CHARTA LEARNING · LESSON NOTES

Backtesting VaR

FRM Part II

1 A promise you cannot see

- Ashby's desk: one-day 99% VaR of about $10m
- Promise: a loss above VaR on only 1 day in 100
- VaR is a quantile: it cannot be observed directly
- So how do we know the model is right?
- Answer: compare forecasts with outcomes

IN THE LESSON

- Ashby Bank's equity desk reports a one day value at risk, written V a R, of about ten million dollars at ninety nine per cent confidence.
- In plain words, the model promises that on only one day in a hundred will the desk lose more than that.
- Here is the difficulty.
- V a R is a quantile, a cut off point in the distribution of tomorrow's profit and loss, and nobody can observe a distribution directly.
- So how does anyone know whether the model is right?
- The only way is to compare what the model forecast with what actually happened, day after day.
- That comparison is called backtesting.

2 Backtesting and exceptions

- Each day: VaR made beforehand, P&L realised after
- Exception: a day when the loss is larger than VaR
- Also called a failure or a violation
- Correct model: chance of exception = p = 1 − confidence
- 99% VaR: about 1% of days, or 2.5 in 250

IN THE LESSON

- Here is the method.
- Each day, record the V a R forecast made beforehand, at the start of the day, and the profit or loss realised afterwards.
- An exception is a day on which the loss is larger than that V a R.
- It is also called a failure or a violation.
- If the model is right, the chance of an exception on any day equals the tail probability, which we write p.
- It is one minus the confidence level.
- So a ninety nine per cent one day V a R should be exceeded on about one per cent of days.
- Over a trading year of two hundred and fifty days, that is two and a half exceptions on average.

3 Example 1: spotting an exception

Four of Ashby's days, $m. Exception if the loss > that morning's VaR

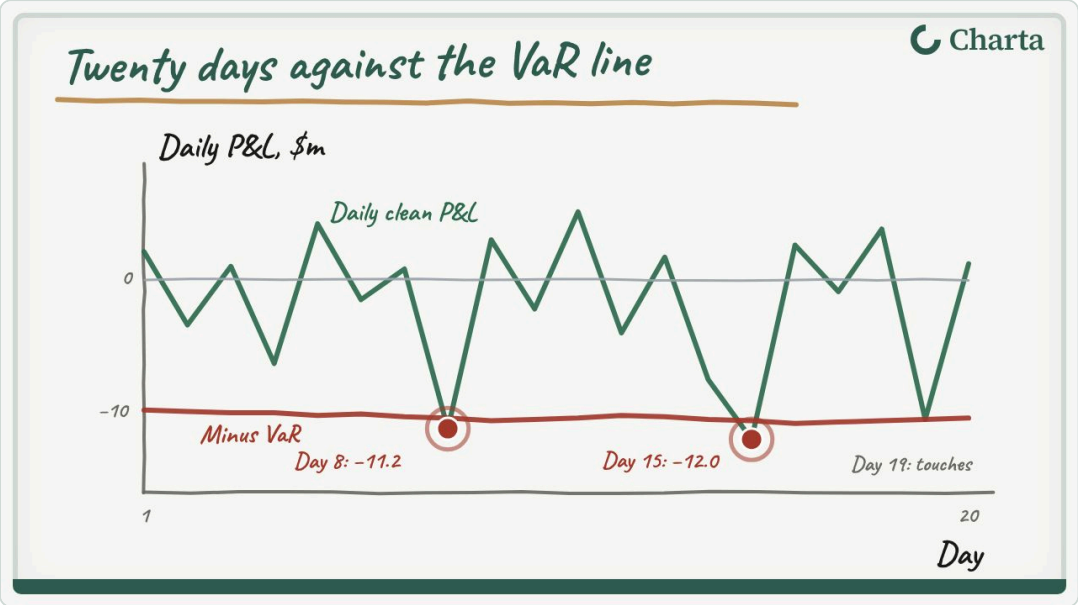
Day 4	VaR 10.0, loss 6.3	**No**
Day 8	VaR 10.4, loss 11.2	**Exception**
Day 11	VaR 10.4, gain 5.1	**No**
Day 19	VaR 10.5, loss 10.5	**No: not larger**

Compare each loss with the VaR set for that same day

IN THE LESSON

- Here are four days from a twenty day stretch on Ashby's desk, in millions of dollars.
- On day four the morning V a R is ten point zero and the desk loses six point three.
- No exception.
- On day eight the V a R is ten point four and the loss is eleven point two.
- That is an exception.
- On day eleven the V a R is ten point four and the desk makes a gain of five point one.
- A gain can never be an exception.
- On day nineteen the V a R is ten point five and the loss is exactly ten point five.
- An equal loss is not larger, so no exception.
- Notice that each loss is compared with the V a R set for that same day.

4 Twenty days against the VaR line



IN THE LESSON

- Here is the whole stretch of twenty days.
- Along the bottom are days one to twenty.
- Up the side is the daily profit or loss in millions of dollars, with zero and minus ten marked.
- The green line is the desk's clean profit and loss.
- The red line underneath is minus the V a R.
- It drifts between about minus ten and minus eleven, because the forecast changes every morning.
- On two days the green line breaks through the red one.
- On day eight the desk lost eleven point two, and on day fifteen it lost twelve against a V a R of ten point six.
- Those are the exceptions.
- On day nineteen the green line only touches the red one, so it is not an exception.

5 Example 2: how many to expect

Ashby's year: 250 trading days at 99%; 6 exceptions recorded

Tail probability, 99% VaR	1 − 0.99	**1%**
Expected in 250 days	1% × 250	**2.5**
At 95%, for comparison	5% × 250	**12.5**
Ashby's exception rate	6 ÷ 250	**2.4%**
Against expectation	6 ÷ 2.5	**2.4 times**

IN THE LESSON

- Now the whole year.
- Ashby's desk ran its ninety nine per cent V a R for two hundred and fifty trading days and recorded six exceptions.
- First the tail probability: one minus point nine nine, which is one per cent.
- Expected exceptions in two hundred and fifty days: one per cent times two hundred and fifty, which is two point five.
- For comparison, a ninety five per cent V a R has a tail probability of five per cent, so it should be exceeded about twelve point five times in the same year.
- Ashby's exception rate is six over two hundred and fifty, two point four per cent.
- That is two point four times what the model promised.
- Is that enough to reject the model, or just bad luck?
- The tests in this lesson decide.

6 Why backtesting matters

- VaR cannot be seen: only outcomes can test it
- Too many exceptions: risk understated, capital inadequate
- Too few exceptions: risk overstated, capital wasted
- Too few is a warning sign, just as too many is
- Regulators use the count to set capital multipliers

IN THE LESSON

- Why does this matter so much?
- Because V a R cannot be seen, only outcomes can test it.
- The errors go both ways.
- Too many exceptions mean the model understates risk, so the capital held against it is inadequate and the bank is weaker than it looks.
- Too few exceptions mean the model overstates risk.
- That sounds safe, but it wastes capital, and it may mean the model is badly built in other ways.
- So a backtest is two sided: too few is a warning sign, just as too many is.
- Finally, regulators use the count of exceptions to set capital multipliers, which we meet at the end of the lesson.

7 Obstacle 1: the P&L is contaminated

- VaR describes the portfolio held at the start of the day
- Actual P&L adds intraday trading, fees, commissions
- Also reserve changes and carry: it is contaminated
- Hypothetical (clean) P&L: revalue the unchanged positions
- VaR is compared with both actual and clean P&L

IN THE LESSON

- A V a R model is hard to backtest, for four reasons.
- The first is that the profit and loss is contaminated.
- V a R describes the portfolio held at the start of the day, frozen.
- But the reported, or actual, profit and loss also contains trading during the day, fees and commissions, changes in reserves, and carry, which is the income from simply holding positions, such as interest.
- None of that was in the V a R.
- So banks also compute a hypothetical, or clean, profit and loss.
- They revalue the unchanged positions on which the V a R was based.
- Banks compare V a R with both figures.

8 Example 3: an exception hidden by trading

Day 8 again: VaR $10.4m on the start-of-day portfolio

Hypothetical P&L	unchanged positions revalued	**−$11.2m**
Exception on clean P&L?	11.2 > 10.4	**Yes**
Add intraday trading gains	made during the day	**+$1.4m**
Add fees and commissions	earned during the day	**+$0.6m**
Actual P&L	−11.2 + 1.4 + 0.6	**−$9.2m**

Actual P&L: 9.2 < 10.4, no exception. Contamination can hide or create one

IN THE LESSON

- Here is contamination at work on day eight.
- That morning's V a R was ten point four million dollars.
- The loss of eleven point two million we counted earlier is the hypothetical profit and loss, from revaluing the unchanged positions.
- Eleven point two is larger than ten point four, so it is an exception.
- But during the day the traders also made one point four million from intraday trading, and the desk earned point six million of fees and commissions.
- Add those in, and the actual profit and loss is minus eleven point two plus one point four plus point six, which is minus nine point two million.
- On the actual figure the exception disappears.
- Contamination can hide an exception or create a false one.

9 Obstacles 2 to 4

- Positions change daily: each VaR is a different portfolio
- So the sample is not repeated draws from one distribution
- Exceptions are rare at 99%: 2.5 expected in 250 days
- Little power, and one cluster can dominate the count
- VaR is one quantile: silent on the size of losses beyond

IN THE LESSON

- The second obstacle is that positions change constantly.
- Each day's V a R belongs to a different portfolio, so the year of results is not two hundred and fifty repeated draws from one distribution.
- The third obstacle is rarity.
- At ninety nine per cent only about two and a half exceptions are expected in two hundred and fifty days.
- With so few events, a test has little power to tell a good model from a bad one, and a single cluster of bad days can dominate the whole count.
- The fourth obstacle is that V a R is a single quantile.
- An exception tells you the line was crossed, but nothing about how far beyond it the loss went.

10 Example 4: same count, very different tails

Two desks, each with VaR $10m and 3 exceptions in 250 days

Desk A: losses on exception days	10.5, 10.8, 11.1	**$m**
Desk A: average excess	(0.5 + 0.8 + 1.1) ÷ 3	**$0.8m**
Desk B: losses on exception days	18, 25, 32	**$m**
Desk B: average excess	(8 + 15 + 22) ÷ 3	**$15m**
Backtest verdict for both	same count: 3	**Same**

A count ignores tail size: one reason capital now uses expected shortfall

IN THE LESSON

- Two desks each have a V a R of ten million dollars, and each records three exceptions in two hundred and fifty days.
- On desk A's exception days the losses were ten point five, ten point eight and eleven point one million.
- The excesses over V a R are point five, point eight and one point one, an average of point eight million.
- On desk B's exception days the losses were eighteen, twenty five and thirty two million.
- The excesses are eight, fifteen and twenty two, an average of fifteen million.
- A count based backtest gives both the same verdict.
- Yet desk B's tail is far more dangerous.
- That is the price of testing a single quantile.
- A count ignores the size of the tail, which is one reason capital rules now use expected shortfall.

11 The proportion of failures test

- Kupiec: are x exceptions in N days consistent with p?
- Null hypothesis: the true failure probability equals p
- Compare the likelihood at x ÷ N with the likelihood at p
- Two-sided: too many or too few can be rejected
- Statistic is chi-square with 1 degree of freedom

IN THE LESSON

- The proportion of failures test, due to Kupiec, asks whether x exceptions in N days is consistent with the tail probability p.
- The null hypothesis is that the model is correct, so the true chance of an exception each day equals p.
- The idea is to compare two likelihoods.
- One uses the failure rate the data suggest, x over N.
- The other uses the rate the model promises, p.
- If they are far apart, in either direction, the model is rejected, so the test is two sided: too many or too few exceptions can both fail it.
- If the model is correct, the statistic follows a chi square distribution with one degree of freedom.

12 The Kupiec likelihood ratio

ln L(q) = (N − x) ln(1 − q) + x ln q
LR = 2 × [ln L(x ÷ N) − ln L(p)]
Reject if LR > 3.84 (5% level) or 6.63 (1% level)

Days, N	Trading days in the backtest
Exceptions, x	Exceptions observed
Tail probability, p	1 − confidence level
Likelihood ratio, LR	Chi-square, 1 degree of freedom

IN THE LESSON

- Here is the machinery.
- The log likelihood of seeing x exceptions in N days, if the failure probability is q, is N minus x, all times the natural log of one minus q, plus x times the natural log of q.
- The likelihood ratio, L R, is twice the difference between two log likelihoods: the unrestricted one, at q equal to x over N, minus the restricted one, at q equal to p.
- Because x over N fits the data best, L R is never negative.
- Reject the model when L R exceeds three point eight four at the five per cent level, or six point six three at the one per cent level.
- In the boxes: N, the days; x, the exceptions; p, one minus the confidence level; and L R, the chi square statistic.

13 Example 5: the Kupiec test on Ashby

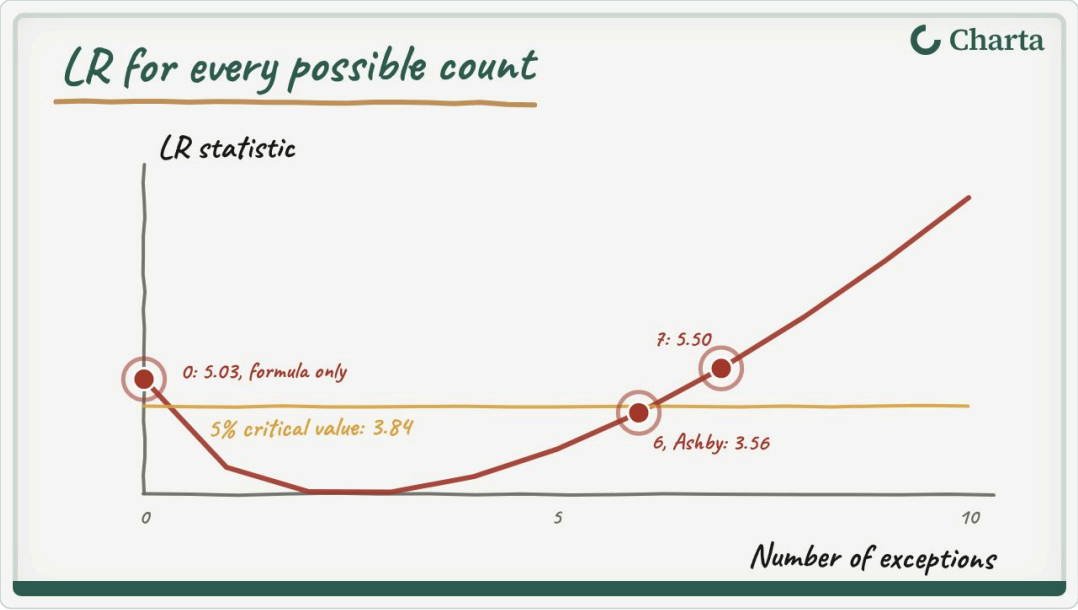
Ashby: N = 250 days, x = 6 exceptions, p = 0.01

Observed failure rate	x ÷ N = 6 ÷ 250	**0.024**
ln L at 0.024	244 ln 0.976 + 6 ln 0.024	**−28.3056**
ln L at 0.01	244 ln 0.99 + 6 ln 0.01	**−30.0833**
LR	2 × (30.0833 − 28.3056)	**3.56**
Verdict at 5%	3.56 < 3.84	**Do not reject**

IN THE LESSON

- On Ashby, N is two hundred and fifty, x is six and p is point zero one.
- The observed failure rate is six over two hundred and fifty, point zero two four.
- The log likelihood at that rate is two hundred and forty four times the natural log of point nine seven six, plus six times the natural log of point zero two four, which is minus twenty eight point three zero five six.
- At the model's rate it is two hundred and forty four times the natural log of point nine nine, plus six times the natural log of point zero one, which is minus thirty point zero eight three three.
- L R is twice the gap, two times one point seven seven seven seven, about three point five six.
- That is below three point eight four, so at the five per cent level we do not reject Ashby's model.

14 LR for every possible count



IN THE LESSON

- This chart shows L R for every count, with two hundred and fifty days at ninety nine per cent.
- Along the bottom is the number of exceptions, from zero to ten.
- Up the side is L R.
- The horizontal line is the five per cent critical value, three point eight four.
- The curve is shaped like a U, lowest near the expected count of two point five.
- At zero exceptions, L R is five point zero three, just above the line, so the formula alone would reject a spotless record.
- But with only two hundred and fifty days, the usual non rejection table at ninety nine per cent, Jorion's, accepts anything below seven and has no lower bound.
- At six, Ashby's count, L R is three point five six, just under the line.
- At seven it jumps to five point five zero.
- So seven or more exceptions are rejected at five per cent.

15 A large-sample shortcut

z = (x − pN) ÷ √(p(1 − p)N)
Reject if abs(z) > 1.96 (5% level, two-sided)

Exceptions, x	Exceptions observed
Tail probability, p	Tail probability under the null
Days, N	Days in the sample

IN THE LESSON

- There is a quicker test for large samples.
- Under the null hypothesis the count of exceptions is binomial, with mean p times N and variance p times one minus p times N.
- The z statistic is x minus p N, divided by the square root of p times one minus p times N.
- Notice that p, the null probability, is used in both the mean and the variance, not the observed rate.
- Reject when the absolute value of z exceeds one point nine six, the two sided five per cent value of the normal distribution.
- This is only an approximation, and with two and a half expected exceptions a rough one.
- The boxes: x, the exceptions; p, the null tail probability; N, the days.

16 Example 6: the shortcut on Ashby

Same data: N = 250, x = 6, p = 0.01

Expected count	pN = 0.01 × 250	**2.5**
Standard deviation	√(0.01 × 0.99 × 250) = √2.475	**1.573**
z	(6 − 2.5) ÷ 1.573	**2.22**
Shortcut verdict	2.22 > 1.96	**Reject**
LR verdict (Example 5)	3.56 < 3.84	**Do not reject**

0 exceptions: LR 5.03 > 3.84 by formula, yet the 250-day table accepts it

IN THE LESSON

- Now the shortcut on the same data.
- The expected count is p N, point zero one times two hundred and fifty, two point five.
- The standard deviation is the square root of point zero one times point nine nine times two hundred and fifty.
- That is the square root of two point four seven five, about one point five seven three.
- So z is six minus two point five, which is three point five, divided by one point five seven three, giving two point two two.
- That is above one point nine six, so the shortcut rejects Ashby's model.
- But the likelihood ratio test, in example five, did not reject it.
- The two tests disagree at six exceptions.
- The z score is the approximation; the likelihood ratio is the test itself.
- And the lower tail: zero exceptions gives an L R of five point zero three, a reject by formula, yet the two hundred and fifty day table accepts it.

17 Two ways to be wrong

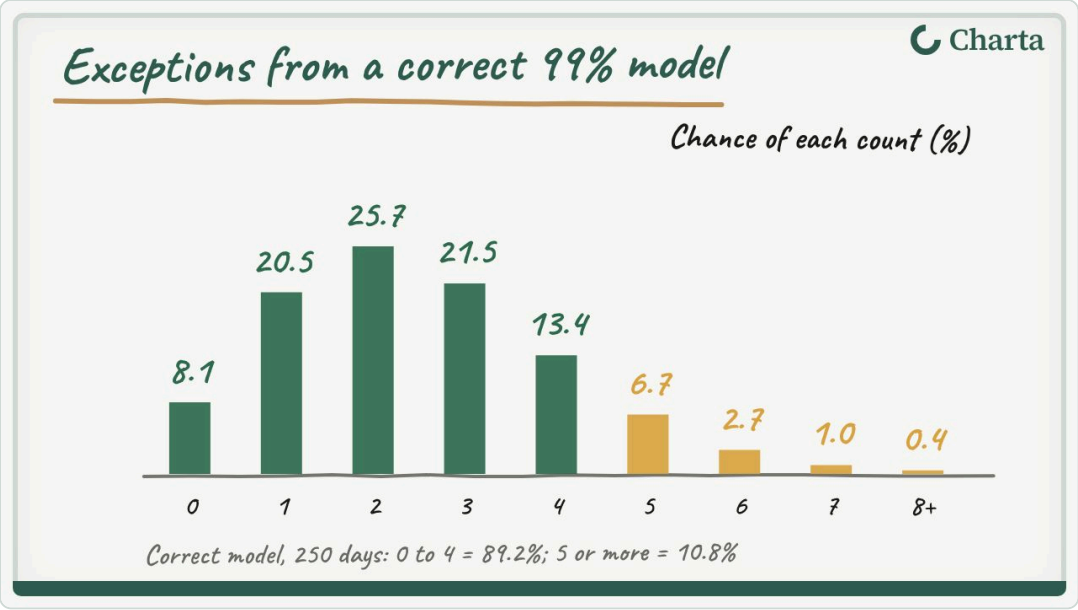
Backtesting is a statistical decision, so it can go wrong in two ways

Decision	Model is correct	Model is wrong
Accept the model	Right decision	Type II error
Reject the model	Type I error	Right decision

IN THE LESSON

- A backtest ends in a decision, accept or reject the model, and like any statistical decision it can go wrong in two ways.
- The columns are the decision, then a model that is correct, then a model that is wrong.
- The first row is accepting the model.
- If the model is correct, accepting it is the right decision.
- If the model is wrong, accepting it is a Type Two error, such as keeping a model that understates risk.
- The second row is rejecting the model.
- If the model is correct, rejecting it is a Type One error, throwing out a good model through bad luck.
- If the model is wrong, rejecting it is the right decision.
- The chance of a Type One error is the significance level, and one minus the chance of a Type Two error is the power of the test.

18 Exceptions from a correct 99% model



IN THE LESSON

- How large are these errors?
- With a true probability of one per cent each day, independent days, and two hundred and fifty days, the count is binomial.
- The bars give the percentages.
- Zero exceptions, eight point one.
- One, twenty point five.
- Two, twenty five point seven.
- Three, twenty one point five.
- Four, thirteen point four.
- Those five bars are green.
- Then the amber bars: five exceptions, six point seven.
- Six, two point seven.
- Seven, one point zero.
- Eight or more, point four.
- The green bars, zero to four, add to eighty nine point two per cent.
- So even a perfect model shows five or more exceptions in ten point eight per cent of years, through pure bad luck.

19 Example 7: the Type I error of a cut-off

A correct 99% model, 250 days. Reject when the count reaches a cut-off

Cut-off 5: P(x ≥ 5)	1 − 0.892	**10.8%**
Cut-off 10: P(x ≥ 10)	binomial sum from 10 up	**0.03%**
What each figure means	good model rejected by luck	**Type I**

A higher cut-off: fewer Type I errors

IN THE LESSON

- Now suppose Ashby's supervisor rejects the model whenever the count reaches a cut off.
- With a cut off of five, the Type One error is the chance that a correct model shows five or more: one minus point eight nine two, which is ten point eight per cent.
- With a cut off of ten, we add the binomial probabilities from ten upwards, and the Type One error is only about point zero three per cent.
- Each figure is the chance that a good model is rejected purely by luck, which is exactly what a Type One error is.
- So raising the cut off makes Type One errors much rarer.

20 Example 8: the Type II error, and the trade-off

Suppose the true probability is 3%, not 1%: the model understates risk

Expected exceptions	3% × 250	**7.5**
Cut-off 10: P(x ≤ 9)	accepted although wrong	**77.9%**
Power at cut-off 10	1 − 77.9%	**22.1%**
Cut-off 5: P(x ≤ 4)	accepted although wrong	**12.8%**
Power at cut-off 5	1 − 12.8%	**87.2%**

Tighter cut-off: fewer Type II errors, more Type I

IN THE LESSON

- Now suppose Ashby's desk truly has a three per cent chance of an exception each day, while the model claims one.
- Expected exceptions are now three per cent times two hundred and fifty, seven point five.
- With a cut off of ten, the model is accepted whenever the count is nine or fewer, which happens seventy seven point nine per cent of the time.
- That is the Type Two error, and the power is only twenty two point one per cent.
- With a cut off of five, the model is accepted only at four or fewer, twelve point eight per cent of the time, so the power rises to eighty seven point two per cent.
- The trade off: the tighter cut off catches the bad model far more often, but rejects a good model ten point eight per cent of the time.

21 Which error costs more?

- Type I: correct model rejected; chance = significance level
- Type II: wrong model accepted; 1 − its chance = power
- Type II depends on how wrong the model is, and on N
- For a given sample, the two errors trade off
- Regulator: keeping a risk-understating model costs more

IN THE LESSON

- A Type One error rejects a correct model, and its probability is the significance level, or, for a fixed cut off, the chance that a correct model reaches it by luck.
- A Type Two error accepts a wrong model, and one minus its probability is the power.
- The Type Two error depends on how wrong the model is, because a badly wrong model is easier to catch, and on the sample size, because more days give more evidence.
- For a given sample the two errors trade off.
- For a regulator, wrongly keeping a model that understates risk is usually the costlier mistake, since the bank then holds too little capital.
- Wrongly rejecting a good model wastes effort and capital.

22 Right count, wrong timing

- Unconditional coverage: the frequency of exceptions only
- Conditional coverage: frequency and independence
- Independence: today's exception says nothing of tomorrow
- Clusters: the model reacts too slowly to volatility
- Right count but clustered: still a poor model

IN THE LESSON

- The Kupiec test looks only at how many exceptions there were, which is called unconditional coverage.
- But a model can have the right number of exceptions and still be poor, if they arrive in clusters.
- Conditional coverage adds a second requirement, independence.
- The chance of an exception tomorrow should not depend on whether one occurred today.
- Clustering usually means the model reacts too slowly to changing volatility: when markets turn stormy, the V a R stays low for days while losses pile up.
- A model with the right count but clustered exceptions is still a poor model.

23 Christoffersen's test

$\pi_{01} = T_{01} \div (T_{00} + T_{01})$ $\pi_{11} = T_{11} \div (T_{10} + T_{11})$
LRind = 2 × [ln L(π_{01}, π_{11}) − ln L(π)]
LRcc = LRuc + LRind (LRuc is the Kupiec LR)
Good model: $\pi_{01} = \pi_{11} = p$. Reject LRcc above 5.99 (5%)

π_{01}	Exception after a quiet day
π_{11}	Exception after an exception
LRind	Independence: chi-square, 1 df
LRcc	Joint: chi-square, 2 df

IN THE LESSON

- Christoffersen's approach counts transitions from one day to the next.
- T zero one is the number of quiet days followed by an exception, and T one one the number of exceptions followed by another.
- Pi zero one, the chance of an exception after a non exception, is T zero one over T zero zero plus T zero one.
- Pi one one, the chance of an exception after an exception, is T one one over T one zero plus T one one.
- L R ind is twice the gap between the log likelihood with these two probabilities and the one with a single shared rate, pi.
- L R c c is the sum of L R u c, which is simply the Kupiec statistic, and L R ind.
- For a good model both pis equal p, and we reject when L R c c exceeds five point nine nine.
- In the boxes, L R ind is chi square with one degree of freedom, and L R c c with two.

24 Example 9: counting Ashby's transitions

Ashby's 6 exceptions came as two runs of 3 days; 249 transitions

Exception then exception, T_{11}	2 per run × 2 runs	**4**
Quiet then exception, T_{01}	start of each run	**2**
Exception then quiet, T_{10}	end of each run	**2**
Quiet then quiet, T_{00}	249 − 4 − 2 − 2	**241**
π_{01} and π_{11}	2 ÷ 243 and 4 ÷ 6	**0.0082, 0.667**

IN THE LESSON

- Ashby's six exceptions were not scattered.
- They came in two runs of three days in a row, with quiet days at the start and end of the year.
- Two hundred and fifty days give two hundred and forty nine transitions.
- Within each run of three there are two moves from an exception to another exception, so T one one is two times two, four.
- Each run starts with a quiet day followed by an exception, so T zero one is two.
- Each run ends with an exception followed by a quiet day, so T one zero is also two.
- Quiet followed by quiet is two hundred and forty nine minus four minus two minus two, two hundred and forty one.
- So pi zero one is two over two hundred and forty three, about point zero zero eight two, and pi one one is four over six, about point six six seven.

25 Example 10: Ashby fails conditional coverage

Shared rate $\pi = 6 \div 249 = 0.0241$. $LR_{uc} = 3.56$ from Example 5

ln L, one shared rate	$243 \ln(1 - \pi) + 6 \ln \pi$	**−28.2813**
ln L, two separate rates	T_{00}, T_{01} at π_{01}; T_{10}, T_{11} at π_{11}	**−15.4107**
LRind	$2 \times (28.2813 - 15.4107)$	**25.74**
LRcc	$3.56 + 25.74$	**29.30**
Verdict	$29.30 > 5.99$	**Reject**

IN THE LESSON

- The shared exception rate, pi, is six over two hundred and forty nine, point zero two four one.
- The log likelihood with one shared rate is two hundred and forty three times the natural log of one minus pi, plus six times the natural log of pi, which is minus twenty eight point two eight one three.
- With separate rates it is two hundred and forty one times the log of one minus pi zero one, plus two times the log of pi zero one, plus two times the log of one third, plus four times the log of two thirds: minus fifteen point four one zero seven.
- L R ind is twice the gap, twenty five point seven four.
- Adding L R u c from example five, L R c c is three point five six plus twenty five point seven four, twenty nine point three zero.
- That is far above five point nine nine: Ashby passes Kupiec but fails Christoffersen.

26 Example 11: each part passes, the sum fails

Ashby's rates desk: $LR_{uc} = 3.2$ and $LR_{ind} = 3.1$

Unconditional test	$3.2 < 3.84$	**Pass**
Independence test	$3.1 < 3.84$	**Pass**
Joint statistic, LRcc	$3.2 + 3.1$	**6.3**
Conditional coverage test	$6.3 > 5.99$	**Reject**

Each part passes, yet the joint test fails

IN THE LESSON

- Ashby's rates desk has a different record.
- Its unconditional statistic is three point two, and its independence statistic is three point one.
- Taken alone, the unconditional test passes, since three point two is below three point eight four.
- The independence test also passes, since three point one is below three point eight four.
- But the joint statistic is their sum, six point three, and that is above the two degree of freedom critical value of five point nine nine.
- So the conditional coverage test rejects the model.
- The exam point: two components that each pass can still fail in sum.

27 Basel: what is backtested

- Legacy Basel internal-models framework for VaR
- Backtest the 1-day, 99% one-tailed VaR
- Against daily trading outcomes, last 250 business days
- Capital VaR is 10-day: count exceptions on the 1-day VaR
- Legacy regime: FRTB now uses expected shortfall

IN THE LESSON

- Under the legacy Basel internal models framework for V a R, a bank backtests its one day, ninety nine per cent, one tailed V a R.
- It compares it with daily trading outcomes over the most recent two hundred and fifty business days.
- Notice the horizon.
- The V a R used for capital has a ten day horizon, but the backtest does not use it.
- Exceptions are counted on the one day V a R, never on the ten day figure, which is much larger and would hide almost every exception.
- Keep in mind that this is the legacy regime.
- The current framework, the F R T B internal models approach, uses expected shortfall for capital, so treat what follows as the V a R era rules.

28 The Basel traffic light

Exceptions in the last 250 days set the zone and the multiplier k

Zone	Exceptions	Multiplier k
Green	0 to 4	3.00
Yellow	5, 6, 7	3.40, 3.50, 3.65
Yellow	8, 9	3.75, 3.85
Red	10 or more	4.00, automatic

IN THE LESSON

- The count of exceptions puts the bank in a traffic light zone, which sets a multiplier we call k.
- The columns are the zone, the number of exceptions, and the multiplier.
- Green is zero to four exceptions, and the multiplier stays at its minimum of three point zero zero.
- Yellow covers five to nine.
- Five, six and seven exceptions give three point four zero, three point five zero and three point six five.
- Eight and nine give three point seven five and three point eight five.
- Red is ten or more, with a multiplier of four point zero zero.
- In the red zone the increase is automatic and the model is presumed flawed.
- Ashby, with six exceptions, sits in the yellow zone with a multiplier of three point five zero.

29 The market risk charge

Charge = max(previous day's VaR , k × 60-day average VaR)
k ≥ 3, set by the count of exceptions, not their size

Previous day's VaR	10-day 99% VaR, yesterday
60-day average VaR	10-day 99% VaR, last 60 days
Multiplier, k	At least 3; set by the backtest

IN THE LESSON

- Here is where the multiplier bites.
- The market risk charge is the larger of two amounts.
- The first is the previous day's V a R.
- The second is the multiplier k times the average V a R of the previous sixty business days.
- Both use the ten day, ninety nine per cent V a R.
- Because k is at least three, the second amount is almost always the larger, so the multiplier drives the charge.
- Note what sets k: the count of exceptions in the one day backtest, not the size of the losses on those days.
- A bank with five small exceptions pays a higher multiplier than a bank with four enormous ones.

30 Example 12: what six exceptions cost Ashby

10-day 99% VaR: previous day $34m, 60-day average $31m. 6 exceptions

If green: k = 3.00	3.00 × 31	**$93.0m**
Ashby, yellow: k = 3.50	3.50 × 31	**$108.5m**
Charge	max(34, 108.5)	**$108.5m**
Cost of 6 exceptions	108.5 − 93.0	**$15.5m**
If red: k = 4.00	4.00 × 31	**$124.0m**

IN THE LESSON

- Now price Ashby's six exceptions.
- Its ten day, ninety nine per cent V a R was thirty four million dollars on the previous day, and the average over the last sixty days was thirty one million.
- Had the backtest been green, k would be three point zero zero, and the multiplied average would be three times thirty one, ninety three million.
- In fact Ashby is yellow, with k of three point five zero, so the multiplied average is one hundred and eight point five million.
- The charge is the larger of thirty four and one hundred and eight point five, so one hundred and eight point five million.
- The six exceptions cost Ashby fifteen point five million dollars of extra capital.
- Had it fallen into the red zone, the charge would have been four times thirty one, one hundred and twenty four million.

31 Mistakes to avoid

- Exception: loss larger than that day's VaR, not equal
- Zero exceptions: 250-day table accepts; maybe too cautious
- LR rejects at 7 or more (5%); the z shortcut at 6
- The right count can still fail conditional coverage
- Basel: count on 1-day VaR; k follows count, not size

IN THE LESSON

- Mistakes to avoid.
- An exception is a loss larger than that day's V a R, so a loss exactly equal to it does not count.
- Zero exceptions gives an L R above three point eight four, but the two hundred and fifty day table accepts it, and regulators do not penalise too few.
- Still, a long run with none hints that the model may be too cautious.
- With two hundred and fifty days at ninety nine per cent, the likelihood ratio needs seven or more exceptions to reject at five per cent, while the z shortcut already rejects at six.
- A model with the right count can still fail conditional coverage if its exceptions cluster.
- And under Basel, count exceptions on the one day V a R, and remember that the multiplier depends on the count, not on how large the losses were.

32 Try three quick questions

1. 97.5% one-day VaR, backtested over 400 days. Expected exceptions?
   Answer: p = 2.5%. 2.5% × 400 = 10.
2. 99% VaR, 8 exceptions in 250 days. Basel zone and multiplier?
   Answer: Yellow (5 to 9): k = 3.75.
3. LRuc = 2.0 and LRind = 4.5. Pass conditional coverage at 5%?
   Answer: LRcc = 6.5 > 5.99: it fails.

IN THE LESSON

- Now try three quick questions.
- Pause, think, then listen.
- Question one.
- A ninety seven point five per cent one day V a R is backtested over four hundred days.
- How many exceptions should a correct model show?
- The tail probability is two point five per cent, and two point five per cent of four hundred is ten.
- Question two.
- A bank's ninety nine per cent V a R had eight exceptions in the last two hundred and fifty days.
- Which Basel zone, and what multiplier?
- Eight is in the yellow zone, five to nine, so the multiplier is three point seven five.
- Question three.
- A backtest gives an L R u c of two point zero and an L R ind of four point five.
- Does the model pass conditional coverage at five per cent?
- L R c c is six point five, above five point nine nine, so it fails.