CHARTA LEARNING · LESSON NOTES

Credit Scoring and Rating

FRM Part II

1 Two borrowers, two tools

- Thousands of card applicants: a model gives each a number
- Fennick Foods, one big loan: analysts give it a grade
- Both rank borrowers by credit quality
- Score = model number. Rating = grade on a defined scale

IN THE LESSON

- Picture Monday morning at Larkspur Bank.
- Five thousand people have applied for a credit card overnight, and there is also one request from Fennick Foods, a listed food manufacturer asking for a large loan.
- Nobody can read five thousand files, so a statistical model reads them and returns a number for each applicant.
- Fennick is different.
- A team of analysts studies its accounts, its industry and its management, and then gives it a grade on a letter scale.
- Both tools do the same basic job: they rank borrowers by credit quality, so that the bank can lend more to the safest and less to the riskiest.
- But the number and the grade differ in form and in use, and that difference is our first topic.

2 Scoring versus rating

Both rank credit quality, but they differ in form and use

	Score	Rating
Form	Number	Ordinal grade
Built by	Statistical model	Analysts, judgement
Suits	Consumers, small firms	Corporates, sovereigns
Examples	Logistic regression	S&P, Moody's, Fitch
Is it a PD?	Not until calibrated	Not until mapped

IN THE LESSON

- Here are the two side by side.
- A credit score is a number produced by a statistical model from borrower characteristics, such as income, past late payments or how much of a credit limit is used.
- It suits large homogeneous populations, meaning many borrowers who look alike, such as consumers and small businesses.
- A credit rating is an ordinal grade on a defined scale.
- Ordinal means the grades are in order, best to worst, but the gaps between them are not fixed amounts.
- Ratings are given to corporates, banks, sovereigns and individual debt issues, usually with analyst judgement.
- The last row is the key warning.
- Neither is a probability of default by itself.
- A score is only a ranking until it is mapped and calibrated to a P D, and a rating needs mapping to a P D in the same way.

3 The four varieties of scoring

- Application: accept or refuse a new applicant
- Behavioural: manage an existing account
- Collection: whom to chase first, and how hard
- Fraud: flag suspicious applications or use
- Models: discriminant analysis, logistic regression

IN THE LESSON

- Scoring comes in four varieties, named by the decision they support.
- An application score helps decide whether to accept a new applicant, and it is what runs on those five thousand card forms.
- A behavioural score is used on existing customers, using how they actually use and repay the account, to set limits and watch for trouble.
- A collection score is used once an account is late, to decide whom to chase first and how hard.
- A fraud score flags applications or transactions that look suspicious.
- Behind them sit statistical models.
- Discriminant analysis finds the weighted combination of characteristics that best separates good borrowers from bad ones.
- Logistic regression estimates the probability of default directly, as a number between zero and one.
- We will build one later.

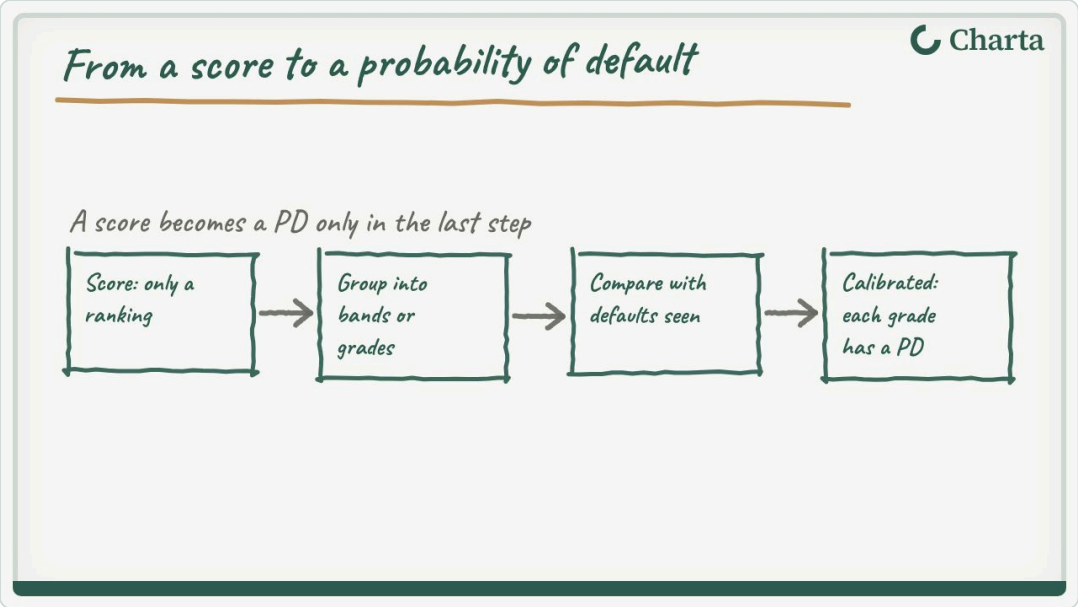
4 The varieties of rating

- External: S&P, Moody's, Fitch rate issuers and debt
- Internal: the bank's own grades for its borrowers
- Internal systems: statistical, expert-based or hybrid
- Statistical: model decides. Expert: officers decide
- Hybrid: model proposes, analyst may adjust

IN THE LESSON

- Ratings divide first by who produces them.
- External ratings come from agencies such as S and P, Moody's and Fitch, who rate corporates, banks, sovereigns and individual debt issues, and publish the result for everyone.
- Internal ratings are the bank's own grades for its own borrowers, used for lending limits, pricing and capital.
- Internal rating systems themselves come in three types.
- A statistical system lets a model decide the grade.
- An expert-based system lets credit officers decide, using a checklist and their judgement.
- A hybrid system combines the two: the model proposes a grade and an analyst may adjust it, for reasons the model cannot see, such as a change of management.
- Larkspur's grade for Fennick Foods would be an internal rating, and the agency grade would be external.

5 From a score to a probability of default



IN THE LESSON

- Now the idea that trips up many candidates.
- Suppose Larkspur's model gives one applicant a score of seven hundred and another six hundred and fifty.
- We know the first is probably safer, but not how safe.
- A score is only a ranking.
- To turn it into a probability of default we group scores into bands, then compare each band with the defaults that were actually seen in the past, and so attach a default rate to it.
- That step is calibration.
- Ratings work the same way: the grade is a label, and a P D has to be attached to each grade by looking at the default history of borrowers who held it.
- Ranking tells you who is riskier.
- Calibration tells you by how much.

6 The Altman Z score

Z = 1.2×WC/TA + 1.4×RE/TA + 3.3×EBIT/TA + 0.6×MVE/TL + 1.0×Sales/TA

Working capital / total assets	Short-term liquidity cushion
Retained earnings / total assets	Profits kept over the years
EBIT / total assets	Operating profit over assets
Market equity / total liabilities	Equity cushion at market prices
Sales / total assets	Asset turnover

IN THE LESSON

- The best known score for companies is the Altman Z score, which uses discriminant analysis.
- It combines five ratios.
- W C is working capital, current assets less current liabilities, and T A is total assets.
- R E is retained earnings, the profits kept in the business over its life.
- E B I T is earnings before interest and tax.
- M V E is the market value of equity, and T L is total liabilities.
- Each ratio gets a weight: one point two, one point four, three point three, point six and one point zero.
- Add them up.
- A higher score is safer.
- Below one point eight one the firm is in the distress zone, and above two point nine nine it is in the safe zone.
- Between those two cut-offs, the result is uncertain.

7 Example 1: Z score for Fennick Foods

Fennick: WC/TA 0.10, RE/TA 0.20, EBIT/TA 0.08, MVE/TL 0.90, Sales/TA 1.10

WC and RE terms	1.2×0.10 + 1.4×0.20	**0.40**
Earnings term	3.3 × 0.08	**0.264**
Market value term	0.6 × 0.90	**0.54**
Sales term	1.0 × 1.10	**1.10**
Z score	0.40+0.264+0.54+1.10	**2.304: middle**

IN THE LESSON

- Let us score Fennick Foods.
- Suppose working capital is ten per cent of assets, retained earnings twenty per cent, earnings before interest and tax eight per cent, the market value of equity is point nine of total liabilities, and sales are one point one times assets.
- Multiply each by its weight.
- One point two times point one is point twelve, and one point four times point two is point twenty eight, together point four.
- Three point three times point oh eight is point two six four.
- Point six times point nine is point five four.
- And one times one point one is one point one.
- The total is two point three oh four.
- That is above one point eight one and below two point nine nine, so Fennick is in the uncertain middle zone, not clearly safe.

8 Scoring and rating: traps to avoid

- A score ranks; it is not a PD until calibrated
- A rating is ordinal: steps are not equal PD gaps
- Higher Z is safer, the opposite of a higher PD
- Below 1.81 distress, above 2.99 safe
- Cut-offs belong to the original manufacturing model

IN THE LESSON

- Four traps appear again and again in this part of the syllabus.
- First, a score is not a probability.
- It ranks borrowers, and only calibration to observed default rates turns it into a P D.
- Second, a rating is ordinal.
- It tells you the order of grades, not that each step up is the same drop in default risk.
- Third, with the Altman score, direction matters: a higher Z is safer, the opposite of a P D, where higher is riskier.
- Fourth, the cut-offs of one point eight one and two point nine nine belong to the original public manufacturing model.
- They are not universal limits for every industry, and an exam question will normally give them to you.

9 A story: the weather and the climate

- TTC: judge the borrower across a full economic cycle
- PIT: default chance over the next year, as now
- In a recession PIT grades fall; TTC grades barely move
- Same borrower, two questions, two behaviours

IN THE LESSON

- Two risk teams at Larkspur grade the same thousand companies.
- Team Tide asks: how will each firm fare across a whole economic cycle, including a nasty trough?
- It assumes stressed trough conditions, so its grades hardly change when the news changes.
- Team Spot asks a different question: given today's conditions, how likely is this firm to default in the next twelve months?
- When a recession arrives, Spot marks many firms down, because the answer to its question really has changed.
- Tide barely reacts, because it was already looking at the bad times.
- It is like the difference between climate and weather.
- Tide is a through-the-cycle system, T T C.
- Spot is a point-in-time system, P I T.

10 Through the cycle versus point in time

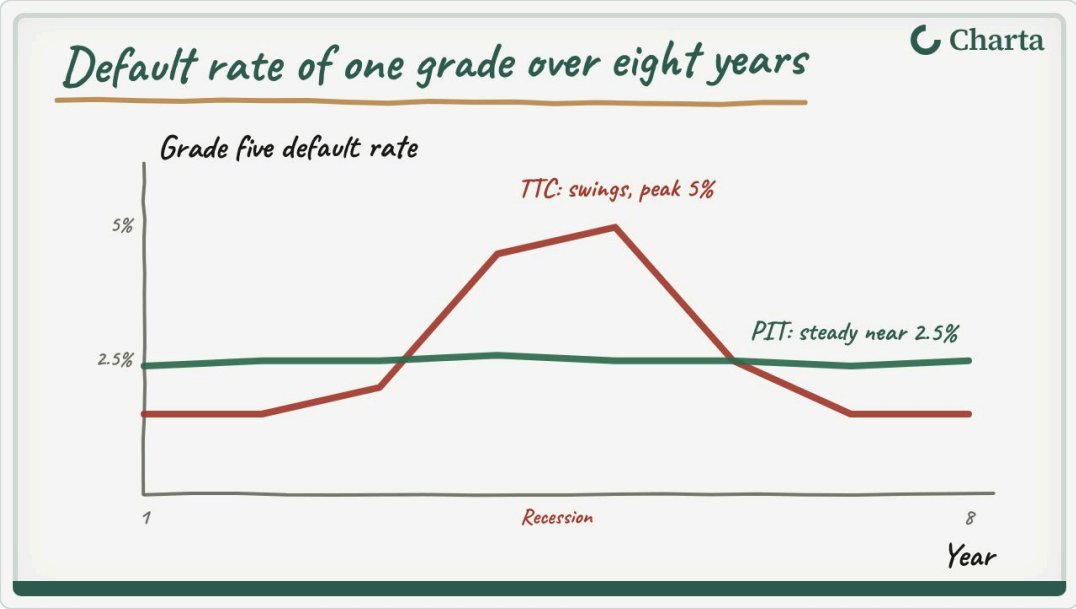
Stable grades with swinging default rates, or moving grades with steady default rates

	TTC	PIT
Horizon	Full cycle, trough	Next year, today
Inputs	Lasting strength	Prices, macro data
Grades	Stable, few moves	Migrate with cycle
Grade default rate	Swings with cycle	Roughly stable
Example	Agency ratings	Equity-implied PD

IN THE LESSON

- Here is the comparison.
- A through-the-cycle rating looks at the borrower's lasting credit quality, typically assuming stressed trough conditions.
- The grade moves only when that lasting quality changes, so ratings are stable and migrate little.
- The price is that the realised default rate of each grade swings with the cycle, higher in recessions.
- A point-in-time rating uses inputs such as market prices and macro data to estimate default probability over the next year.
- In a downturn borrowers migrate to worse grades, while the P D attached to each grade stays roughly stable.
- Agency ratings are mostly through the cycle.
- Market-based measures, such as a P D implied by an equity price, are point in time.

11 Default rate of one grade over eight years



IN THE LESSON

- This chart follows one grade, say grade five, over eight years, with a recession in years four and five.
- The red line is the realised default rate of that grade under a through-the-cycle system.
- Because the grade does not react to the downturn, firms that default in the recession are still sitting in the grade, so its default rate climbs to five per cent at the worst point, then falls back to one and a half per cent.
- The green line is the same grade under a point-in-time system.
- Weak firms were moved out into worse grades as conditions worsened, so what remains in grade five defaults at about two and a half per cent in every year.
- Stable grades or stable default rates, but not both.

12 Example 2: procyclical provisions

$400m loans to grade five, LGD 45%. Amounts in $ millions

TTC PD: 8-year average	20.0% ÷ 8 years	**2.5%**
Loss at TTC PD, every year	2.5% × 45% × 400	**4.5**
PIT port. PD 1.5%: good year	1.5% × 45% × 400	**2.7**
PIT port. PD 5.0%: recession	5.0% × 45% × 400	**9.0**
Recession vs good year	9.0 ÷ 2.7	**3.3 times**

IN THE LESSON

- Now the consequence for money.
- Larkspur lends four hundred million dollars to firms in grade five, with loss given default of forty five per cent.
- Over the eight years just shown, the through-the-cycle default rates average two point five per cent.
- That is the T T C P D, the long-run average, and it gives an expected loss of four point five million dollars, in every year.
- Under the point-in-time system, the P D follows conditions.
- In a good year at one and a half per cent, expected loss is two point seven million.
- In the recession at five per cent, it is nine million.
- Provisions and capital based on point-in-time numbers are more than three times larger in the recession than in the good year, exactly when the bank is weakest.

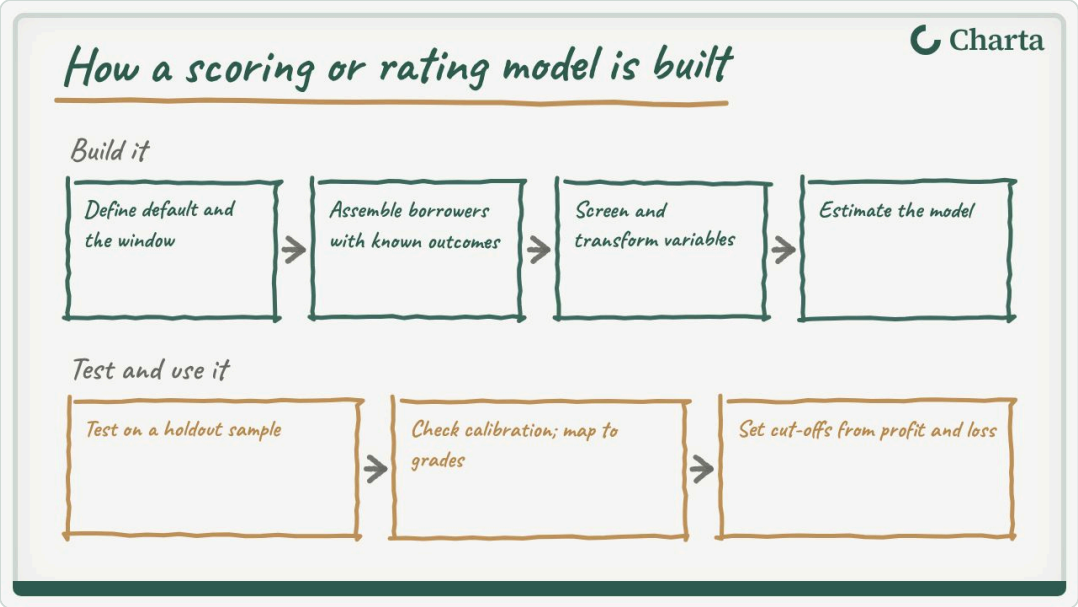
13 TTC and PIT: the exam points

- Agency ratings: mostly TTC. Equity-implied PDs: PIT
- TTC PD = long-run average, not this year's rate
- PIT: grades migrate, grade PDs roughly stable
- TTC: grades stable, grade default rates swing
- PIT makes capital and provisions more procyclical

IN THE LESSON

- Four points to memorise.
- First, agency ratings are mostly through the cycle, while market-based measures such as equity-implied default probabilities are point in time.
- Second, a T T C P D is the long-run average default rate, not the rate for the current year, so do not use it as a forecast for next year.
- Third, in a point-in-time system the grades migrate, borrowers moving down in a downturn, but the P D attached to each grade stays roughly stable.
- In a through-the-cycle system it is the other way round.
- Fourth, and the usual exam point: point-in-time systems make capital and provisions more procyclical, meaning they rise in recessions and fall in booms.

14 How a scoring or rating model is built



IN THE LESSON

- Now we open the workshop.
- Building a scoring or rating model follows a sequence, and the exam likes the order.
- First, define default, because default is the thing we want to predict, called the dependent variable, and fix the observation window.
- Second, assemble historical data on borrowers whose outcomes are known.
- Third, screen and transform the candidate variables.
- Fourth, estimate the model, most often a logistic regression.
- The second row happens once the model exists.
- Test how well it separates good from bad on a holdout sample.
- Check that its probabilities match observed default rates.
- Map scores to grades.
- And finally set cut-offs by comparing profit on good accounts with loss on bad ones.
- We take these in turn.

15 Steps one and two: default and data

- Default is the dependent variable: the thing we predict
- Fix one rule, e.g. 90 days past due, throughout
- Window: record traits, then watch for a year
- Assemble borrowers with known good/bad outcomes
- Few defaults means little for the model to learn from

IN THE LESSON

- Start with the definition of default, because everything else depends on it.
- Larkspur might say a borrower defaults if it is ninety days late on a payment.
- Whatever the rule, it must be fixed and applied consistently, since a vague definition makes the model learn a blurry target.
- Next comes the observation window.
- We look at borrowers at one date, record their characteristics, and then watch them for a period, typically a year, to see who defaulted.
- Then we assemble historical data: many borrowers with known outcomes, each marked good or bad.
- If there are very few defaults, the model has little to learn from.
- The raw data is also messy, with missing values and extreme figures, which brings us to the next step.

16 Example 3: binning and weight of evidence

900 good and 100 bad accounts, binned by credit-limit use

Under 30% used: 450 G, 20 B	ln(50% ÷ 20%)	**+0.92**
30% to 70% used: 360 G, 40 B	ln(40% ÷ 40%)	**0.00**
Over 70% used: 90 G, 40 B	ln(10% ÷ 40%)	**−1.39**

IN THE LESSON

- Candidate variables are screened and transformed.
- A common method is binning: split a variable into ranges and code each range by its weight of evidence.
- Weight of evidence is the natural log of the share of all good accounts in the bin divided by the share of all bad accounts in the bin.
- A positive value means the bin is safer than average, and a negative one riskier.
- Larkspur has nine hundred good and one hundred bad accounts.
- Under thirty per cent of the limit used holds half the goods and a fifth of the bads, so the log of two point five, which is plus point nine two.
- The middle bin holds forty per cent of each, so zero.
- Over seventy per cent holds ten per cent of the goods and forty of the bads, so minus one point three nine.

17 Logistic regression

ln(PD ÷ (1 − PD)) = z
PD = 1 ÷ (1 + e^(−z))

Probability of default	Always between 0 and 1
Odds of default	PD ÷ (1 − PD)
Score, z	Intercept + weights × variables

IN THE LESSON

- Now we estimate the model, and the standard choice is logistic regression.
- The idea is to model not the probability itself, which is stuck between zero and one, but the log-odds of default.
- Odds are the probability of default divided by the probability of no default.
- The log-odds, written z, are a linear function of the variables: an intercept plus a weight times each variable.
- Because z can be any number, positive or negative, we are free to add up weights without breaking any rule.
- Then we solve for the probability.
- Rearranging gives P D equals one over one plus e to the power of minus z.
- That formula turns the score z back into a probability, always between zero and one.
- A big negative z means a tiny P D.

18 Example 4: from score to probability

z = −6.0 + 0.9 × late payments + 2.0 × utilisation. Applicant: 2 late, 80% used

Score z	−6.0 + 0.9×2 + 2.0×0.8	**−2.6**
e raised to −z	e^2.6	**13.46**
Default probability	1 ÷ (1 + 13.46)	**6.9%**
Check: odds	0.069 ÷ 0.931	**0.0741**
Check: log-odds	ln(0.0741)	**−2.6**

IN THE LESSON

- Suppose Larkspur's fitted model is: z equals minus six, plus point nine times the number of late payments, plus two times the credit-limit utilisation.
- An applicant has two late payments and uses eighty per cent of the limit.
- Then z is minus six, plus one point eight, plus one point six, which is minus two point six.
- To get the probability, compute e to the power of plus two point six, which is about thirteen point four six.
- Then one over one plus thirteen point four six is about point oh six nine.
- The applicant has a default probability of about six point nine per cent.
- As a check, the odds are point oh six nine divided by point nine three one, about point oh seven four, and the log of that odds is minus two point six, our z.

19 Example 5: correcting for oversampling

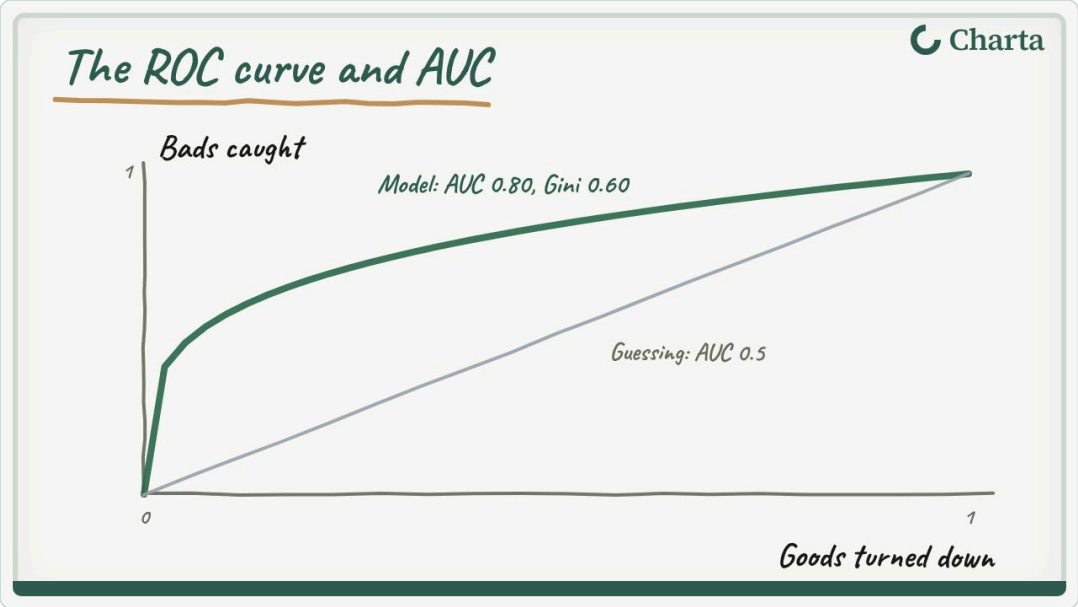
Sample 50% bad. True default rate 2%. Model says PD 60% on the sample

Model odds on the sample	0.60 ÷ 0.40	**1.5**
Population odds	0.02 ÷ 0.98	**0.0204**
Corrected odds	1.5 × 0.0204 ÷ 1	**0.0306**
Corrected PD	0.0306 ÷ 1.0306	**2.97%**

IN THE LESSON

- A practical wrinkle.
- Real defaults are rare, so to give the model enough bad accounts to learn from, builders often oversample defaults.
- Suppose Larkspur's sample holds five hundred bad and five hundred good accounts, a default share of fifty per cent, while the true default rate of the population is two per cent.
- The model, fitted on that sample, gives an applicant a probability of sixty per cent, which is odds of one point five.
- That is far too high for the real world.
- The fix is to correct the odds back to the true population default rate.
- Multiply them by the population odds, point oh two over point nine eight, which is point oh two oh four, divided by the sample odds, which are one.
- Corrected odds are point oh three oh six, a probability of two point nine seven per cent.

20 The ROC curve and AUC



IN THE LESSON

- Next, test discrimination, meaning how well the model separates bad borrowers from good ones.
- We do this on a holdout sample, accounts the model never saw.
- Rank them by score, then slide a cut-off down the list.
- At each cut-off, plot the share of bad accounts caught against the share of good accounts wrongly turned down.
- That is the R O C curve.
- A model with no skill gives the dashed diagonal, equal to guessing.
- A perfect model shoots straight up the left edge.
- The area under the curve is the A U C: point five for guessing, one for perfect.
- The green curve here has an A U C of point eight.
- The Gini coefficient rescales it: Gini equals two times A U C minus one, so point six, with zero meaning guessing.

21 Example 6: judge the model on a holdout

Development AUC 0.86. Holdout AUC 0.74. Gini = 2 × AUC − 1

Development Gini	2 × 0.86 − 1	**0.72**
Holdout Gini	2 × 0.74 − 1	**0.48**
Gini lost on new data	0.72 − 0.48	**0.24**
Honest figure to quote	the holdout	**Gini 0.48**

IN THE LESSON

- This brings us to the most common error in model building: judging a model on the sample that built it.
- A model can memorise the quirks of its own data and look excellent there, while being much weaker on new borrowers.
- Larkspur builds a model with an A U C of point eight six on its development sample.
- That gives a Gini of two times point eight six minus one, which is point seven two.
- On the holdout sample, the A U C falls to point seven four, a Gini of point four eight.
- A gap of point two four in Gini is a warning that the model has over-fitted.
- The honest measure of discrimination is the holdout figure, and it is the one a validator would quote.

22 Calibration and grades

Holdout results: observed default rate in each score band

Score band	Default rate seen
Grade 1: 700 and up	0.5%
Grade 2: 620 to 699	2%
Grade 3: 540 to 619	6%
Grade 4: below 540	15%

IN THE LESSON

- Discrimination is not enough.
- A model can rank perfectly and still give probabilities that are wrong, so we also check calibration against observed default rates.
- Take the holdout accounts in each score band, count how many actually defaulted, and compare with what the model predicted.
- If they match, the probabilities can be trusted.
- If they do not, we adjust.
- Then we map scores to grades.
- Larkspur groups scores into four bands and attaches the observed default rate to each, one half of one per cent for the best band, two per cent, six per cent, and fifteen per cent for the worst.
- Now each grade carries a calibrated P D, which is just what the first part of this lesson said a score needs.

23 Example 7: setting the cut-off

Good account earns $120. Bad account loses $1,500

Break-even default probability	120 ÷ (120 + 1,500)	**7.4%**
Grade 3 (PD 6%): expected profit	0.94×120 − 0.06×1,500	**+$22.80**
Grade 4 (PD 15%): expected profit	0.85×120 − 0.15×1,500	**−$123.00**
Decision	cut-off at 7.4%	**Accept 1 to 3**

IN THE LESSON

- The last step is the cut-off: the line between accepting and refusing.
- Larkspur makes the decision by comparing profit with loss.
- Suppose a good account earns one hundred and twenty dollars over its life, while a bad account loses one thousand five hundred dollars.
- At what default probability is the expected profit zero?
- One hundred and twenty divided by the sum of one hundred and twenty and one thousand five hundred is about seven point four per cent.
- Check the grades.
- Grade three, at six per cent, earns point nine four times one hundred and twenty minus point oh six times one thousand five hundred, a profit of twenty two dollars eighty, so accept.
- Grade four, at fifteen per cent, loses one hundred and twenty three dollars, so refuse.
- The cut-off falls between grades three and four.

24 Model building: mistakes to avoid

- Judging a model on the sample that built it
- Discrimination (ROC, Gini) is not calibration
- Forgetting to correct odds after oversampling defaults
- A vague default definition
- A cut-off set without profit versus loss

IN THE LESSON

- Here are the traps.
- The first and most common is judging a model on the sample that built it.
- Always test on a holdout, or on a later period, and quote that figure.
- The second is to confuse discrimination with calibration.
- The R O C curve and Gini tell you about ranking only.
- A model can rank well and still be badly calibrated, so check predicted against observed default rates.
- The third is to forget to correct the odds when defaults were oversampled, which leaves every probability far too high.
- The fourth is a loose definition of default.
- And the fifth is setting the cut-off from the model alone, instead of weighing the profit on good accounts against the loss on bad ones.

25 A story: one issuer, three bonds

- Fennick Foods has one issuer rating
- It has several bonds: secured, senior, subordinated
- Same borrower, but not the same recovery if it fails
- So each bond gets its own issue rating

IN THE LESSON

- Now the other side of ratings, the agencies.
- Larkspur holds bonds issued by Fennick Foods.
- Fennick has a rating as a company, an issuer rating, which is an opinion on its overall capacity to meet its financial obligations.
- But Fennick has several bonds.
- One is secured on its factories.
- One is ordinary senior debt.
- One is subordinated, meaning it is paid only after the others if Fennick fails.
- These bonds do not have the same chance of getting their money back, even though they have the same borrower.
- So agencies also publish an issue rating, one for each specific instrument.
- To see how the two relate, we compare them.

26 Issuer rating versus issue rating

Issuer: how likely is the obligor to pay? Issue: and what if it does not?

	Issuer	Issue
Applies to	The obligor	One instrument
Question	Capacity to pay	Pay, and recovery
Reflects	Overall strength	Seniority, collateral
Versus issuer	Is the base	Notched up or down

IN THE LESSON

- Here is the comparison.
- An issuer rating, such as the S and P issuer credit rating, expresses an obligor's overall capacity to meet its financial obligations.
- It is about the borrower.
- An issue rating applies to a specific instrument.
- It starts from the issuer rating, but it also reflects the instrument's seniority, any collateral, and the recovery that is expected if default happens.
- That is why bonds of the same borrower are notched.
- Subordinated debt, which recovers less, is notched below the issuer rating.
- Well-secured debt, which recovers more, may be notched above it.
- A notch is one step on the scale, for example from triple B to triple B minus.
- So the issue rating blends the chance of default with the loss if it happens.

27 Example 8: notching Fennick's bonds

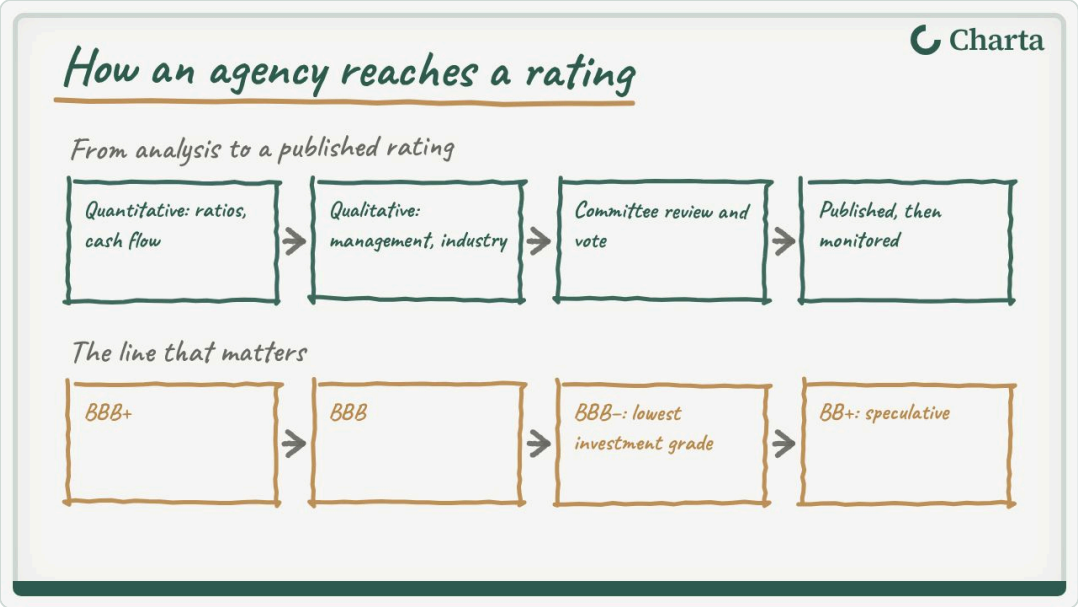
Issuer rating BBB. Notch counts are illustrative

Issuer rating	Fennick Foods	**BBB**
Well-secured bond	1 notch up	**BBB+**
Senior unsecured bond	no notching	**BBB**
Subordinated bond	2 notches down	**BB+**
Investment grade?	lowest is BBB−	**No: speculative**

IN THE LESSON

- Suppose Fennick's issuer rating is triple B.
- The notch counts here are for illustration only, as the agencies decide them case by case.
- Its well-secured bond is notched one step above, to triple B plus.
- The ordinary senior bond has no notching and stays at triple B.
- The subordinated bond is notched two steps below.
- Counting down the scale, triple B minus is one notch, and double B plus is two.
- Now the important line.
- Triple B minus on the S and P scale, or Baa three at Moody's, is the lowest investment grade.
- Double B plus is below it, so the subordinated bond is speculative grade, even though the company that issued it is investment grade.

28 How an agency reaches a rating



IN THE LESSON

- How does an agency arrive at a rating?
- It combines quantitative analysis with qualitative judgement and committee review.
- Quantitative analysis covers the numbers: leverage, cash flow, profitability and so on.
- Qualitative judgement covers what numbers miss: the quality of management, the strength of the industry and the legal terms of the instrument.
- Then a rating committee, not one analyst, votes on the grade, which is meant to keep the result consistent and to reduce individual bias.
- Once published, the rating is kept under review.
- The second row repeats the key landmark on the scale.
- Triple B minus or Baa three is the lowest investment grade, and anything below it is speculative, sometimes called junk.
- Many investors cannot hold anything below that line.

29 The chief complaints about ratings

Five chief complaints

Complaint	In short
Conflict of interest	Issuer pays; rating shopping
Slow and procyclical	Lag markets, then cliff edges
Overreliance	Regulators and mandates
Structured products	Weak in 2007 to 2009
Few agencies	Opaque, little competition

IN THE LESSON

- Agencies are widely criticised, and the complaints are a favourite exam list.
- The first is the conflict of interest in the issuer-pays model: the issuer who is being rated also pays for the rating, which can tempt an agency to be generous, and lets issuers go rating shopping, meaning choose the agency that gives the best result.
- The second is slowness and procyclicality.
- Ratings lag the market, then arrive as cliff-edge downgrades.
- The third is overreliance, by regulators and by investment mandates, which hard-wire ratings into rules.
- The fourth is weak performance on structured products in two thousand and seven to two thousand and nine.
- The fifth is limited transparency and little competition, since a few agencies dominate.

30 Cliff edges and why they matter

- Markets move first; the rating lags
- Then one big step: BBB to BB+, across the line
- Mandates and rules tied to ratings force sales
- Forced sales push prices down further
- Treat a rating as one input, not the whole answer

IN THE LESSON

- Let us tie several complaints together with Fennick.
- Suppose its senior debt has long been triple B, but its sales have been falling for a year.
- Markets have noticed: the bond price has already dropped and its spread has widened.
- The agency, slower and focused on the long run, moves only later, and when it does, the downgrade is a big step across the investment grade line, to double B plus.
- Now overreliance bites.
- A bond fund whose mandate allows only investment grade, and banks whose capital rules key off ratings, must sell or hold more capital at once.
- Prices fall further, in a cliff edge that comes after the market has moved.
- A rating is one useful input, not a substitute for your own analysis.

31 Recap: the main ideas

- Score = number; rating = grade; PD needs calibration
- Altman Z: below 1.81 distress, above 2.99 safe
- TTC: stable grades. PIT: migration, more procyclical
- Build: default, data, logit, holdout ROC/Gini, cut-off
- Issue ratings notched from issuer; five chief complaints

IN THE LESSON

- Let us pull it together.
- A score is a number from a statistical model, with application, behavioural, collection and fraud varieties, and a rating is an ordinal grade on a defined scale, external or internal, with internal systems statistical, expert-based or hybrid.
- Either one becomes a P D only after mapping and calibration.
- The Altman Z score is below one point eight one in distress and above two point nine nine when safe.
- Through-the-cycle ratings are stable with swinging grade default rates, point-in-time ratings migrate, and point-in-time makes capital more procyclical.
- Models follow a clear sequence and must be tested on a holdout.
- Issue ratings are notched from issuer ratings, and agencies face well-known complaints.

32 Try three quick questions

1. PIT system, recession starts. Grades, grade PDs, capital?
   Answer: Grades migrate down; grade PDs roughly stable; capital rises (procyclical).
2. Logistic model, score z = −3. Default probability?
   Answer: $1 \div (1 + e^{3}) = 1 \div 21.09 \approx 4.7\%$.
3. Issuer A−; subordinated bond notched 3 below. Rating?
   Answer: BBB+, BBB, BBB−: the lowest investment grade.

IN THE LESSON

- Now try three quick questions.
- Pause the video, think, then listen to the answers.
- Question one.
- A bank uses a point-in-time rating system, and a recession begins.
- What happens to its borrowers' grades, to the P D attached to each grade, and to its capital?
- Borrowers migrate to worse grades, the P D attached to each grade stays roughly stable, and capital rises, which is procyclical.
- Question two.
- A logistic model gives a score z of minus three.
- What is the default probability?
- One over one plus e to the power of three.
- E cubed is about twenty point one, so the answer is about four point seven per cent.
- Question three.
- An issuer is rated A minus, and its subordinated bond is notched three below.
- What is the bond's rating?
- Triple B minus, the lowest investment grade.