CHARTA LEARNING · LESSON NOTES

Common Univariate Random Variables

FRM Part I

1 A toolbox of shapes

- Every uncertain quantity needs a shape
- Counts, waiting times, returns, prices, recoveries
- Each named distribution has a job
- Mixtures blend shapes when one is not enough

IN THE LESSON

- Think of a workshop with a wall of tools.
- A good carpenter does not hammer every job; she picks the tool that fits.
- Risk modelling is the same.
- Every uncertain quantity needs a shape that describes how likely each outcome is.
- How many of our loans default this year?
- How long until the next failure?
- What will a share price be in a year?
- What fraction of a bond will we get back after a default?
- Each question has a distribution that fits it naturally.
- Learning the named distributions is learning the tools, and learning what each is for.
- Mixtures then let us blend tools when one shape is not enough.

2 What a distribution tells us

- A random variable is a number we cannot know in advance
- Discrete: countable values. Continuous: any value in a range
- Density: probability is area. CDF: probability up to a point
- Parameters (p, n, λ, μ, σ) pin down the exact shape

IN THE LESSON

- A random variable is a number we cannot know in advance, such as tomorrow's return.
- Its distribution tells us how likely each value is.
- If the values can be counted, like the number of defaults, the variable is discrete, and we give each value a probability.
- If it can take any value in a range, like a return, it is continuous, and we use a density, where probability is the area under the curve.
- The cumulative distribution function adds up probability to a given point.
- Every named distribution has parameters, a few numbers such as a mean or a rate, that pin down its exact shape.

3 Bernoulli and binomial

- Bernoulli: one trial, success with probability p
- Mean p, variance p(1 − p); a default flag is Bernoulli
- Binomial: counts successes in n independent trials
- Same p each time, and the trials must be independent
- Binomial = sum of n Bernoulli variables

IN THE LESSON

- Start with the simplest random variable.
- A Bernoulli variable describes one trial with two outcomes, success with probability p, and failure otherwise.
- We code success as one and failure as zero.
- A loan either defaults or it does not, so a default flag is Bernoulli.
- Its mean is p and its variance is p times one minus p.
- Now repeat the trial many times.
- The binomial counts the successes in a fixed number of trials, say n.
- It needs two conditions: the probability of success is the same each time, and the trials are independent.
- A binomial is just the sum of n Bernoulli variables.

4 The binomial formula

P(X = k) = C(n,k) · p^k · (1 − p)^{n − k}
Mean = np Variance = np(1 − p)

n: trials	Number of independent trials
p: chance	Chance of success on each trial
k: successes	Successes we are counting
Mean, variance	Mean np, variance np(1−p)

IN THE LESSON

- Here is the binomial.
- The probability of exactly k successes in n trials is the number of ways to choose which k trials succeed, times p to the power k, times one minus p to the power n minus k.
- The number of ways is the combination, n choose k.
- The mean is n times p, which is just the average success rate times the number of trials.
- The variance is n times p times one minus p, which is the Bernoulli variance added up n times, because independent variances add.
- Notice the variance is largest when p is one half, and shrinks as p moves toward zero or one.

5 Example 1: defaults in a loan book

Illustration: 10 independent loans, each with a 5% default probability

Mean, np	10 × 0.05	**0.5**
Variance, np(1 − p)	10 × 0.05 × 0.95	**0.475**
P(no default)	0.95^{10}	**59.87%**
P(exactly two)	$45 \times 0.05^2 \times 0.95^8$	**7.46%**

IN THE LESSON

- Illustration.
- A bank holds ten independent loans, each with a five per cent chance of default.
- The mean number of defaults is ten times zero point zero five, which is zero point five.
- The variance is ten times zero point zero five times zero point nine five, which is zero point four seven five.
- The chance of no default at all is zero point nine five to the tenth power, about fifty nine point eight seven per cent.
- The chance of exactly two defaults uses forty five ways to choose which two, times zero point zero five squared, times zero point nine five to the eighth power.
- That comes to seven point four six per cent.

6 Poisson and exponential

- Poisson: number of events in a period, at average rate λ
- P(N = n) = e^{−λ} · λ^n ÷ n!
- Mean = variance = λ; events arrive independently
- Exponential: the waiting time between events
- Exponential mean = 1 ÷ λ

IN THE LESSON

- Sometimes we count events with no fixed number of trials, such as defaults in a year, or fraud alerts in a day.
- If events arrive independently at an average rate lambda per period, the count follows a Poisson distribution.
- The probability of exactly n events is e to the minus lambda, times lambda to the n, divided by n factorial.
- The Poisson has a neat property: its mean and its variance are both equal to lambda.
- Now flip the question.
- Instead of counting events, ask how long until the next one.
- That waiting time is exponential, and its mean is one over lambda.
- If three defaults arrive per year, the average wait is a third of a year.

7 Example 2: Poisson defaults

Illustration: on average 3 defaults per year, so λ = 3

Mean and variance	both equal λ	**3**
P(no default)	e^{-3}	**4.98%**
P(exactly two)	$e^{-3} \times 3^2 \div 2!$	**22.40%**
P(at least one)	$1 - e^{-3}$	**95.02%**

IN THE LESSON

- Illustration.
- A portfolio suffers on average three defaults per year, so lambda is three.
- The mean and the variance of the yearly count are both three.
- The chance of no defaults is e to the minus three, about four point nine eight per cent.
- The chance of exactly two defaults is e to the minus three, times three squared, divided by two factorial, which is twenty two point four zero per cent.
- The chance of at least one default is one minus e to the minus three, which is ninety five point zero two per cent.
- A quiet year with no defaults is possible, but rare, at about one chance in twenty.

8 Example 3: the wait for a default

Illustration: same portfolio, λ = 3 defaults per year

Mean wait	1 ÷ 3 years	**0.333 yrs**
P(wait > half a year)	e^{−3 × 0.5}	**22.31%**
Poisson check, mean 1.5	P(no default) = e^{−1.5}	**22.31%**
P(wait ≤ half a year)	1 − 0.2231	**77.69%**

IN THE LESSON

- Same portfolio, lambda of three defaults per year.
- How long do we wait for the next default?
- The mean wait is one divided by three, which is zero point three three three years, about four months.
- What is the chance the wait is longer than half a year?
- For an exponential it is e to the minus lambda times time, so e to the minus three times zero point five, which is twenty two point three one per cent.
- The chance the wait is half a year or less is one minus that, seventy seven point six nine per cent.
- This matches the Poisson view: no defaults in half a year, when the expected count is one and a half, has the same probability.

9 The uniform distribution

- Uniform on a to b: every value equally likely
- Flat density 1 ÷ (b − a) inside, zero outside
- Mean (a + b) ÷ 2, variance $(b - a)^2 \div 12$
- P(c to d) = (d − c) ÷ (b − a): a ratio of lengths
- Standard uniform (0 to 1) feeds simulations

IN THE LESSON

- The uniform is the distribution of complete ignorance within a range.
- Every value between a lower limit a and an upper limit b is equally likely, so the density is flat, equal to one over b minus a inside the interval and zero outside.
- The mean is the midpoint, a plus b over two.
- The variance is b minus a, squared, over twelve.
- Probabilities are easy: the chance of landing between two points is the length of that stretch divided by the full length.
- The standard uniform runs from zero to one, and it matters in practice because simulation programs start from uniform random numbers and transform them into other distributions.

10 Example 4: a loss spread evenly

Illustration: claim uniform between 0 and 100 (millions of dollars)

Mean	(0 + 100) ÷ 2	**50**
Variance	(100 − 0)² ÷ 12	**833.3**
Standard deviation	√833.3	**28.87**
P(claim > 80)	(100 − 80) ÷ 100	**20%**

IN THE LESSON

- Illustration.
- Suppose a claim could be any amount from zero to one hundred million dollars, with every amount equally likely.
- The mean is zero plus one hundred, divided by two, which is fifty million.
- The variance is one hundred minus zero, squared, divided by twelve, which is eight hundred thirty three point three.
- The standard deviation is the square root of that, twenty eight point eight seven million.
- What is the chance the claim exceeds eighty million?
- The remaining stretch is one hundred minus eighty, which is twenty, divided by the full length of one hundred, giving twenty per cent.
- No tables needed, just lengths.

11 The normal distribution

- Symmetric bell, set by the mean μ and σ
- About 68% within 1σ and 95% within 2σ
- Standardize: z = (x − μ) ÷ σ, then use the tables
- A sum of independent normals is itself normal
- Thin tails: extreme moves are rare in the model

IN THE LESSON

- The normal is the bell curve, and the most important distribution in risk.
- It is symmetric around its mean mu, and its width is set by the standard deviation sigma.
- About sixty eight per cent of outcomes fall within one sigma of the mean, and about ninety five per cent within two.
- To find probabilities we standardize: subtract the mean and divide by sigma, giving a z score that follows the standard normal, with mean zero and standard deviation one.
- Sums of independent normal variables are themselves normal, which makes the normal convenient.
- Its weakness is thin tails: very large moves are extremely rare under a normal, rarer than they often are in real markets.

12 Standard normal points to know

z scores that appear again and again in risk work

Standard normal z	Probability below z
0	50.00%
1.000	84.13%
1.645	95.00%
1.960	97.50%
2.326	99.00%

IN THE LESSON

- Here are the standard normal points worth memorizing.
- At z equal to zero, half the probability lies below.
- At one, about eighty four per cent lies below.
- At one point six four five, ninety five per cent lies below, so five per cent lies in the upper tail.
- That is the usual cut-off for a ninety five per cent one sided value at risk.
- At one point nine six, ninety seven and a half per cent lies below, which leaves two and a half per cent in each tail of a two sided interval.
- And at two point three two six, ninety nine per cent lies below, the cut-off for ninety nine per cent value at risk.
- By symmetry, the lower tail uses the same numbers with a minus sign.

13 Example 5: a normal monthly return

Illustration: monthly return ~ Normal(mean 1%, std dev 4%)

P(return < 0)	z = (0 − 1) ÷ 4 = −0.25	**40.13%**
P(return < −7%)	z = (−7 − 1) ÷ 4 = −2	**2.28%**
5% worst cut-off	1% − 1.645 × 4%	**−5.58%**
1% worst cut-off	1% − 2.326 × 4%	**−8.30%**

IN THE LESSON

- Illustration.
- A portfolio's monthly return is normal, with a mean of one per cent and a standard deviation of four per cent.
- What is the chance of a loss?
- The z score is zero minus one, over four, which is minus zero point two five, and the probability is forty point one three per cent.
- What is the chance of falling below minus seven per cent?
- The z score is minus seven minus one, over four, which is minus two, giving two point two eight per cent.
- Now work backwards.
- The five per cent worst cut-off is one per cent minus one point six four five times four per cent, which is minus five point five eight per cent.
- The one per cent cut-off is one per cent minus two point three two six times four per cent, minus eight point three zero per cent.

14 The lognormal distribution

- Lognormal: the natural log of the variable is normal
- Always positive and right-skewed: the model for prices
- Median = exp(μ); mean = exp(μ + σ²/2)
- Mean exceeds median, and the gap grows with σ
- Slip to avoid: calling exp(μ) the mean

IN THE LESSON

- A normal variable can be negative, but a share price cannot.
- The fix is the lognormal: a variable is lognormal if its natural logarithm is normal.
- Raising e to a normal number always gives a positive value, and the result is right skewed, with a long tail of large values.
- That is why the lognormal is the standard model for prices.
- Now the key point.
- If the log has mean mu and standard deviation sigma, the median is e to the mu.
- The mean is e to the mu plus sigma squared over two.
- The mean is larger than the median, and the gap grows with sigma.
- The usual slip is to call e to the mu the mean.

15 Example 6: a lognormal share price

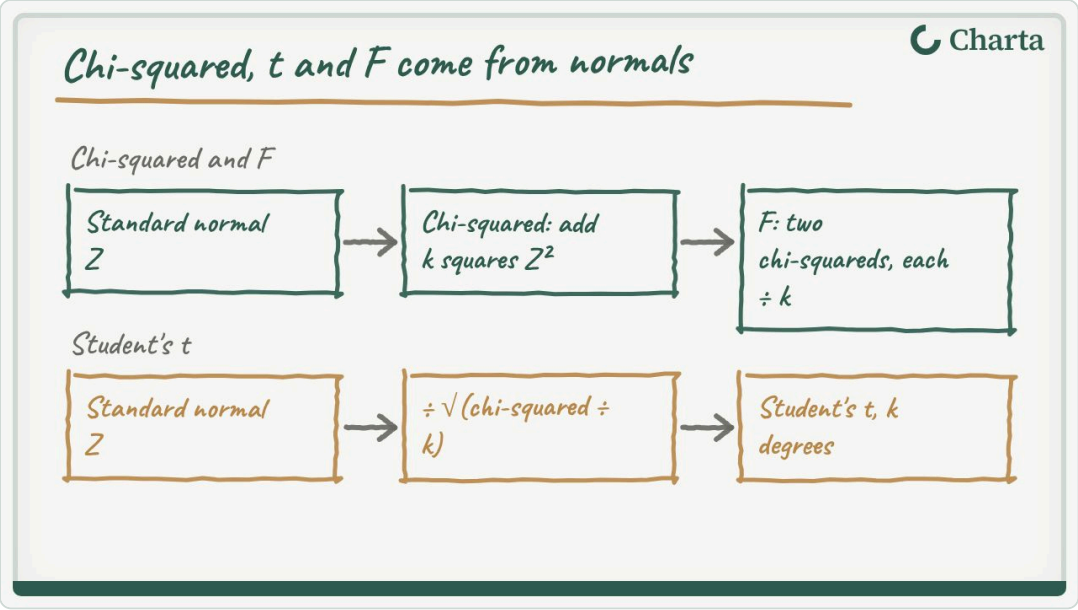
Illustration: price today 100; log growth ~ Normal(mean 8%, std dev 20%)

Median price	100 × e^{0.08}	**108.33**
Mean price	100 × e^{0.08 + 0.02}	**110.52**
P(price < 100)	z = (0 − 0.08) ÷ 0.2	**34.46%**
P(price > 150)	z = (ln 1.5 − 0.08) ÷ 0.2	**5.18%**

IN THE LESSON

- Illustration.
- A share is priced at one hundred today.
- The log of next year's price relative to today is normal, with mean eight per cent and standard deviation twenty per cent.
- The median price is one hundred times e to the zero point zero eight, which is one hundred eight point three three.
- The mean price adds sigma squared over two, which is zero point zero two, so one hundred times e to the zero point one, one hundred ten point five two.
- The mean sits above the median.
- The chance the price ends below one hundred uses z of minus zero point four, giving thirty four point four six per cent.
- The chance it exceeds one hundred fifty is five point one eight per cent.

16 Chi-squared, t and F come from normals



IN THE LESSON

- Three more distributions are all built from the standard normal.
- Take independent standard normal variables and square each.
- Add up k of these squares and you have a chi squared variable with k degrees of freedom.
- Take two independent chi squared variables, divide each by its own degrees of freedom, and form their ratio.
- That is an F variable.
- Student's t takes a different route.
- Divide a standard normal by the square root of an independent chi squared variable divided by its degrees of freedom.
- Student was the pen name of William Gosset, who worked at the Guinness brewery.
- These three appear constantly in testing and in building confidence intervals.

17 Chi-squared and Student's t

$$\chi^2_k = Z_1^2 + Z_2^2 + \dots + Z_k^2$$
$$t_k = Z \div \sqrt{(\chi^2_k \div k)}$$

Standard normal	Independent standard normal
Degrees of freedom	k, the degrees of freedom
Chi-squared	Chi-squared: mean k, variance 2k
Student's t	t: variance k ÷ (k − 2), k > 2

IN THE LESSON

- Here are the definitions side by side.
- Chi squared with k degrees of freedom is the sum of k squared independent standard normal variables.
- It is never negative, it is right skewed, its mean is k and its variance is two k.
- As k grows it looks more and more normal.
- Student's t is a standard normal divided by the square root of an independent chi squared over its degrees of freedom.
- It is symmetric like the normal, but with fatter tails.
- Its variance is k over k minus two, defined when k is above two.
- As k grows, that variance falls toward one, and the t becomes the standard normal.

18 Fatter tails: the t versus the normal

97.5% points: 2.5% of probability lies beyond each one

Degrees of freedom k	97.5% point
5	2.571
10	2.228
30	2.042
Infinite (the normal)	1.960

IN THE LESSON

- To see the fatter tails, compare cut-off points.
- For a two sided ninety five per cent interval the standard normal uses one point nine six.
- A t with thirty degrees of freedom needs two point zero four two, a little more.
- With ten degrees it needs two point two two eight.
- With five degrees of freedom it needs two point five seven one.
- The lower the degrees of freedom, the more probability sits in the tails, so you must go further out to catch ninety five per cent.
- With an infinite number of degrees of freedom the t is exactly the normal, at one point nine six.
- That is why small samples use the t.

19 Example 7: moments of chi-squared and t

Illustration: chi-squared with k = 4; Student's t with k = 5

Chi-squared mean	k	**4**
Chi-squared variance	2k = 2 × 4	**8**
t variance, k = 5	5 ÷ (5 − 2)	**1.67**
t std dev, k = 5	√1.67	**1.29**

IN THE LESSON

- Illustration.
- Take a chi squared with four degrees of freedom.
- Its mean equals the degrees of freedom, which is four.
- Its variance is twice that, which is eight.
- Now take a Student's t with five degrees of freedom.
- Its variance is k over k minus two, which is five over three, or one point six seven.
- That is more than the variance of one for a standard normal, which is the fat tail effect.
- The standard deviation is the square root, about one point two nine.
- With thirty degrees of freedom the variance would be thirty over twenty eight, only one point zero seven, close to the normal.

20 The F distribution

- F = (A ÷ k_1) ÷ (B ÷ k_2), A and B chi-squared
- Two degrees-of-freedom numbers: top and bottom
- Positive and right-skewed; near 1 when variances match
- Used to compare variances, and to test regressions

IN THE LESSON

- The F distribution compares variances.
- It is the ratio of two independent chi squared variables, each divided by its own degrees of freedom.
- So an F has two degrees of freedom numbers, one for the numerator and one for the denominator.
- It is positive and right skewed, and when the two underlying variances really are equal, the ratio of sample variances tends to land near one.
- A ratio far above one is evidence that the top variance is larger.
- This is the basis of the test for equal variances, and it also appears in regression, where it tests whether a group of coefficients is jointly zero.
- For the course, remember what it is built from and what it is used for.

21 Example 8: comparing two variances

Illustration: sample variances of monthly returns for two funds

Fund A variance	3% squared = 0.03^2	**0.0009**
Fund B variance	2% squared = 0.02^2	**0.0004**
F statistic	0.0009 ÷ 0.0004	**2.25**
If variances were equal	F would sit near	**1**

IN THE LESSON

- Illustration.
- We estimate the variance of monthly returns for two funds.
- Fund A has a sample standard deviation of three per cent, so a sample variance of zero point zero zero zero nine.
- Fund B has a standard deviation of two per cent, so a variance of zero point zero zero zero four.
- The F statistic puts the larger on top: zero point zero zero zero nine divided by zero point zero zero zero four, which is two point two five.
- If the true variances were equal, we would expect a value near one.
- Whether two point two five is big enough depends on the degrees of freedom, which come from the two sample sizes, and a table or software gives the cut-off.

22 The beta distribution

- Beta lives on 0 to 1; shape set by parameters a and b
- Mean = a ÷ (a + b)
- Variance = ab ÷ [(a + b)² (a + b + 1)]
- Humped, skewed or U-shaped: flexible on a bounded range
- Suits recovery rates; not the market beta of the CAPM

IN THE LESSON

- The beta distribution lives on the interval from zero to one, so it suits anything that is a proportion, and in credit risk that means recovery rates, the fraction of a bond's value recovered after default.
- Two shape parameters, usually called a and b, control it.
- Its mean is a over a plus b.
- Its variance is a times b, divided by a plus b squared times a plus b plus one.
- By changing a and b it can be humped, skewed, or even U shaped, which gives a lot of flexibility on a bounded range.
- Do not confuse this distribution with the market beta of the capital asset pricing model.
- They share a name and nothing else.

23 Example 9: a recovery-rate distribution

Illustration: recovery rate ~ Beta(a = 2, b = 3)

Mean	2 ÷ (2 + 3)	**40%**
Variance	6 ÷ (25 × 6)	**0.04**
Standard deviation	√0.04	**20%**
Beta(4, 6): std deviation	√(24 ÷ (100 × 11))	**14.8%**

IN THE LESSON

- Illustration.
- Suppose the recovery rate on defaulted bonds is modelled as a beta with a equal to two and b equal to three.
- The mean is two over five, which is forty per cent.
- The variance is two times three, which is six, divided by five squared, which is twenty five, times six.
- That is six over one hundred fifty, which is zero point zero four.
- The standard deviation is the square root, twenty per cent.
- Now suppose we gain more evidence and double both parameters to four and six.
- The mean is still forty per cent, but the variance falls to zero point zero two one eight, so the standard deviation drops to about fourteen point eight per cent.

24 The named distributions at a glance

Mean ; variance. Also: normal μ ; σ², chi-squared k ; 2k

Distribution	Mean ; variance
Bernoulli(p)	p ; p(1 − p)
Binomial(n, p)	np ; np(1 − p)
Poisson(λ)	λ ; λ
Exponential(λ)	1/λ ; 1/λ²
Uniform(a, b)	(a+b)/2 ; (b−a)²/12

IN THE LESSON

- Here is a summary of the means and variances you should know cold.
- A Bernoulli has mean p and variance p times one minus p.
- A binomial multiplies both by n.
- A Poisson has mean and variance both equal to lambda.
- An exponential has mean one over lambda and variance one over lambda squared.
- A uniform from a to b has the midpoint as its mean and b minus a, squared, over twelve as its variance.
- Add to these the normal with mean mu and variance sigma squared, the chi squared with mean k and variance two k, and the t with variance k over k minus two.
- With these you can answer most questions on the named distributions.

25 Mixture distributions: the idea

- Step 1: draw which component applies (weights sum to 1)
- Step 2: draw a value from that component
- Density and CDF: weighted averages of the components
- Arises from regimes: calm and stressed, safe and risky
- Not a weighted sum: a mixture picks one, a sum adds all

IN THE LESSON

- Now the second topic.
- Sometimes no single named distribution fits, because the data come from more than one situation.
- A market is calm on most days and stressed on a few.
- A loan book has safe borrowers and risky borrowers.
- A mixture handles this in two stages.
- First, a random draw decides which component applies, using weights that add up to one.
- Second, we draw a value from that component.
- The density of the mixture is the weighted average of the component densities, and the cumulative distribution is the weighted average of the component cumulative distributions.
- This is different from a weighted sum of random variables, where you add the draws together.
- In a mixture you pick one component, you do not add.

26 Mixture mean and variance

Mean μ = Σ w_i μ_i
Variance = Σ w_i (σ_i^2 + μ_i^2) − μ^2

Weights	Weights w_i, adding up to one
Means	Component means μ_i
Variances	Component variances σ_i^2
Mixture mean	Mixture mean μ
Mixture variance	Weighted σ_i^2 + μ_i^2, less μ^2

IN THE LESSON

- The mean of a mixture is the weighted average of the component means.
- So far, easy.
- The variance needs more care.
- For each component, take its variance plus its mean squared, which is its second moment.
- Weight those second moments, add them up, then subtract the square of the mixture mean.
- The variance is not the weighted average of the variances.
- That shortcut works only if all the component means are equal, because otherwise it ignores the spread between the means.
- That extra spread is real.
- A mixture of two tight clusters far apart is very dispersed overall, even though each cluster on its own is tight.

27 Example 10: mixture with different means

Illustration: 60% Normal(mean 0, sd 2) mixed with 40% Normal(mean 5, sd 3)

Mixture mean	$0.6 \times 0 + 0.4 \times 5$	**2**
Mixture second moment	$0.6 \times 4 + 0.4 \times 34$	**16**
Mixture variance	$16 - 2^2$	**12**
Averaging variances (wrong)	$0.6 \times 4 + 0.4 \times 9$	**6**

IN THE LESSON

- Illustration.
- A mixture has two components.
- With weight sixty per cent, the mean is zero and the standard deviation is two.
- With weight forty per cent, the mean is five and the standard deviation is three.
- The mixture mean is point six times zero plus point four times five, which is two.
- The second moments are four for the first component, and nine plus twenty five, thirty four, for the second.
- Weighted, that is point six times four plus point four times thirty four, which is sixteen.
- The variance is sixteen minus two squared, which is twelve.
- The tempting shortcut, averaging the variances, gives point six times four plus point four times nine, which is six, only half the right answer.

28 Example 11: calm and stressed markets

Illustration: 80% calm (sd 1%), 20% stressed (sd 3%), both with mean 0

Variance	$0.8 \times 1 + 0.2 \times 9$	**2.6**
Standard deviation	$\sqrt{2.6}$	**1.61%**
Fourth moment	$0.8 \times 3 + 0.2 \times 243$	**51**
Kurtosis	$51 \div 2.6^2$	**7.54**
Excess kurtosis	$7.54 - 3$	**4.54**

IN THE LESSON

- Illustration.
- Daily returns come from two regimes with the same mean of zero.
- On eighty per cent of days the market is calm, with a standard deviation of one per cent.
- On twenty per cent of days it is stressed, with a standard deviation of three per cent.
- Because the means are equal, the variance is the weighted average: point eight times one plus point two times nine, which is two point six.
- The standard deviation is the square root, one point six one per cent.
- For a normal, the fourth moment is three sigma to the fourth: three times one for the calm regime, and three times eighty one, which is two hundred forty three, for the stressed one.
- Weighting gives fifty one.
- Kurtosis is fifty one divided by two point six squared, which is seven point five four, far above the normal's three.
- The excess kurtosis is about four point five.

29 Example 12: the tail probability

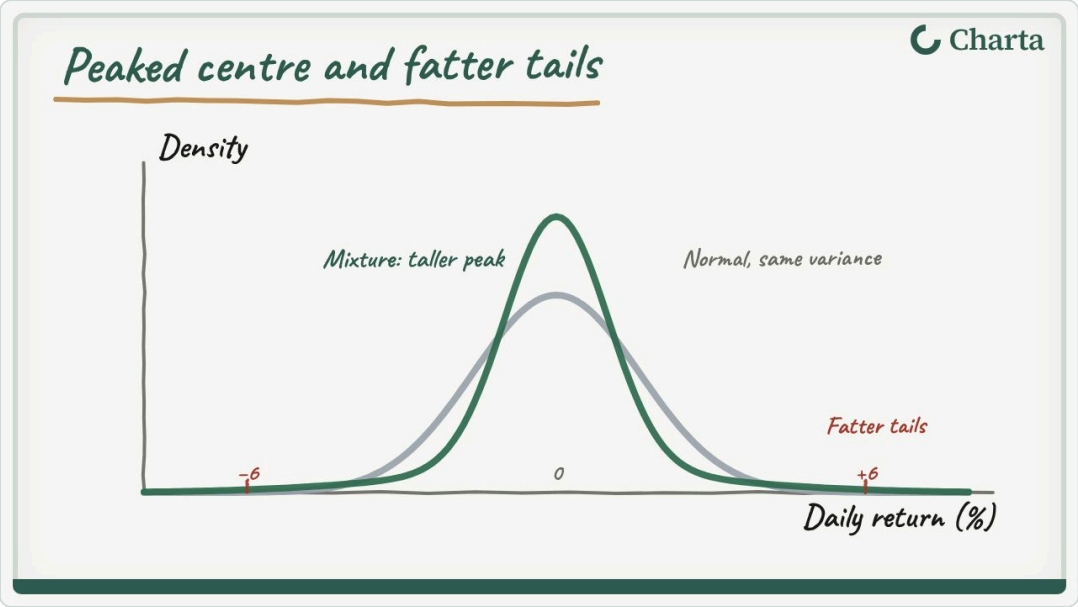
Illustration: P(daily loss worse than 6%), same mixture as Example 11

Calm: P(X < −6%)	z = −6	**about 0%**
Stressed: P(X < −6%)	z = −2	**2.28%**
Mixture	0.8 × 0 + 0.2 × 2.28%	**0.46%**
Normal, same sd 1.61%	z = −3.72	**0.0099%**
Mixture ÷ normal	0.455 ÷ 0.0099	**about 46**

IN THE LESSON

- Illustration.
- Using the same mixture, what is the chance of a daily loss worse than six per cent?
- Weight the probabilities, not the standard deviations.
- In the calm regime, six per cent is six standard deviations away, essentially zero.
- In the stressed regime it is two standard deviations, which is two point two eight per cent.
- The mixture is point eight times zero plus point two times two point two eight, about point four six per cent.
- Compare a plain normal with the same overall standard deviation of one point six one.
- Six per cent is three point seven two standard deviations, a probability of about point zero zero nine nine per cent.
- The mixture makes the extreme loss roughly forty six times more likely.

30 Peaked centre and fatter tails



IN THE LESSON

- Let's see the shape.
- The green curve is the calm and stressed mixture from the last examples.
- The grey curve is a normal with exactly the same mean and variance.
- Look at the centre first.
- The mixture is taller and more sharply peaked, because most days come from the calm regime with its small standard deviation.
- Now look to the sides.
- Far out, near minus six and plus six per cent, the green curve sits above the grey one, because the stressed regime puts real probability there.
- The two curves cannot match in variance and also match in shape.
- The mixture is symmetric, but it has the fat tails and the peak, which is what excess kurtosis means.

31 What mixtures can produce

Mixtures of normal components

If the components have	The mixture is
Same mean, different volatility	Symmetric with fat tails
Well separated means	Bimodal: two humps
Unequal means and weights	Can be skewed
Any normals mixed together	Not itself a normal

IN THE LESSON

- Mixing changes the shape in predictable ways.
- Mix normals with the same mean but different volatilities, and you get a symmetric distribution with excess kurtosis.
- That is the textbook way to produce fat tails, and regime switching between calm and stressed markets is the economic story behind it.
- Mix components with well separated means, and the result can have two humps, so it is bimodal.
- Mix components with unequal means and weights, and the result can be skewed.
- Whatever you do, a mixture of normals is not itself a normal.
- In the FRM, mixture problems usually ask for a tail probability or a standard deviation, so weight the probabilities, and compute the variance from second moments.

32 Four mistakes to avoid

- Lognormal mean is not exp(μ); that is the median
- Mixture variance: w_i(σ_i^2 + μ_i^2) − μ^2, not an average
- Mixture is not a weighted sum: weight probabilities
- Exponential mean is 1 ÷ λ, not λ; match time units

IN THE LESSON

- Mistake one: treating e to the mu as the mean of a lognormal.
- It is the median.
- Add sigma squared over two inside the exponent to get the mean, which is always larger.
- Mistake two: averaging variances in a mixture.
- Use each component's variance plus its mean squared, weight them, then subtract the square of the mixture mean.
- Mistake three: confusing a mixture with a weighted sum of random variables.
- A mixture weights probabilities, so for tails weight the tail probabilities, not the standard deviations.
- Mistake four: mixing up the exponential rate and its mean.
- The mean wait is one over lambda, not lambda, and the time units must match the Poisson rate.

33 Recap: the main ideas

- Binomial np, np(1 − p); Poisson λ; exponential mean 1 ÷ λ
- Uniform flat; normal symmetric; lognormal for prices
- Lognormal mean exp(μ + σ²/2) exceeds median exp(μ)
- χ^2, t, F from normals; t fatter tails; beta on 0 to 1
- Mixture: weight probabilities; same mean gives fat tails

IN THE LESSON

- Let's pull it together.
- Bernoulli is one trial, the binomial counts successes with mean np and variance np times one minus p.
- The Poisson counts events with mean and variance equal to lambda, and the exponential times the gaps, with mean one over lambda.
- The uniform is flat, the normal is the symmetric bell, and the lognormal is the model for prices, with a mean above its median.
- Chi squared, t and F are built from normals, and the t has fatter tails.
- The beta lives on zero to one.
- Finally, a mixture weights probabilities, its variance needs second moments, and mixing volatilities creates fat tails.

34 Try three quick questions

1. 20 loans, 10% default each. Mean and std dev of defaults?
   Answer: Mean 2; variance 20 × 0.1 × 0.9 = 1.8; std dev 1.34.
2. ln(price) ~ Normal(0.10, 0.30²). Median and mean multiples?
   Answer: Median e^0.10 = 1.105; mean e^0.145 = 1.156.
3. 50/50 mixture: means 0 and 4, both variance 4. Variance?
   Answer: Mean 2; 0.5×4 + 0.5×20 = 12; 12 − 4 = 8.

IN THE LESSON

- Now try three quick questions yourself.
- Pause after each one.
- Question one.
- Twenty independent loans each default with probability ten per cent.
- What are the mean and standard deviation of the number of defaults?
- The mean is two, the variance is twenty times point one times point nine, one point eight, and the standard deviation is about one point three four.
- Question two.
- The log of a price is normal with mean ten per cent and standard deviation thirty per cent.
- What are the median and the mean multiples?
- The median is e to the point one, one point one oh five; the mean is e to the point one four five, one point one five six.
- Question three.
- A fifty fifty mixture has components with means zero and four, both with variance four.
- What is the mixture variance?
- The mean is two, the second moment is twelve, so the variance is eight.