CHARTA LEARNING · LESSON NOTES

Fundamentals of Probability

FRM Part I

1 A story: one question, three answers

- Same loan, three different chances of default
- Ordinary: low. Recession: higher. Missed payments: more
- The loan did not change, the information did
- Probability tools tell us how to update, step by step

IN THE LESSON

- Imagine a bank asks, what is the chance that a loan defaults next year?
- Ask a loan officer on an ordinary day and she might say three per cent.
- Ask her again in the middle of a recession and the answer is far higher.
- Ask once more, after telling her that the borrower has already missed two payments, and it climbs again.
- The loan is the same each time.
- What changed is the information.
- Probability in risk management is really about this: how a chance should change when we learn something.
- The tools in this lesson are the careful way to do that.

2 The sample space and its outcomes

- Sample space = every outcome that could occur
- Bond rating in a year: upgrade, same, downgrade
- An outcome is one single result; exactly one occurs
- The list must be complete: leave nothing out

IN THE LESSON

- Every probability model starts by listing what could happen.
- The set of every possible outcome is called the sample space.
- Take the credit rating of a bond one year from now.
- The outcomes might be that it is upgraded, that it stays the same, or that it is downgraded.
- Those three are the whole sample space.
- An outcome is one single, specific result, and exactly one outcome will occur.
- The sample space must be complete, with nothing missing, and the outcomes must not overlap.
- If your list leaves out a possibility, such as default, then every probability you build on it will be wrong.
- So the first discipline is to write the list down.

3 An event is a set of outcomes

- Event = any subset of the sample space
- One outcome, several, none, or all of them
- Example: 'not upgraded' = {same, downgrade}
- Outcome = one point. Event = a set of points

IN THE LESSON

- An event is any collection of outcomes, in other words any subset of the sample space.
- It can be a single outcome, such as the bond stays the same.
- It can be several outcomes gathered together, such as the bond is downgraded, which in a richer model would include every rating lower than today's.
- It can even be the empty set, which contains no outcomes and can never occur, or the whole sample space, which is certain.
- The key distinction is this.
- An outcome is one point.
- An event is a set of points.
- When we say the event occurred, we mean that the outcome that actually happened is one of the members of the set.

4 The event space and its three rules

- Event space = the events that are given probabilities
- Rule 1: it contains the whole sample space
- Rule 2: it contains the complement of each event
- Rule 3: it contains any countable union of events
- Closed: 'not A' and 'A or B' always have a probability

IN THE LESSON

- The event space is the collection of events to which we assign probabilities.
- For the probabilities to make sense, this collection must follow three rules.
- First, it must contain the whole sample space.
- Second, for every event it contains, it must also contain the complement, meaning the event that the first one does not happen.
- Third, for any countable collection of events it contains, it must contain their union, meaning the event that at least one of them happens.
- These rules mean the collection is closed.
- Statements such as not A, or A or B, always have a probability, because they are themselves events inside the space.

5 Example 1: counting the events

Illustration: outcomes are upgrade, same, downgrade; every subset is an event

Outcomes in the sample space	up, same, down	**3**
Number of events	2^3 subsets	**8**
Same, with four ratings	2^4 subsets	**16**
Complement of 'upgrade'	{same, downgrade}	**one of the 8**

IN THE LESSON

- Let's count the events for the bond.
- The sample space has three outcomes: upgrade, same and downgrade.
- If every subset counts as an event, the number of events is two raised to the power of the number of outcomes.
- Two cubed is eight.
- These are the empty set, the three single outcomes, the three pairs, and the whole space.
- If we had four ratings, it would be two to the fourth, or sixteen.
- Now test closure.
- Take the event upgrade.
- Its complement is same or downgrade, and that pair is one of the eight, so rule two holds.
- The count grows very fast, but the rule is always the same: count the subsets, then check closure.

6 Mutually exclusive events

- Mutually exclusive: they cannot occur together
- Example: upgraded and downgraded in the same year
- Then P(A and B) = 0 and P(A or B) = P(A) + P(B)
- If events can overlap, subtract the overlap

IN THE LESSON

- Now to the relationships between events.
- Two events are mutually exclusive when they cannot happen together.
- A bond cannot be both upgraded and downgraded in the same year.
- The probability that both occur is therefore zero.
- That makes the arithmetic for either one occurring very simple: the probability of the union is just the sum of the two probabilities.
- Add them, with nothing to subtract.
- Exclusive events are the pieces of a clean split of the sample space.
- The danger is applying that simple addition to events that can overlap, such as a bond that is downgraded and a bond issued by a bank.
- They can both happen, so you must subtract the overlap.

7 Independent events

- Learning one tells you nothing about the other
- P(A | B) = P(A)
- Joint: P(A and B) = P(A) × P(B)
- Union: P(A or B) = P(A) + P(B) − P(A) × P(B)
- Independent events can still happen together

IN THE LESSON

- Independence is a different idea.
- Two events are independent when learning that one occurred does not change the probability of the other.
- Flip a fair coin twice.
- Knowing the first flip was heads tells you nothing about the second.
- In symbols, the probability of A given B is just the probability of A.
- The consequence is a product rule: the probability that both happen equals the probability of the first times the probability of the second.
- And the probability of either one is the sum minus the overlap, where the overlap is that product.
- Independence is about information, not about whether events can happen together.
- In fact, independent events with positive probabilities can always occur together.

8 The union and the joint, side by side

Exclusive: P(A or B) = P(A) + P(B)
Independent: P(A and B) = P(A)×P(B)
General: P(A or B) = P(A) + P(B) − joint

Union, either occurs	At least one of the two happens
Joint, both occur	Both happen together
Each event on its own	Its own chance, ignoring the other

IN THE LESSON

- Here are the two rules.
- For mutually exclusive events, the union equals the probability of the first plus the probability of the second.
- For independent events, the joint probability, meaning both occurring, equals the product of the two.
- Notice what each formula needs.
- The union is the chance that at least one event happens.
- The joint is the chance that both happen together.
- The general union rule, which always holds, is the sum of the two probabilities minus the joint.
- If the events are exclusive, the joint is zero and the subtraction disappears.
- If they are independent, the joint is the product.
- So the general rule contains both special cases.

9 Example 2: independent versus exclusive

Illustration: P(A) = 0.20, P(B) = 0.30

Independent: joint	0.20 × 0.30	**0.06**
Independent: union	0.20 + 0.30 − 0.06	**0.44**
Exclusive: joint	zero by definition	**0**
Exclusive: union	0.20 + 0.30	**0.50**
Exclusive: P(A | B)	A cannot occur if B has	**0, not 0.20**

IN THE LESSON

- Take two events with made up probabilities: A is twenty per cent and B is thirty per cent.
- First suppose they are independent.
- The joint is point two times point three, which is point zero six.
- The union is point two plus point three, minus point zero six, which is point four four.
- Now suppose instead that they are mutually exclusive.
- The joint is zero by definition, and the union is simply point two plus point three, point five.
- Same two events, different relationship, different answer.
- Finally, check dependence for the exclusive pair.
- If B has occurred, A cannot, so the probability of A given B is zero, not point two.
- Exclusive events with positive probabilities are therefore dependent.

10 Exclusive versus independent

Independent events can co-occur; exclusive ones cannot

	What it says
Exclusive: the idea	Cannot occur together
Independent: the idea	Learning one tells nothing
Exclusive: joint, union	0 and P(A) + P(B)
Independent: joint, union	P(A)×P(B), sum − joint
Positive chances: exclusive	Always dependent

IN THE LESSON

- Let's put the two ideas next to each other, because they are easy to mix up.
- Mutually exclusive is about the events being able to occur together, and the answer is no.
- Independent is about information, and the answer is that learning one does not change the other.
- The joint probability for exclusive events is zero, and for independent events it is the product.
- Exclusive events with positive chances are always dependent, because knowing one happened forces the other to zero.
- Independent events with positive chances can always happen together, so they are never exclusive.
- They are close to opposites.
- Before choosing a formula, check which condition the question actually states.

11 Conditional independence

- Given C, learning A adds nothing about B
- P(A and B | C) = P(A | C) × P(B | C)
- Plain independence and this one are separate properties
- Firms independent given the economy, dependent overall

IN THE LESSON

- Now a subtler idea.
- Two events are conditionally independent given a third event C when, once C is known, learning A tells you nothing more about B.
- In symbols, the probability of A and B given C equals the probability of A given C times the probability of B given C.
- This is separate from plain independence, and neither implies the other.
- Think of two firms.
- Given the state of the economy, their defaults might be independent, since each firm's fate then depends on its own business.
- But the state itself is shared.
- In a recession, both defaults become more likely.
- So unconditionally, the two defaults are positively dependent, even though they are independent within each state.

12 Example 3: two firms and one economy

Illustration: recession 20%. Default in recession 30%, in a normal year 5%

Each firm's default chance	0.2×0.30 + 0.8×0.05	**0.10**
Both default, recession	0.2 × 0.30 × 0.30	**0.018**
Both default, normal year	0.8 × 0.05 × 0.05	**0.002**
True joint probability	0.018 + 0.002	**0.020**
Naive: multiply marginals	0.10 × 0.10	**0.010**

IN THE LESSON

- Here is a made up economy.
- There is a twenty per cent chance of recession and an eighty per cent chance of a normal year.
- In a recession each firm defaults with probability thirty per cent.
- In a normal year, five per cent.
- Given the state, the firms are independent.
- Each firm's unconditional default chance is point two times point three plus point eight times point zero five, which is point one zero.
- Now the joint.
- In a recession, both default: point two times point three times point three, equals point zero one eight.
- In a normal year: point eight times point zero five times point zero five, which is point zero zero two.
- The total joint is point zero two.
- The naive product of the marginals is only point zero one, half the truth.

13 Example 4: independent, yet dependent given C

Illustration: two fair coins. C = at least one head

P(C)	3 of 4 outcomes	**3/4**
P(A | C) and P(B | C)	(1/2) ÷ (3/4)	**2/3 each**
Product of the two	2/3 × 2/3	**4/9**
Actual P(A and B | C)	(1/4) ÷ (3/4)	**1/3**
Verdict	3/9 is not 4/9	**dependent**

IN THE LESSON

- The reverse can also happen.
- Flip two fair coins.
- Let A be that the first is heads and B that the second is heads.
- They are independent: each is one half, and the joint is one quarter.
- Now let C be the event that at least one flip is heads, which covers three of the four outcomes, so it has probability three quarters.
- Given C, the probability that the first is heads is one half divided by three quarters, which is two thirds.
- The same holds for the second.
- If they were still independent, the joint given C would be two thirds times two thirds, four ninths.
- But both heads, given C, is one quarter divided by three quarters, one third.
- Three ninths is not four ninths, so conditioning created dependence.

14 Plain versus conditional independence

Never assume independence the question has not stated

	What it says
Plain independence	Joint = P(A) × P(B)
Conditional independence	Joint | C = product | C
Firms and the economy	Conditionally independent
Two coins and 'at least one'	Independent, not given C
Unconditional joint	Weight each state's joint

IN THE LESSON

- Let's summarize the difference.
- Plain independence is a statement about the marginal probabilities: the joint equals the product of the two probabilities.
- Conditional independence is the same statement but inside a fixed condition: the joint given C equals the product of the conditional probabilities.
- Neither implies the other.
- The firm example showed events that are conditionally independent but unconditionally dependent, because a common factor drives both.
- The coin example showed events that are independent but dependent once we condition on something both of them influence.
- To find an unconditional joint probability, compute the joint within each state and then weight it by the probability of the state.
- The most common mistake is to multiply unconditional probabilities when the question never said they were independent.

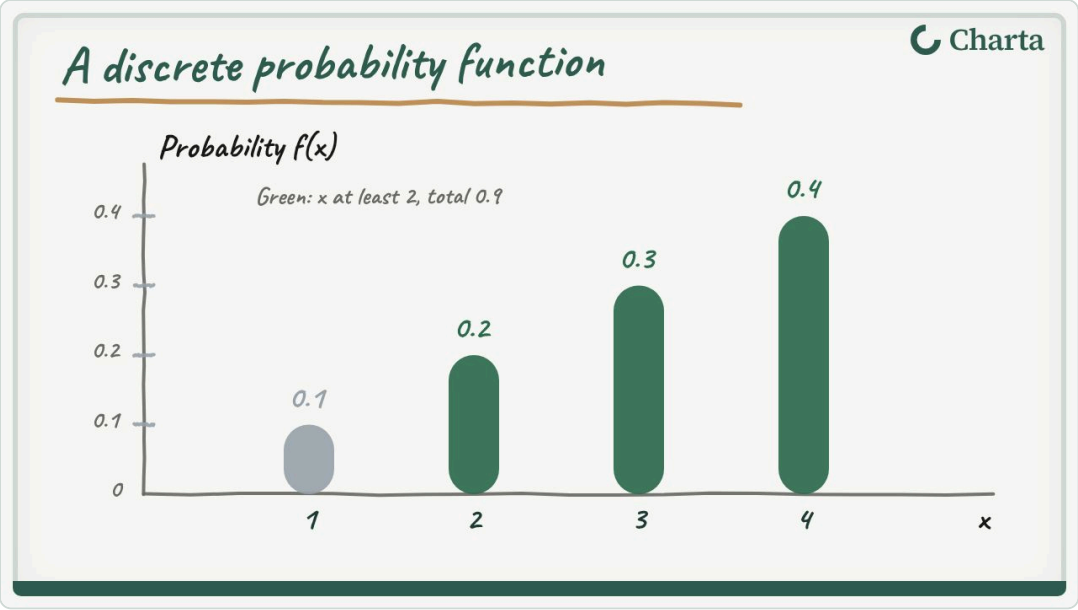
15 Discrete probability functions

- f(x) = probability that the variable equals x
- Rule 1: each f(x) lies between 0 and 1
- Rule 2: probabilities over all values add to exactly 1
- P(event) = the sum of f(x) over the values in the event

IN THE LESSON

- Now we move to random variables that take separate values, such as the number of defaults among a few loans.
- A discrete probability function, also called a probability mass function, tells us for each possible value the probability that the variable takes it.
- We write it as f of the value.
- Two properties must hold.
- Every probability lies between zero and one, and the probabilities over all possible values add to exactly one.
- The probability of any event is then the sum of the probabilities of the values that make it up.
- So the chance that the variable is at least two adds the probability of two, of three, and so on.
- Never use a function whose total is not one.

16 A discrete probability function



IN THE LESSON

- Here is a probability function drawn as bars.
- The variable can take the values one, two, three or four.
- The height of each bar is the probability of that value.
- The first bar is point one, the second point two, the third point three, and the fourth point four.
- Add the four heights and you get exactly one, which is what makes this a valid function.
- Now look at the green bars.
- They show the event that the variable is at least two.
- Its probability is the sum of the three green bars, point two plus point three plus point four, which is point nine.
- The grey bar is the single value one, and the event is simply everything else.

17 Example 5: finding the unknown constant

Illustration: f(x) = c × x for x = 1, 2, 3, 4

Total must be 1	c × (1+2+3+4) = 1	**c = 0.1**
P(X ≥ 2)	1 − f(1)	**0.9**
P(2 ≤ X ≤ 3)	0.2 + 0.3	**0.5**
P(X > 2), strict	0.3 + 0.4	**0.7**

IN THE LESSON

- Suppose a function says that the probability of the value x is a constant times x, for the values one, two, three and four, and zero otherwise.
- The constant is unknown, so set the total to one first.
- The values add to ten, so the constant times ten equals one, and the constant is point one.
- That gives the bars we just saw.
- Now the events.
- The probability the variable is at least two is one minus the probability of one, point nine.
- The probability that it lies from two to three inclusive adds point two and point three, giving point five.
- And strictly greater than two keeps only three and four, point three plus point four, point seven.
- Notice how strict and non strict inequalities change which end values are included.

18 Conditional probability

- Knowing B shrinks the sample space to B
- P(A | B) = P(A and B) ÷ P(B), defined only if P(B) > 0
- Multiplication rule: P(A and B) = P(A | B) × P(B)
- Draws without replacement: 2nd depends on the 1st

IN THE LESSON

- Conditional probability answers: what is the chance of A once we know B has occurred?
- We write it as the probability of A given B.
- The idea is that the knowledge shrinks the world.
- Only outcomes inside B are still possible, so B becomes the new sample space.
- The joint probability of A and B is then rescaled by dividing by the probability of B.
- This only works when the probability of B is positive.
- Rearranged, the same idea gives the multiplication rule: the probability that both occur equals the probability of A given B, times the probability of B.
- In sampling without replacement, the second draw is conditional on the first, because the pool has changed.

19 The conditional probability formula

$P(A \mid B) = P(A \cap B) \div P(B)$
$P(A \cap B) = P(A \mid B) \times P(B)$

Conditional probability	Chance of A once B is known
Joint, both occur	A and B both occur
Probability of the given event	Probability of the given event B

IN THE LESSON

- Here is the formula and its rearrangement.
- The conditional probability of A given B equals the joint probability of A and B, divided by the probability of B.
- The given event, the one we are told has occurred, goes in the denominator.
- That is the single most important thing to remember.
- Turn the formula around and you get the multiplication rule, which builds a joint probability from a conditional one and a marginal one.
- The formula is not symmetric.
- The probability of A given B divides by the probability of B, whereas the probability of B given A divides by the probability of A, and the two answers are usually different.

20 Example 6: drawing loans without replacement

Illustration: 10 loans, 3 bad. Two picks, no replacement

First pick is bad	3 ÷ 10	**0.30**
Second bad, given first bad	2 ÷ 9	**0.2222**
Both are bad	0.30 × 2/9	**0.0667**
Wrong: reuse 3 ÷ 10	0.30 × 0.30	**0.09**

IN THE LESSON

- A small book holds ten loans, of which three are bad.
- We pick two loans at random, one after the other, without putting the first back.
- The probability the first is bad is three over ten, point three.
- Given that the first was bad, the pool now holds nine loans, two of them bad.
- So the probability that the second is also bad is two over nine, point two two two two.
- The multiplication rule gives the joint: point three times two ninths, which is point zero six six seven, or one in fifteen.
- Compare the tempting mistake of using three tenths again for the second draw.
- That would give point zero nine, which treats the draws as independent and overstates the chance.

21 Example 7: a bond that has survived

Illustration: default in year 1 is 3%; default in year 2 (from today) is 4%

Survive year one	1 − 3%	**97%**
Year-2 default, from today	given	**4%**
Year-2 default, survived	4% ÷ 97%	**4.12%**
Wrong: divide by 1	4% ÷ 1	**4%**

IN THE LESSON

- Conditioning matters in credit.
- Suppose a bond has a three per cent chance of defaulting in year one, and an unconditional four per cent chance of defaulting in year two, meaning default specifically in that year as seen from today.
- Now a year goes by and the bond has not defaulted.
- What is the chance it defaults in year two, given survival?
- The survival probability is one minus point zero three, which is ninety seven per cent.
- Divide the four per cent by ninety seven per cent and you get four point one two per cent.
- The error to avoid is dividing by one, which leaves four per cent.
- We must divide by the probability of the given event, which is survival.

22 Conditioning is not symmetric

- P(default | recession) is not P(recession | default)
- Same joint; they differ by P(recession) ÷ P(default)
- Ask first: which event has been given?
- Put the given event's probability in the denominator

IN THE LESSON

- A frequent trap is to flip a conditional probability the wrong way round.
- The probability of default given a recession is not the probability of recession given default.
- The first divides the joint by the chance of a recession.
- The second divides the same joint by the chance of default.
- So the two answers differ by the ratio of the chance of a recession to the chance of default.
- When one of those events is much more common than the other, the two conditional probabilities can be very far apart, and they say very different things.
- Always ask which event is the given one, and put its probability underneath.

23 Unconditional versus conditional

- Unconditional (marginal): no extra information used
- Conditional: uses a given fact, e.g. 'in a recession'
- Unconditional = weighted average of conditional ones
- If independent, conditioning changes nothing
- Never apply a conditional figure to everyone

IN THE LESSON

- An unconditional probability, also called a marginal probability, uses no extra information.
- The chance that a loan in the book defaults next year is an example.
- A conditional probability uses extra information, such as the chance of default given that the economy is in recession.
- The two are linked.
- The unconditional probability is a weighted average of the conditional ones, with the probabilities of the states as the weights.
- If the events are independent, the condition changes nothing, so the conditional probability equals the unconditional one.
- Otherwise conditioning moves the figure up or down.
- The usual mistake is to take a conditional number, such as the default rate in a recession, and apply it as if it described the whole book.

24 The law of total probability

$P(A) = \Sigma\, P(A \mid B_i) \times P(B_i)$
states B_i: exclusive and exhaustive

Unconditional probability	Overall chance of A in any state
Conditional probability	Chance of A inside state i
Probability of the state	Chance of state i: the weight
Sum over all states	Add the products, state by state

IN THE LESSON

- The link between the two is the law of total probability.
- Split the world into states that are mutually exclusive and exhaustive, meaning exactly one of them will occur.
- For each state, take the probability of A within that state and multiply by the probability of the state itself.
- Add these products over all the states.
- The result is the unconditional probability of A.
- It is a weighted average, and the weights are the state probabilities, which add to one.
- This formula is the engine behind the earlier firm example, and it is also the denominator in Bayes' rule, as we will see.
- If the states do not cover everything, or overlap, the formula fails.

25 Example 8: default across three states

Illustration: state probabilities 50%, 30%, 20% add to 100%

Boom	50% × 1%	**0.50%**
Normal	30% × 3%	**0.90%**
Recession	20% × 10%	**2.00%**
Unconditional default	0.50 + 0.90 + 2.00	**3.40%**

IN THE LESSON

- Here is a made up loan book in three economic states.
- A boom has probability fifty per cent, and the default rate within a boom is one per cent.
- A normal state has probability thirty per cent and a default rate of three per cent.
- A recession has probability twenty per cent and a default rate of ten per cent.
- The state probabilities add to one, so the states are valid.
- Multiply each pair.
- Boom: point five times point zero one is half a per cent.
- Normal: point three times point zero three is point nine per cent.
- Recession: point two times point one is two per cent.
- Add them.
- The unconditional default rate is three point four per cent.
- It lies between the lowest and highest conditional rates, as a weighted average must.

26 Bayes' rule: reversing a conditional

- Prior: our belief before the evidence
- Likelihood: P(evidence | hypothesis true)
- Posterior: P(hypothesis | evidence), the updated belief
- Posterior ∝ prior × likelihood, then scaled to total 1

IN THE LESSON

- Bayes' rule is named after Thomas Bayes, an eighteenth century clergyman.
- It lets us reverse a conditional probability.
- We often know the chance of seeing some evidence if a hypothesis is true, but what we want is the chance the hypothesis is true once we have seen the evidence.
- Three names matter.
- The prior is our belief before the evidence.
- The likelihood is the probability of the evidence if the hypothesis is true.
- The posterior is the updated belief after the evidence.
- The prior gets multiplied by the likelihood, and then the result is scaled so that the posteriors over all hypotheses add to one.
- That scaling factor is the total probability of the evidence.

27 Bayes' rule and the evidence

$$P(A \mid B) = P(B \mid A) \times P(A) \div P(B)$$
$$P(B) = P(B \mid A)P(A) + P(B \mid \neg A)P(\neg A)$$

Posterior	Updated chance of A after seeing B
Likelihood	Chance of B if A is true
Prior	Chance of A before the evidence
Total probability of evidence	Chance of seeing B at all

IN THE LESSON

- Here is the rule.
- The posterior equals the likelihood times the prior, divided by the total probability of the evidence.
- The denominator comes straight from the law of total probability.
- The evidence can occur if the hypothesis is true, which has probability likelihood times prior, or if it is false, which has the probability of the evidence given that the hypothesis is false, written the probability of B given not A, times the probability that the hypothesis is false.
- Add those two products to get the total.
- With more than two hypotheses, you divide each likelihood times prior product by the total of all such products.
- A reliable routine is to write down the prior, the two likelihoods and the total before you divide.

28 Example 9: a default warning flag

Illustration: prior default 2%. Flags 90% of defaulters, 10% of non-defaulters

Flagged and defaults	0.02 × 0.90	**0.018**
Flagged, no default	0.98 × 0.10	**0.098**
Total flagged	0.018 + 0.098	**0.116**
Chance of default if flagged	0.018 ÷ 0.116	**15.5%**

IN THE LESSON

- Imagine a model that flags borrowers who are likely to default.
- Two per cent of borrowers actually default.
- The model flags ninety per cent of those who default, and wrongly flags ten per cent of those who do not.
- A borrower is flagged.
- What is the chance that she defaults?
- First, defaulters who are flagged: point zero two times point nine zero, which is point zero one eight.
- Next, non defaulters who are flagged: point nine eight times point one zero, point zero nine eight.
- The total flagged is point one one six.
- The posterior is point zero one eight divided by point one one six, which is fifteen point five per cent.
- Most flags come from the large group of safe borrowers.

29 Example 10: a second flag arrives

Illustration: prior is now 15.5%. Second flag, independent given default status

Default and flagged again	0.1552 × 0.90	**0.1397**
No default, flagged again	0.8448 × 0.10	**0.0845**
Total	0.1397 + 0.0845	**0.2242**
Posterior after two flags	0.1397 ÷ 0.2242	**62.3%**

IN THE LESSON

- Bayes' rule can be applied again and again, and it ties back to conditional independence.
- Suppose the model flags the same borrower a second time, and the two flags are independent given whether she defaults.
- Our new prior is the old posterior, fifteen point five per cent.
- Defaulting and flagged: point one five five two times point nine, equals point one three nine seven.
- Not defaulting and flagged: point eight four four eight times point one, which is point zero eight four five.
- The total is point two two four two.
- The new posterior is point one three nine seven over point two two four two, about sixty two per cent.
- One flag took us from two to fifteen, and a second took us above sixty.

30 Four mistakes to avoid

- Exclusive (chances > 0) is dependent; overlap: subtract
- Never multiply marginals unless independence is stated
- The given event goes below: P(A | B) is not P(B | A)
- A likelihood is not a posterior: use Bayes' rule

IN THE LESSON

- Mistake one: treating mutually exclusive as meaning unrelated, or adding probabilities of events that overlap.
- Exclusive events with positive chances are dependent, and overlapping events need the overlap subtracted.
- Mistake two: assuming independence that the question never stated, and multiplying the marginal probabilities.
- Mistake three: dividing by the wrong event.
- The given event goes in the denominator, and the probability of A given B is not the probability of B given A.
- Mistake four: reading a likelihood as a posterior.
- The share of defaulters that a test flags is not the chance a flagged borrower defaults.
- Use Bayes' rule, with the prior and the total.

31 Recap: the main ideas

- Sample space, events, event space: three closure rules
- Exclusive: union adds. Independent: joint multiplies
- Conditional independence is a separate property
- Discrete function: values in [0, 1], total exactly 1
- Conditional = joint ÷ given; total prob.; Bayes' rule

IN THE LESSON

- Let's pull it together.
- A probability model starts with a sample space of outcomes, and an event is a subset.
- The event space must contain the sample space, every complement, and every countable union.
- Mutually exclusive events cannot co occur, so their union adds.
- Independent events share no information, so their joint multiplies.
- The two are close to opposites.
- Conditional independence is a separate property that can hold or fail on its own.
- A discrete probability function has probabilities between zero and one that add to one.
- Conditional probability divides the joint by the given event.
- Total probability averages over states, and Bayes' rule reverses the conditional to turn a prior into a posterior.

32 Try three quick questions

1. Independent, P(A) = 0.5, P(B) = 0.2. P(A or B)?
   Answer: 0.5 + 0.2 − 0.1 = 0.6.
2. P(A and B) = 0.12, P(B) = 0.4. P(A | B)?
   Answer: 0.12 ÷ 0.4 = 0.30. Divide by the given event.
3. Prior 1%, flags 80% of defaulters, 5% of others. P(default | flag)?
   Answer: 0.008 ÷ (0.008 + 0.0495) = 13.9%.

IN THE LESSON

- Now try three quick questions yourself.
- Pause after each one.
- Question one.
- Two events are independent, with probabilities of point five and point two.
- What is the probability that at least one occurs?
- The joint is point one, so the union is point five plus point two, minus point one, which is point six.
- Question two.
- The joint probability of A and B is point one two, and the probability of B is point four.
- What is the probability of A given B?
- Point one two over point four, which is point three.
- Question three.
- Prior default is one per cent, the test flags eighty per cent of defaulters and five per cent of others.
- The posterior is point zero zero eight over point zero five seven five, about thirteen point nine per cent.