine per cent.
- Ninety nine per cent lies exactly halfway between, so the smoothed V a R is the average of five point six and five point two, which is five point four million dollars.
- The fourth worst, six point one, sits at ninety nine point three per cent.
- So ninety nine point two per cent lies halfway between six point one and five point six, a V a R of five point eight five million.
- Basic historical simulation gave five point six million at ninety nine per cent, so the convention moves the answer.

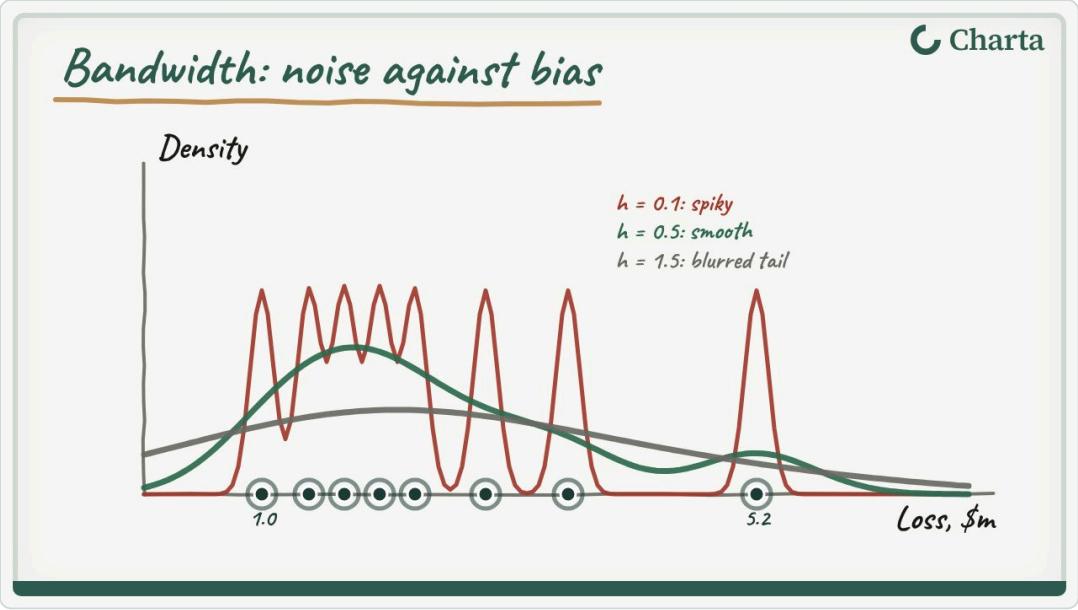
15 Kernel smoothing

- Place a small bump, often a normal curve, on every loss
- Average the bumps: a smooth density, no set shape
- Bandwidth h: the width of each bump, chosen by us
- Read VaR at any level from the smooth cumulative curve

IN THE LESSON

- A second route is the kernel method.
- On every loss in the sample we place a small bump, a kernel, which is often a normal curve centred on that loss.
- We then average all the bumps.
- The result is a smooth density that follows the data wherever the data go, with no shape assumed in advance.
- The width of each bump is called the bandwidth, written h, and it is chosen by the analyst.
- Adding up the area under the smooth density gives a smooth cumulative distribution, and from that curve we can read the V a R, or any other quantile, at any confidence level.

16 Bandwidth: noise against bias



IN THE LESSON

- Here are three kernel densities, all fitted to the same eight illustrative losses, in millions of dollars.
- The red curve uses a bandwidth of point one.
- It is spiky, with a separate peak on every loss, so it is noisy and its quantiles would jump around like the raw data.
- The green curve uses a bandwidth of point five.
- It is smooth, but it still follows the cluster of losses on the left and keeps a small bump for the lone large loss on the right.
- The grey curve uses a bandwidth of one point five.
- It is blurred: the gap before the largest loss has vanished, and the tail spreads far beyond the data.
- That is bias.
- The labels at the top name the three bandwidths, and the dots along the bottom are the eight losses themselves, from one up to five point two.
- The bandwidth trades noise against bias.

17 Smoothing spreads information, it does not add it

- Too small a bandwidth: a spiky, noisy density
- Too large: a blurred tail and biased quantiles
- Both routes read quantiles between the observations
- Tail estimates still rest on the same few extreme losses
- Trap: a smooth curve is not extra data

IN THE LESSON

- So how should the bandwidth be set?
- Too small, and the density is spiky and noisy, and V a R still jumps as days enter and leave.
- Too large, and the tail is blurred, so the extreme quantiles are biased.
- Both kinds of smoothing, the bins and the kernels, help us read quantiles between the observations.
- But neither adds any information.
- Smoothing spreads what the sample knows across the gaps between the losses.
- The tail estimates still rest on the same few extreme days, at Halden just five losses beyond the ninety nine per cent point.
- The exam trap is to think a smooth curve is more data.
- It is the same data, drawn more neatly.

18 Why weight the history?

- Basic HS: each day weighs 0.2%; returns used as they were
- A crash 499 days ago counts as much as yesterday
- Today's volatility and correlations are ignored
- Refinements: one changes weights, three change returns

IN THE LESSON

- Now the third question: should an old crash count as much as yesterday?
- In basic historical simulation every day weighs point two per cent, and every return is used exactly as it happened.
- So a crash four hundred and ninety nine days ago counts as much as yesterday's loss.
- The method also ignores today's conditions.
- If volatility is high now, calm days from the past understate today's risk, and if correlations have risen, the past overstates diversification.
- Four refinements make the sample more relevant to current conditions.
- Age weighting changes the weights.
- Volatility weighting, correlation weighting and filtered historical simulation change the returns.

19 Age-weighted historical simulation

w(i) = λ^{i−1} × (1 − λ) ÷ (1 − λ^{n})
Rank losses; add weights from the worst until 1% is reached

Decay factor, λ	A little below 1, here 0.98
Observation number, i	1 = the most recent day
Number of losses, n	Days in the window, 500

IN THE LESSON

- Age weighting gives recent days more probability.
- Number the days so that observation one is the most recent.
- Here i counts days back in time, not loss rank as before.
- Pick a decay factor lambda, a little below one.
- Observation i gets lambda to the power i minus one, times one minus lambda, divided by one minus lambda to the power n.
- The last part only makes the weights add up to one.
- Each day back in time, the weight shrinks by the factor lambda.
- To find the V a R, rank the losses from the worst, add up their weights as you go, and stop at the loss where the running total first reaches one per cent.
- The losses are unchanged; only their probabilities move.

20 Example 7: age-weighted VaR at Halden

λ = 0.98, n = 500, so 1 − 0.98^500 ≈ 1. Halden's four worst losses, $m

9.4, observation 500	0.98^{499} × 2% ≈ 0.0001%	**total ≈ 0.0001%**
7.8, observation 40	0.98^{39} × 2% = 0.91%	**total 0.91%**
6.9, observation 200	0.98^{199} × 2% = 0.04%	**total 0.95%**
6.1, observation 12	0.98^{11} × 2% = 1.60%	**total 2.55%**
Age-weighted 99% VaR	total first reaches 1%	**$6.1m**

Equal-weight HS gave $5.6m at 99%

IN THE LESSON

- Halden uses lambda equal to point nine eight.
- With five hundred days, one minus point nine eight to the five hundred is practically one, so each weight is about point nine eight to the power i minus one, times two per cent.
- The worst loss, nine point four, is observation five hundred, the oldest day, and its weight is about one ten thousandth of one per cent.
- The seven point eight loss is observation forty, weight point nine one per cent, so the running total is point nine one.
- The six point nine loss is observation two hundred, weight point zero four per cent, total point nine five.
- The six point one loss is observation twelve, weight one point six zero per cent, total two point five five.
- The total first reaches one per cent here, so the age weighted V a R is six point one million, above the equal weight answer of five point six.

21 Volatility-weighted historical simulation

Adjusted return = r_{t} × σ_{now} ÷ σ_{t}
Weights stay equal; the returns are rescaled

Historical return	Return on past day t
Volatility today	Today's estimate
Volatility on that day	The estimate on day t

IN THE LESSON

- Volatility weighting keeps the equal weights but changes the returns.
- Each historical return is multiplied by today's volatility, divided by the volatility on the day it happened.
- For a simple linear position, the loss scales the same way.
- The volatilities usually come from a model such as G A R C H or an exponentially weighted moving average.
- The idea is simple.
- A loss that happened in a storm is scaled down if today is calm, and a loss from a calm day is scaled up if today is stormy.
- After rescaling, we rank the adjusted losses and read the V a R and E S exactly as before, with every day still weighing point two per cent.

22 Example 8: volatility-weighted losses

Halden's σ today: 1.2% a day. Rescale all 500 losses by 1.2% ÷ σ that day, then re-rank

9.4, σ then 2.35%	9.4 × 1.2 ÷ 2.35	**$4.80m**
7.8, σ then 1.5%	7.8 × 1.2 ÷ 1.5	**$6.24m**
6.9, σ then 0.9%	6.9 × 1.2 ÷ 0.9	**$9.20m**
6.1, σ then 1.2%	6.1 × 1.2 ÷ 1.2	**$6.10m**
Largest of these four	the calm-day loss	**$9.20m**

IN THE LESSON

- Halden's daily volatility today is one point two per cent.
- All five hundred losses are rescaled by one point two per cent over the volatility on their own day, then ranked again.
- The nine point four loss happened when volatility was two point three five per cent, so it becomes nine point four times one point two over two point three five, which is four point eight zero million.
- The seven point eight loss came at one point five per cent volatility and becomes six point two four.
- The six point nine loss came on a calm day, at point nine per cent, and grows to nine point two zero.
- The six point one loss came at one point two per cent, the same as today, so it stays at six point one zero.
- The largest of these four is now the loss from the calm day, nine point two zero million dollars.

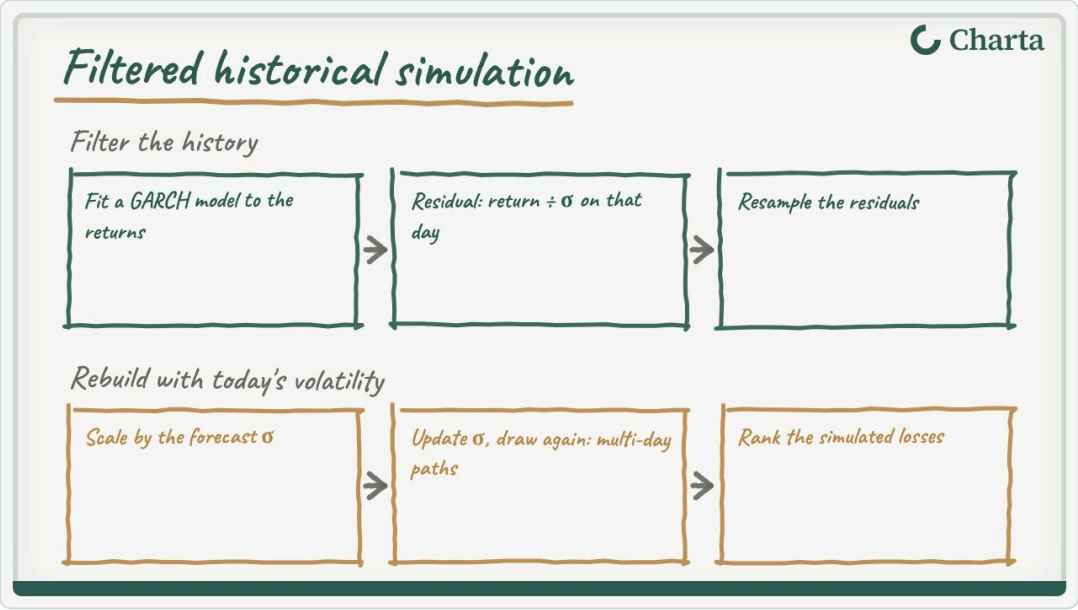
23 Correlation-weighted historical simulation

- Adjust the returns so their covariance matches today's
- Volatility weighting, extended to the covariance matrix
- Two desks: ρ = 0.2 in the sample, 0.7 today
- Unadjusted history overstates diversification
- Equal weights stay; the returns themselves change

IN THE LESSON

- Correlation weighting goes one step further.
- Instead of rescaling each return by its own volatility, it adjusts the whole set of returns so that their covariance structure reflects today's correlations, not the correlations in the sample.
- It is volatility weighting extended to the full covariance matrix.
- Suppose Halden has two desks whose returns had a correlation of point two over the sample, but of point seven today.
- Unadjusted history would show losses on the two desks often offsetting each other, which overstates diversification.
- After correlation weighting, the scenarios show them falling together more often.
- As with volatility weighting, every day keeps an equal weight; the returns themselves are changed.

24 Filtered historical simulation



IN THE LESSON

- The most complete refinement is filtered historical simulation, F H S.
- First row.
- We fit a volatility model such as G A R C H to the returns.
- We standardise each return by dividing it by the model's volatility on that day, which gives a residual.
- Then we resample those residuals, as in the bootstrap.
- Second row.
- We scale each drawn residual by the forecast volatility, giving a simulated return for tomorrow.
- Because the model updates its volatility after each simulated day, we can draw again and build paths over several days.
- Finally, we rank the simulated losses and read off the V a R and E S.

25 Example 9: a two-day filtered path

Portfolio $200m. GARCH, in %²: ω 0.16, α 0.10, β 0.85. Forecast σ 1.2%

Residual: the 9.4 day	−9.4 ÷ 200 = −4.7%; ÷ 2.35%	**z = −2.0**
Day 1 return	−2.0 × 1.2%	**−2.40%: $4.8m**
Day 2 variance	0.16+0.10×2.4^2+0.85×1.2^2	**1.96: σ 1.4%**
Day 2: draw z = −1.5	−1.5 × 1.4%	**−2.10%**
Two-day loss	≈ (2.40% + 2.10%) × 200	**≈ $9.0m**

IN THE LESSON

- Halden's portfolio is worth two hundred million dollars, so its nine point four million loss was a return of minus nine point four over two hundred, or minus four point seven per cent, on a day when the model's volatility was two point three five.
- The residual is minus two point zero.
- Scaled by today's forecast of one point two per cent, day one's return is minus two point four per cent, which is a loss of four point eight million, the same as volatility weighting.
- Now update.
- Day two variance is point one six, plus point one times two point four squared, which is five point seven six, plus point eight five times one point two squared, one point four four.
- That totals one point nine six, a volatility of one point four per cent.
- Draw another residual, say minus one point five, for a return of minus two point one per cent.
- Adding the two returns, a close approximation, gives a two day loss of about nine million dollars.

26 Four refinements side by side

Each makes the history more relevant to today, in a different way

Method	What it changes
Age-weighted	Weights: recent days count more
Volatility-weighted	Returns × σ now ÷ σ then
Correlation-weighted	Returns fit today's covariance
Filtered (FHS)	GARCH residuals × forecast σ

IN THE LESSON

- Each of the four makes the history more relevant to today, in a different way.
- Let us set them side by side, reading down the table.
- Age weighting changes the probability weights, so recent days count more, and leaves the returns alone.
- Volatility weighting rescales each return by volatility now over volatility then, and keeps the weights equal.
- Correlation weighting adjusts the returns so their covariance matches today's.
- Filtered historical simulation rescales G A R C H residuals by forecast volatility, and can build multi day paths.
- Notice that only the first row changes the weights; the other three change the returns.

27 What they share, and how they differ

- Shared: every scenario comes from actual history
- Shared: no assumed shape for the return distribution
- Age weighting alone changes the weights, not the returns
- Volatility, correlation and FHS rescale returns to today
- Only FHS builds multi-day paths with changing volatility

IN THE LESSON

- What do the four share?
- Every scenario still comes from actual historical observations, so the method stays rooted in what really happened.
- And none of them assumes a parametric shape for the distribution of returns, so skew and fat tails in the data survive.
- How do they differ?
- Age weighting alone changes the weights and not the returns.
- Volatility weighting, correlation weighting and filtered simulation rescale the returns to today's conditions.
- Only filtered historical simulation builds multi day paths with a volatility that changes along the way.
- A typical exam question gives you one of these descriptions and asks which method it is.

28 Strengths of non-parametric methods

- Intuitive, easy to implement and to explain
- No assumed distribution: skew and fat tails are kept
- No covariance matrix to estimate
- Options and non-linear positions revalued in each scenario
- Bootstrapping and weighting add precision and relevance

IN THE LESSON

- Now let us weigh the whole family, starting with the strengths.
- Non parametric methods, whether plain historical simulation or its bootstrapped and weighted versions, are intuitive and easy to implement and to explain to a board.
- They assume no distribution for returns, so the skewness and fat tails that really appear in the data are kept.
- They need no covariance matrix, which for a large book is a big saving.
- And they handle options and other non linear positions naturally, because each position is fully revalued in every historical scenario rather than approximated by a straight line.
- The refinements we have seen add precision and relevance on top.

29 Weaknesses: all from relying on the sample

- Results depend on the window chosen
- Assumes the past represents the future
- Equal weights react slowly to regime shifts
- Ghost effect: VaR falls when a big loss leaves the window
- Extreme quantiles are capped by the worst loss observed

IN THE LESSON

- The weaknesses all come from relying on the sample.
- Results depend on the window chosen: five hundred days and one thousand days can give quite different answers.
- The method assumes the past represents the future, so a risk that has not happened in the window is invisible.
- With equal weights it reacts slowly to a regime shift, since one new bad day is only one day in five hundred.
- When a large loss leaves the window, V a R and E S can fall abruptly although nothing has changed in the markets.
- This is called the ghost or drop off effect.
- And the extreme quantiles are capped by the worst loss observed.

30 Example 10: the ghost effect at Halden

The $9.4m loss, observation 500, leaves the window tomorrow; tomorrow's loss is small

Today's tail	9.4, 7.8, 6.9, 6.1, 5.6	**VaR $5.6m**
Today's ES	35.8 ÷ 5	**$7.16m**
Tomorrow's tail	7.8, 6.9, 6.1, 5.6, 5.2	**VaR $5.2m**
Tomorrow's ES	31.6 ÷ 5	**$6.32m**
Change, markets unchanged	6.32 − 7.16	**−$0.84m**

IN THE LESSON

- Let us see the ghost effect at Halden.
- The nine point four million loss is the oldest day in the window, and tomorrow it drops out.
- Suppose tomorrow is an ordinary day that does not enter the tail.
- Today's tail is nine point four, seven point eight, six point nine, six point one and five point six, so V a R is five point six million.
- Today's E S is thirty five point eight over five, seven point one six.
- Tomorrow's tail is seven point eight, six point nine, six point one, five point six and five point two, so V a R falls to five point two.
- Tomorrow's E S is thirty one point six over five, six point three two.
- E S falls by point eight four million, with nothing changed in the markets.
- Age weighting avoids this cliff, because old losses fade gradually.

31 Example 11: thin data in the tail

500 days of history. A 10-day VaR from non-overlapping 10-day periods

Independent 10-day periods	500 ÷ 10	**50**
Tail at 99%, 10-day	1% × 50	**0.5 losses**
Tail at 99%, 1-day	1% × 500	**5 losses**
Overlapping 10-day periods	more numbers	**not independent**
New instrument, 60 days	1% × 60	**0.6 losses**

IN THE LESSON

- The last weakness is thin data.
- Halden's regulator asks for a ten day V a R.
- Five hundred days of history give only five hundred over ten, which is fifty, independent ten day periods.
- The ninety nine per cent tail of fifty periods holds one per cent of fifty, half an observation, so there is not even one loss to read.
- Compare the one day V a R, where the tail held five losses.
- Overlapping ten day periods give more numbers, but they are not independent.
- A new instrument is worse still.
- With sixty days of history, the ninety nine per cent tail holds point six of an observation.
- Non parametric methods cannot extrapolate beyond the data they have.

32 Mistakes to avoid

- A bootstrap cannot create losses worse than the sample
- Smoothing adds no information to the tail
- Equal-weight HS does not react quickly to new risk
- Age weighting changes weights; volatility weighting, returns
- VaR is not coherent; ES and spectral measures are

IN THE LESSON

- Here are the mistakes to avoid.
- First, thinking that a bootstrap can create losses worse than those in the sample.
- It only reshuffles them.
- Second, thinking that smoothing adds information to the tail.
- It spreads the same few points.
- Third, believing that equal weight historical simulation reacts quickly to new risk.
- It is slow, and then it drops old losses with a jolt.
- Fourth, mixing up age weighting, which changes the weights, with volatility weighting, which changes the returns.
- And fifth, forgetting that V a R is not coherent, while expected shortfall and the other spectral measures are.
- The usual exam error is to credit these methods with the power to extrapolate or to react quickly.

33 Try three quick questions

1. Worst losses \$10m, \$8m, \$6m; weights 0.5, 0.3, 0.2. Spectral measure?
   Answer: 5 + 2.4 + 1.2 = \$8.6m.
2. λ = 0.95, long window. Age weight of observation 3?
   Answer: 0.95^2 × 0.05 ≈ 4.51%.
3. n = 200. Quantile level of the 3rd worst loss?
   Answer: i = 198: (198 − 0.5) ÷ 200 = 98.75%.

IN THE LESSON

- Now try three quick questions.
- Pause the video, think, then listen to the answers.
- Question one.
- The three worst losses are ten, eight and six million dollars, with spectral weights of point five, point three and point two, the largest on the worst loss.
- What is the risk measure?
- Five plus two point four plus one point two, which is eight point six million dollars.
- Question two.
- With a decay factor of point nine five and a long window, what weight does age weighting give to observation three?
- Point nine five squared times point zero five, about four point five one per cent.
- Question three.
- In a sample of two hundred losses, at what quantile level does the third worst loss sit?
- It is the one hundred and ninety eighth smallest, so one hundred and ninety seven point five over two hundred, ninety eight point seven five per cent.