Profit and Loss

- **Profit/Loss** $P/L = P_t + D_t - P_{t-1}$

- **Arithmetic Return Data :**

$$r_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \frac{P_t + D_t}{P_{t-1}} - 1$$

- **Geometric Return Data :**

$$R_t = \ln\left(\frac{P_t + D_t}{P_{t-1}}\right) = \ln(1 + r_t)$$

Profit and Loss

- **Profit/Loss** $\underbrace{P/L = P_t + D_t - P_{t-1}}$

P_{t-1} $P_t + D_t$

$t-1$ t

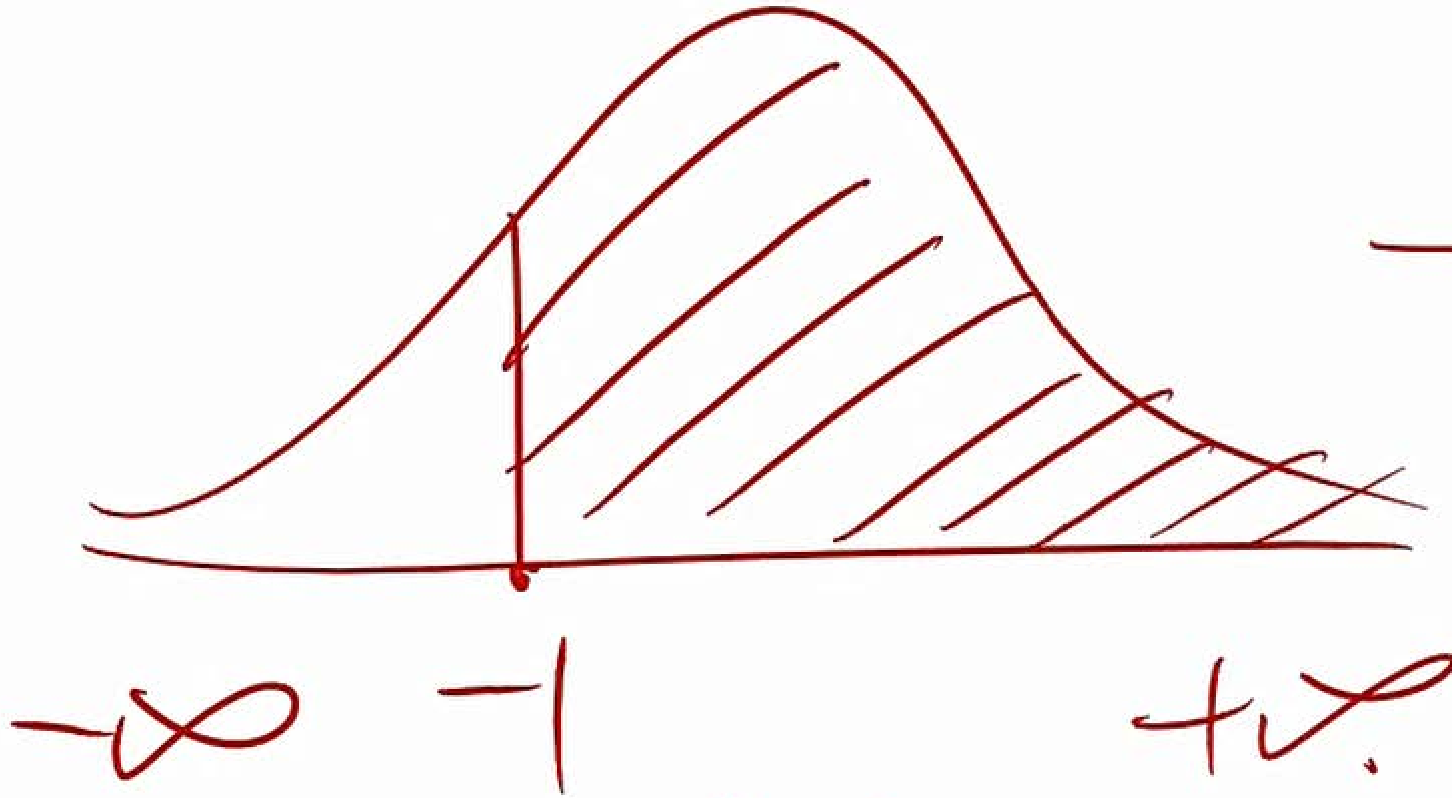
$P_t - P_{t-1}$

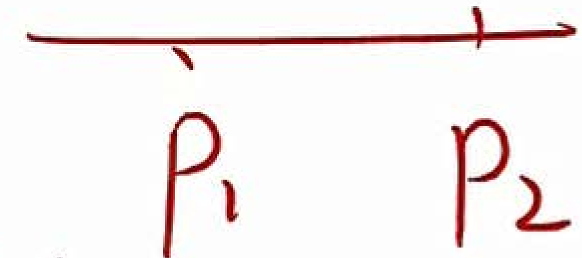
- **Arithmetic Return Data :**

$$r_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \frac{P_t + D_t}{P_{t-1}} - 1$$

- **Geometric Return Data :**

$$R_t = \ln(\frac{P_t + D_t}{P_{t-1}}) = \ln(1 + r_t)$$





$$R = \frac{P_2 - P_1}{P_1} = \frac{P_2}{P_1} - 1 \in (-1, +\infty)$$

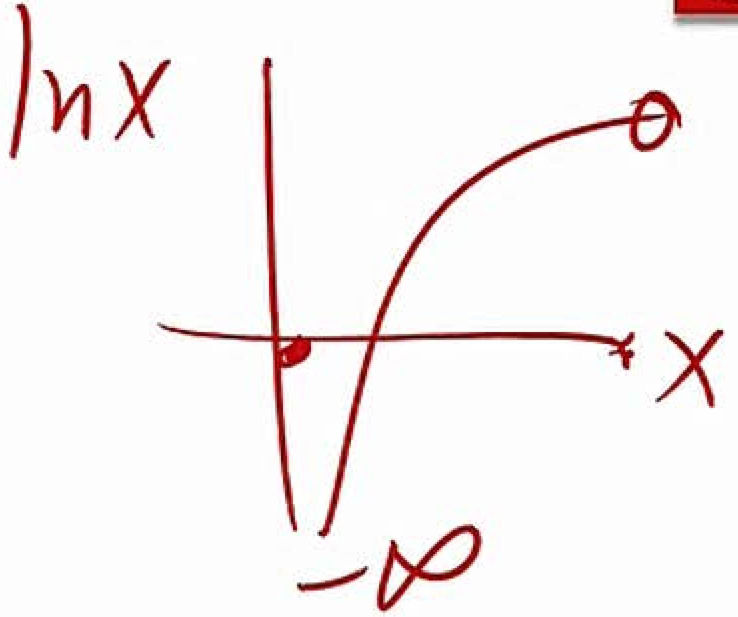
$$R_{FRM} = \ln\left(\frac{P_2}{P_1}\right)$$

$$r = \frac{P_2 - P_1}{P_1} = \frac{P_2}{P_1} - 1$$

$$\boxed{\frac{P_2}{P_1}} = 1 + r$$

$$\lim_{r \to 0} R = \ln(1+r) = r$$

$$R = \ln\left(\frac{p_2}{p_1}\right)$$



$$p_2 \to \infty \qquad \ln \infty = \infty$$

$$p_2 \to 0 \qquad \ln(0) = -\infty$$

GOLDEN FUTURE

Profit and Loss

- **Profit/Loss** $$\underbrace{P/L = P_t + D_t - P_{t-1}}$$

P_{t-1} $\quad$ $P_t + D_t$

$t-1$ $\quad$ t

- **Arithmetic Return Data :**

%

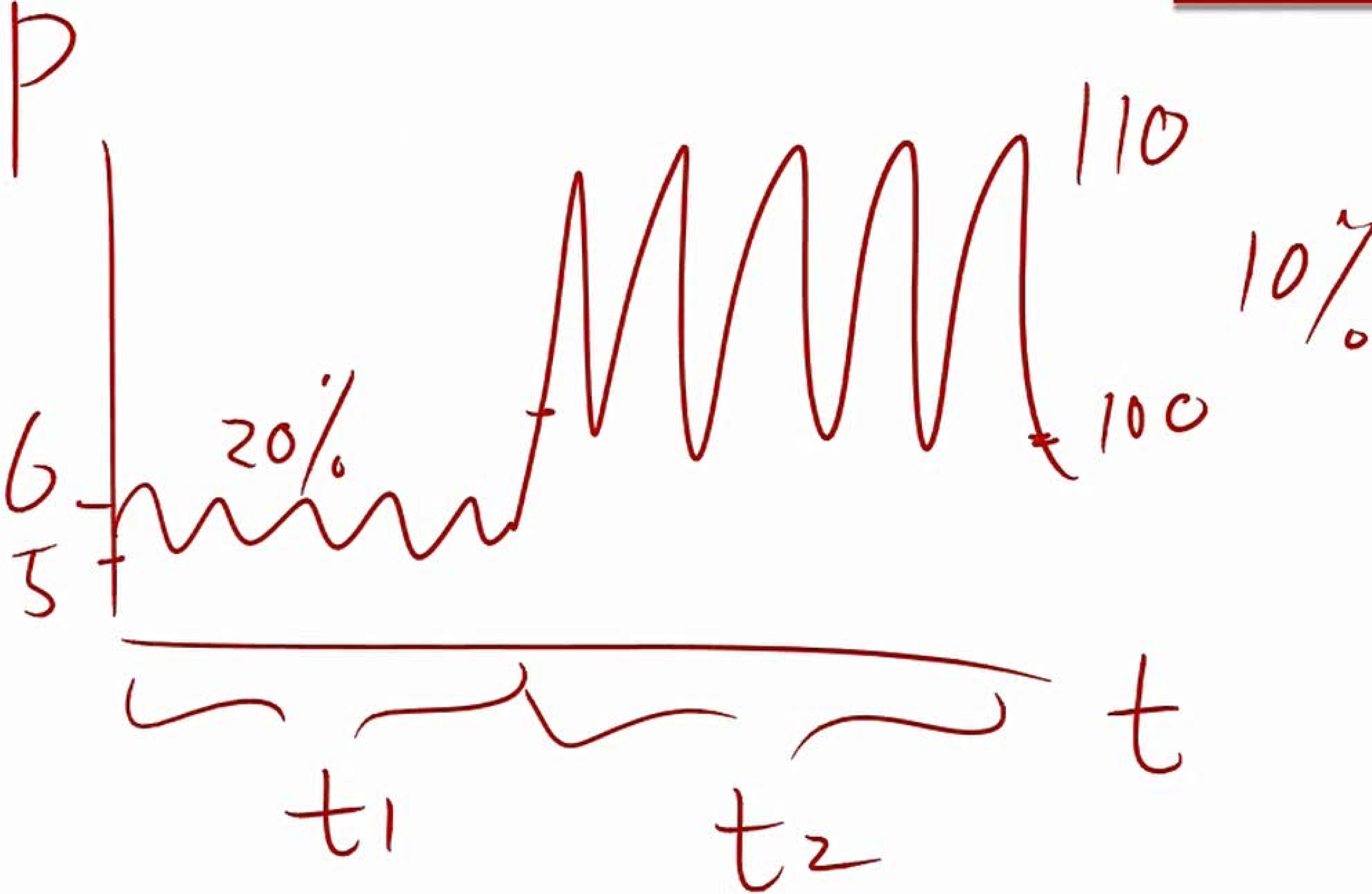
$\frac{P_2 - P_1}{P_1}$

$$r_t = \frac{(P_t + D_t) - P_{t-1}}{P_{t-1}} = \frac{P_t + D_t}{P_{t-1}} - 1$$

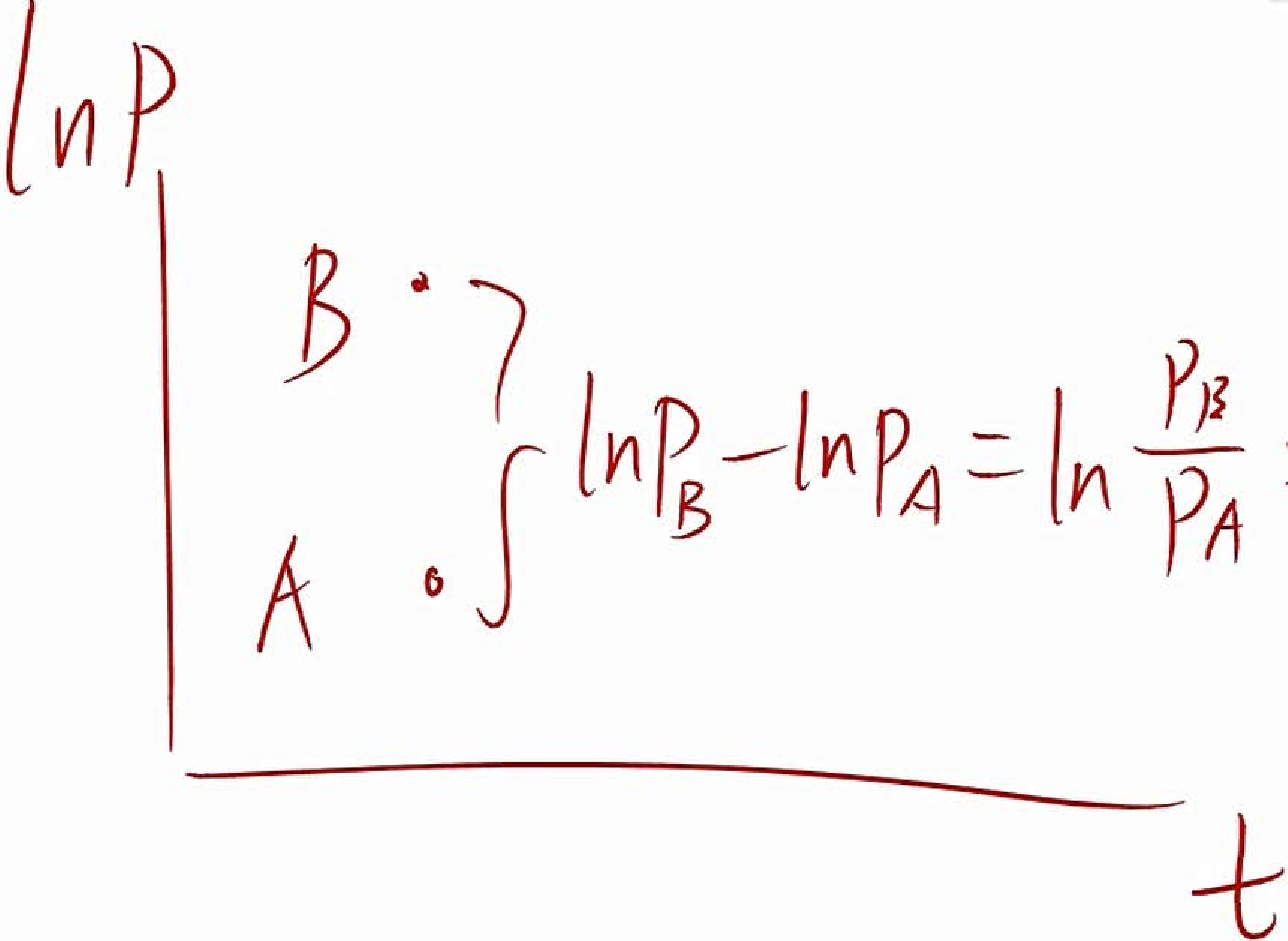
$P_t - P_{t-1}$

- **Geometric Return Data :**

$$R_t = \ln(\frac{P_t + D_t}{P_{t-1}}) = \ln(1 + r_t)$$



P
110
10%
20%
100
6
5
t
t1
t2



$$\ln P_B - \ln P_A = \ln \frac{P_B}{P_A} = R$$

Estimating parametric VaR

➢ **The parametric approach explicitly assumes a distribution for the underlying observations:**

- **Approach 1**: Suppose that we wish to estimate VaR under the assumption that P/L is normally distributed.
- **Approach 2:** Suppose that we wish to estimate VaR under the assumption that P/L is lognormal distributed.

VaR(1天, 95%) = 1万.

1天内, 95% loss ≤ 1万; 5% L > 1万

100天内, 5天 loss > 1万

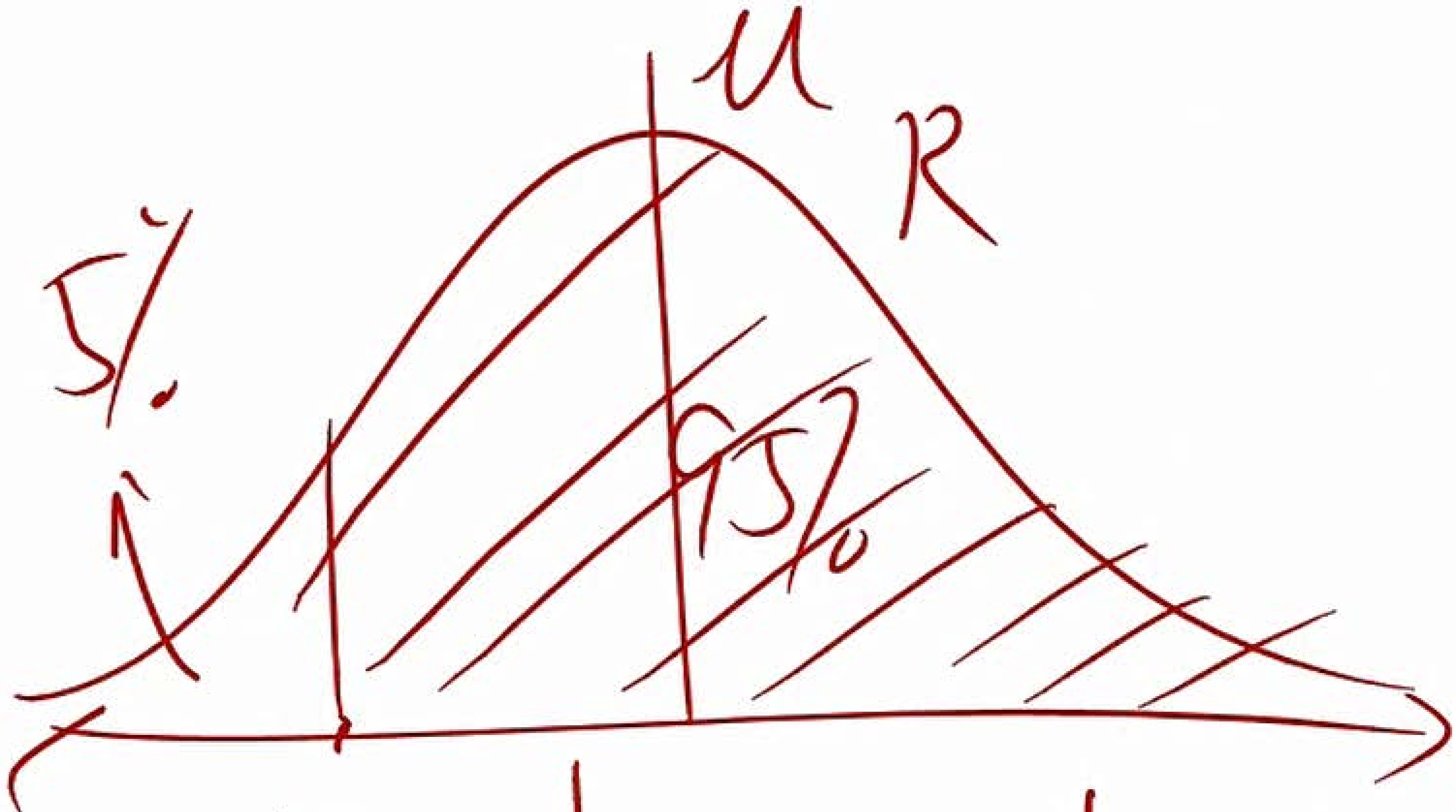
Parametric VaR

μ, σ, ρ

distribution

N

$\ln N$



loss $\quad VaR = |\mu - 1.65\sigma| \quad$ gain

95% VaR

99% VaR

%

$\mu - 1.65\sigma$

$\mu - 2.33\sigma$.

% %

95% VaR	99% VaR

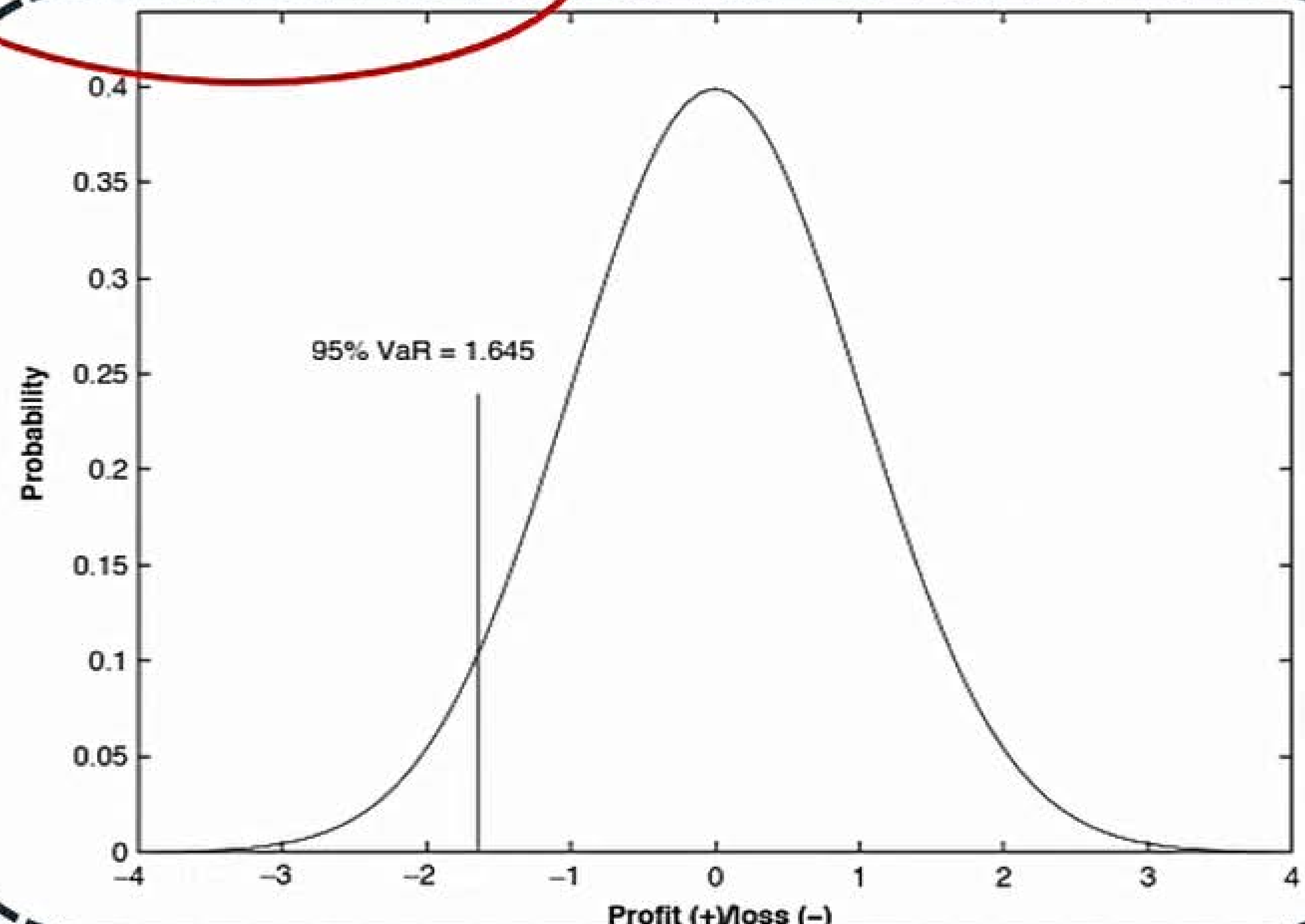
	95% VaR	99% VaR
%	$\mu - 1.65\sigma$ % %	$\mu - 2.33\sigma$
\$	$\lvert \mu - 1.65\sigma \rvert \times P_0$ \$	$\lvert \mu - 2.33\sigma \rvert \cdot P_0$

Estimating VaR with Normally Distributed P/L

➢ **Approach 1: Normal VaR**

- We assume that arithmetic returns are normally distributed with mean μ and standard deviation σ

$$VaR = -(\mu - z_\alpha \sigma) \qquad VaR = -(\mu - z_\alpha \sigma)P_{t-1}$$



Estimating VaR with Normally Distributed P/L

➢ **Example:**

- Assume that the profit/loss distribution for XYZ is normally distributed with an annual mean of \$16million and a standard deviation of \$11 million. Calculate the VaR at the 95% and 99% confidence levels using a parametric approach.

σ μ

$16-1.65\times 11$

$$\begin{aligned} VaR(5\%) &= -\$16\text{million} + \$11\text{million} \times 1.65 \\ &= \$2.15 \text{ million} \end{aligned}$$

$$\begin{aligned} VaR(1\%) &= -\$16\text{million} + \$11\text{million} \times 2.33 \\ &= \$9.63 \text{ million} \end{aligned}$$

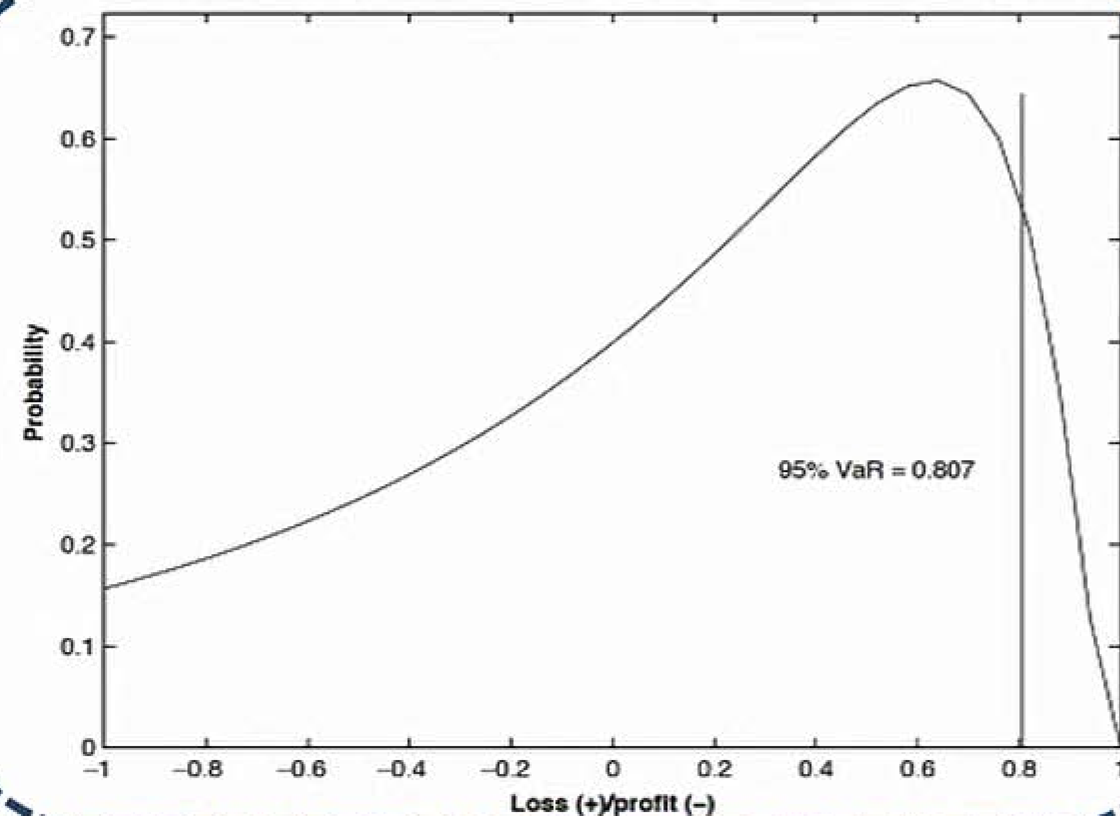
Estimating Lognormal VaR

Lognormal VaR

- Assume that geometric returns are normally distributed with mean μ and standard deviation σ. This assumption implies that the natural logarithm of p_t is normally distributed, or that p_t itself is lognormally distributed. Normally distributed geometric returns imply that the VaR is lognormally distributed.

$$VaR = 1 - e^{\mu - z_\alpha \sigma}$$

$$VaR = (1 - e^{\mu - z_\alpha \sigma})P_{t-1}$$



Estimating Lognormal VaR

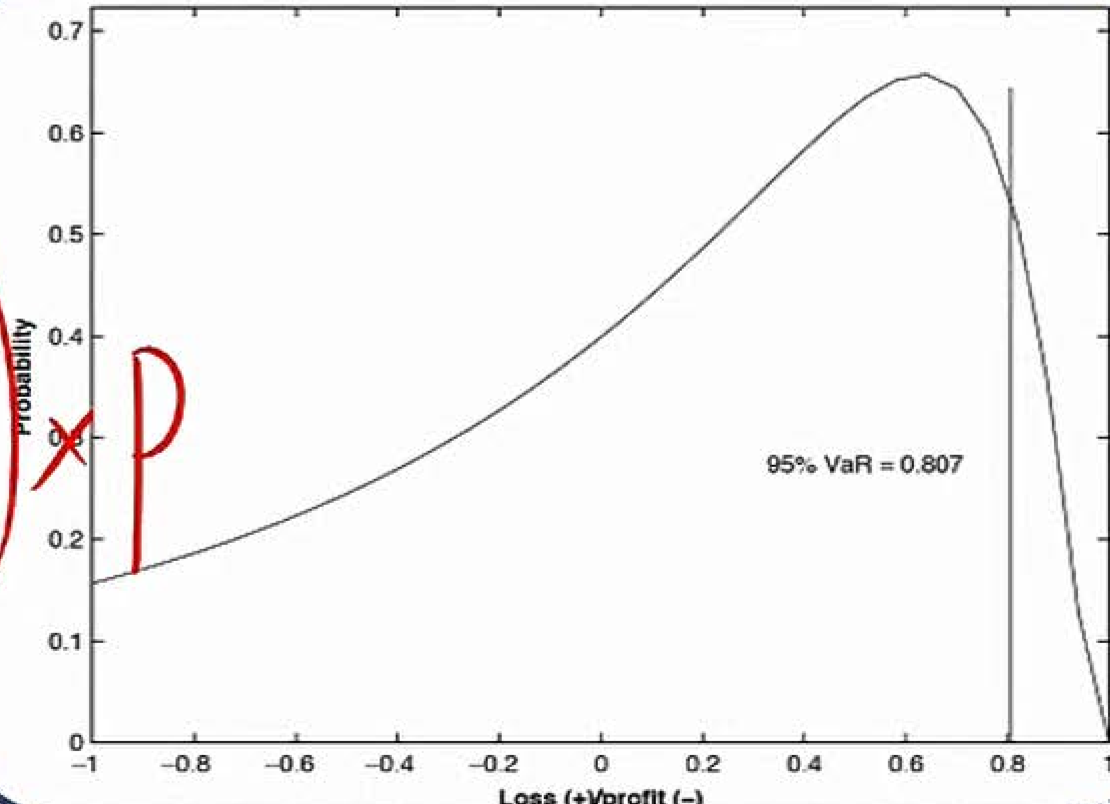
Lognormal VaR

- Assume that geometric returns are normally distributed with mean μ and standard deviation σ. This assumption implies that the natural logarithm of p_t is normally distributed, or that p_t itself is lognormally distributed. Normally distributed geometric returns imply that the VaR is lognormally distributed.

$$VaR = 1 - e^{\mu - z_\alpha \sigma}$$

$$VaR = (1 - e^{\mu - z_\alpha \sigma})P_{t-1}$$

$\ln VaR = (1 - e^{u - 1.65\sigma}) \times P$



Estimating Lognormal VaR

➢ **Example:**

μ σ

- A diversified portfolio exhibits a normally distributed geometric return with mean and standard deviation of 11% and 21%, respectively. Calculate the 5% and 1% lognormal VaR assuming the beginning period portfolio value is $100.

Lognormal VaR(5%) $= 100 \times (1 - e^{0.11-0.21\times1.65}) = \$21.06$

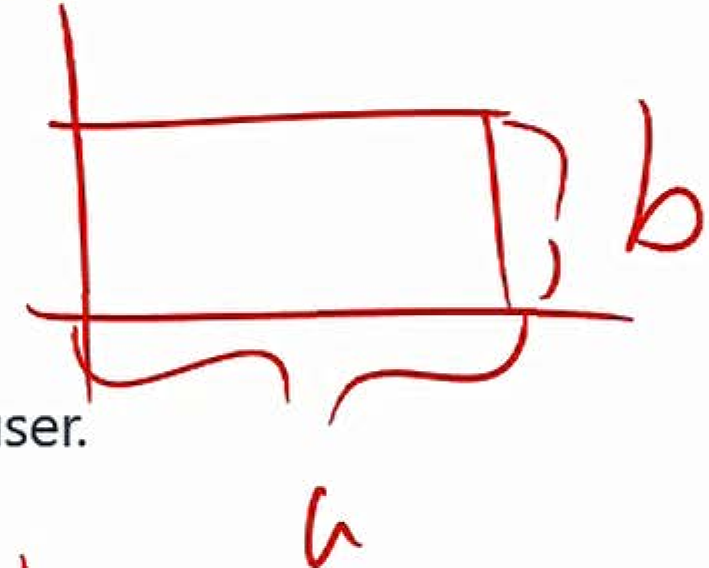
Lognormal VaR(1%) $= 100 \times (1 - e^{0.11-0.21\times2.33]}) = \$31.57$

$$\left(1-e^{11\%-1.65\times 21\%}\right)\times 100$$

Estimating Coherent Risk Measures

- **A coherent risk measure is a weighted average of the quantiles of our loss distribution.**

$$M_{\emptyset} = \int_0^1 \emptyset(\mathrm{p})\, q_p \, d\mathrm{p}$$

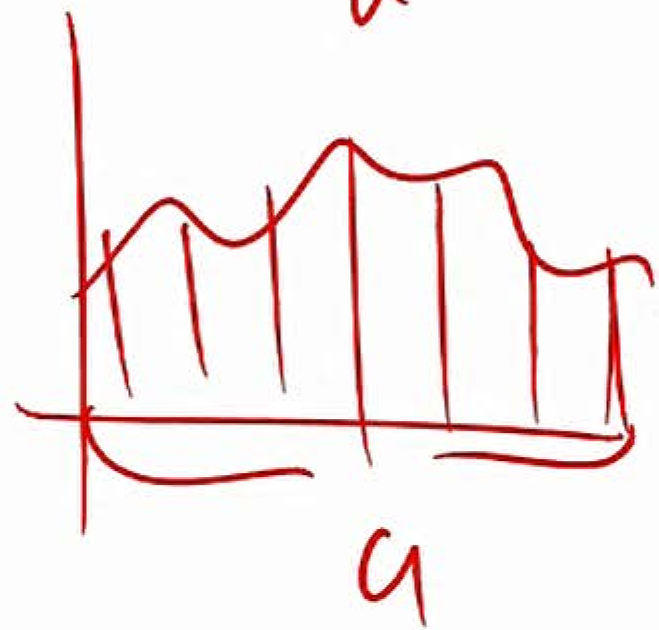


- $\Phi(\mathrm{p})$ = weighing function specified by the user.

- Exponential Weighting Function

$$\emptyset_{\gamma}(\mathrm{p}) = \frac{e^{-(1-\mathrm{p})/\gamma}}{\gamma(1 - e^{-1/\gamma})}$$

✓ γ : the degree of our risk-aversion



$n = 100$ 95% VaR

Loss $\longrightarrow$ Gain

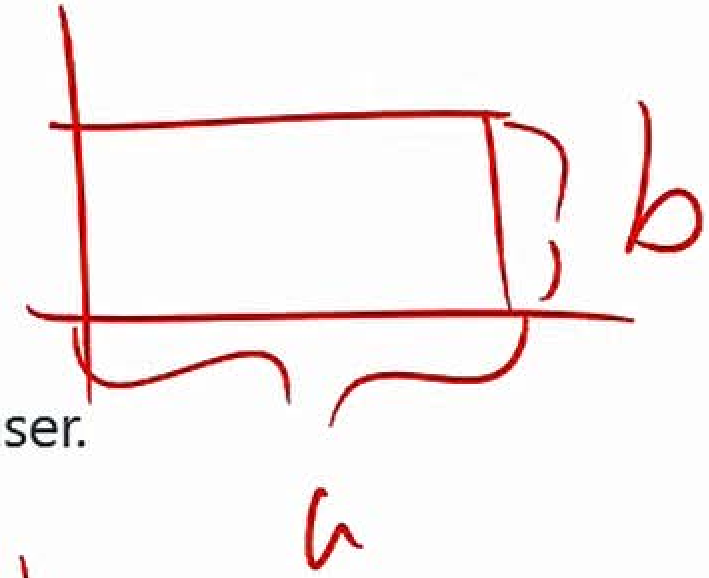
$-17 \quad -15 \quad -10 \quad -9 \quad \underline{\underline{-8}} \quad \cdots \quad 0 \quad +1 \quad +3$

VaR

Estimating Coherent Risk Measures

➢ **A coherent risk measure is a weighted average of the quantiles of our loss distribution.**

$$M_{\emptyset} = \int_0^1 \emptyset(\mathrm{p}) q_p \, d\mathrm{p}$$

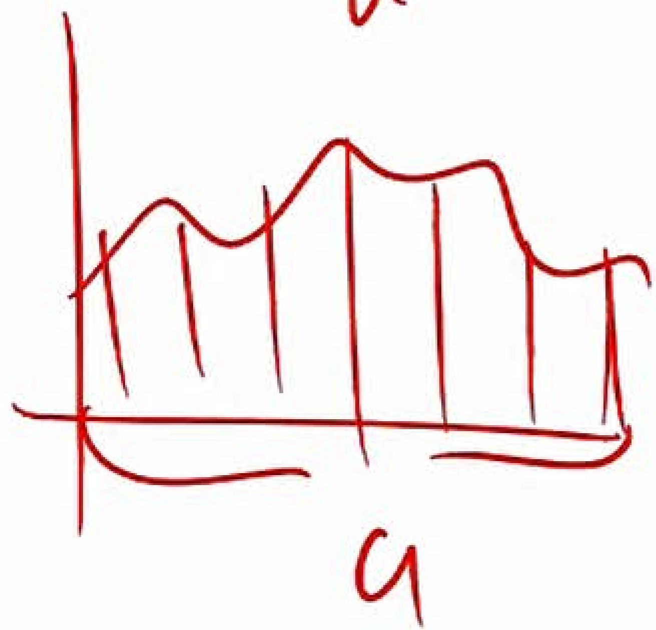


- $\Phi(\mathrm{p})$ = weighing function specified by the user.

➢ Exponential Weighting Function

$$\emptyset_\gamma(\mathrm{p}) = \frac{e^{-(1-\mathrm{p})/\gamma}}{\gamma(1-e^{-1/\gamma})}$$

✓ γ : the degree of our risk-aversion



Estimating Coherent Risk Measures

➢ **Estimating exponential spectral risk measures as a weighted average of VaRs (γ=0.05)**

Li Wi

Confidence level (α)	αVaR	Weight Φ(α)	Φ(α)×αVaR
10%	-1.2816	0	0.0000
20%	-0.8416	0	0.0000
30%	-0.5244	0	0.0000
40%	-0.2533	0.0001	0.0000
50%	0	0.0009	0.0000
60%	0.2533	0.0067	0.0017
70%	0.5244	0.0496	0.0260
80%	0.8416	0.3663	0.3083
90%	1.2816	2.7067	3.4689
Risk measure = mean (Φ(α) times αVaR)			0.4226

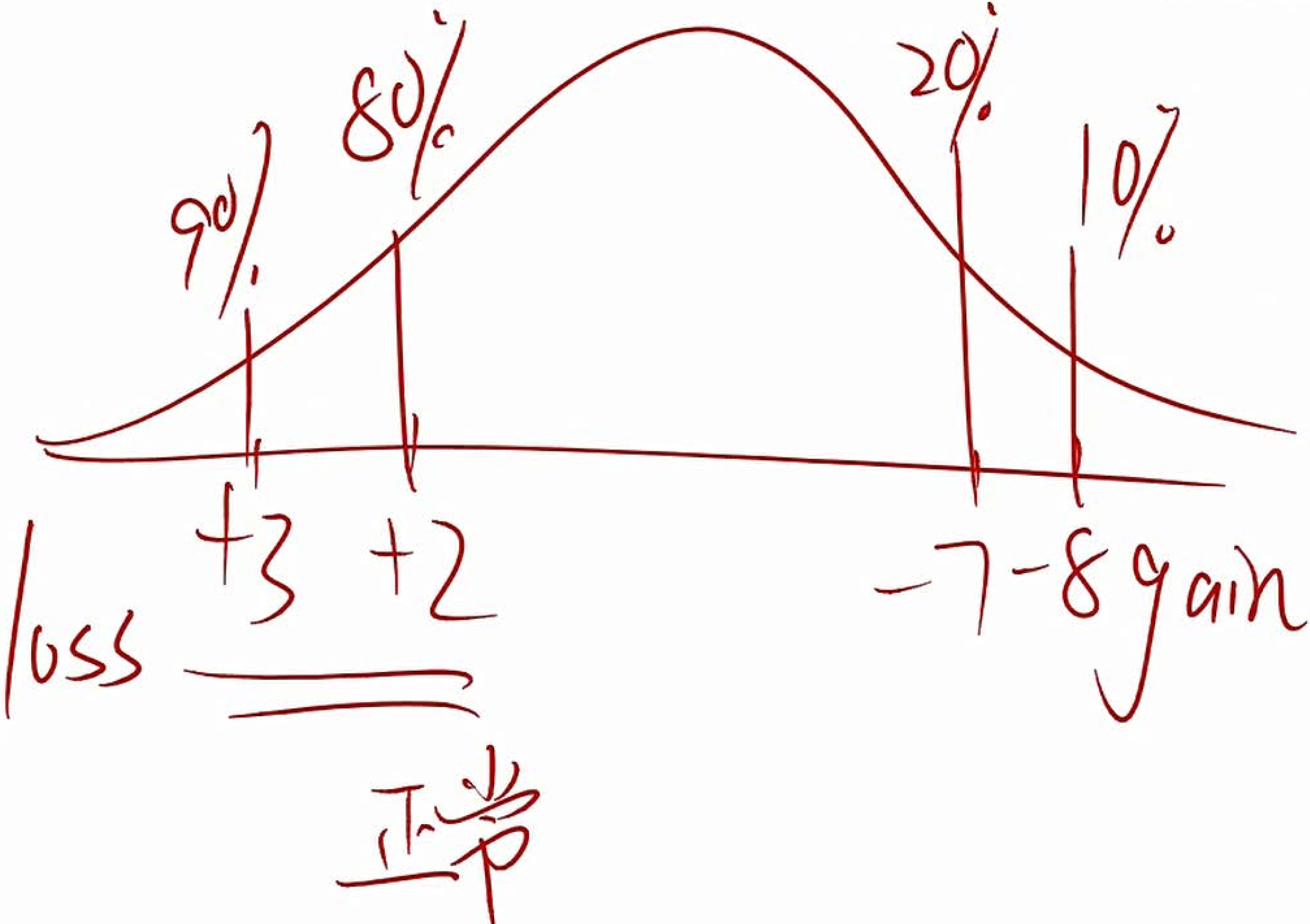
gain

loss

Li×Wi

Li

sum & average


90%
80%
20%
10%
loss
+3
+2
-7
-8
gain
正常

Estimating Coherent Risk Measures

- The estimate **does eventually converge** to the true value as n gets large.

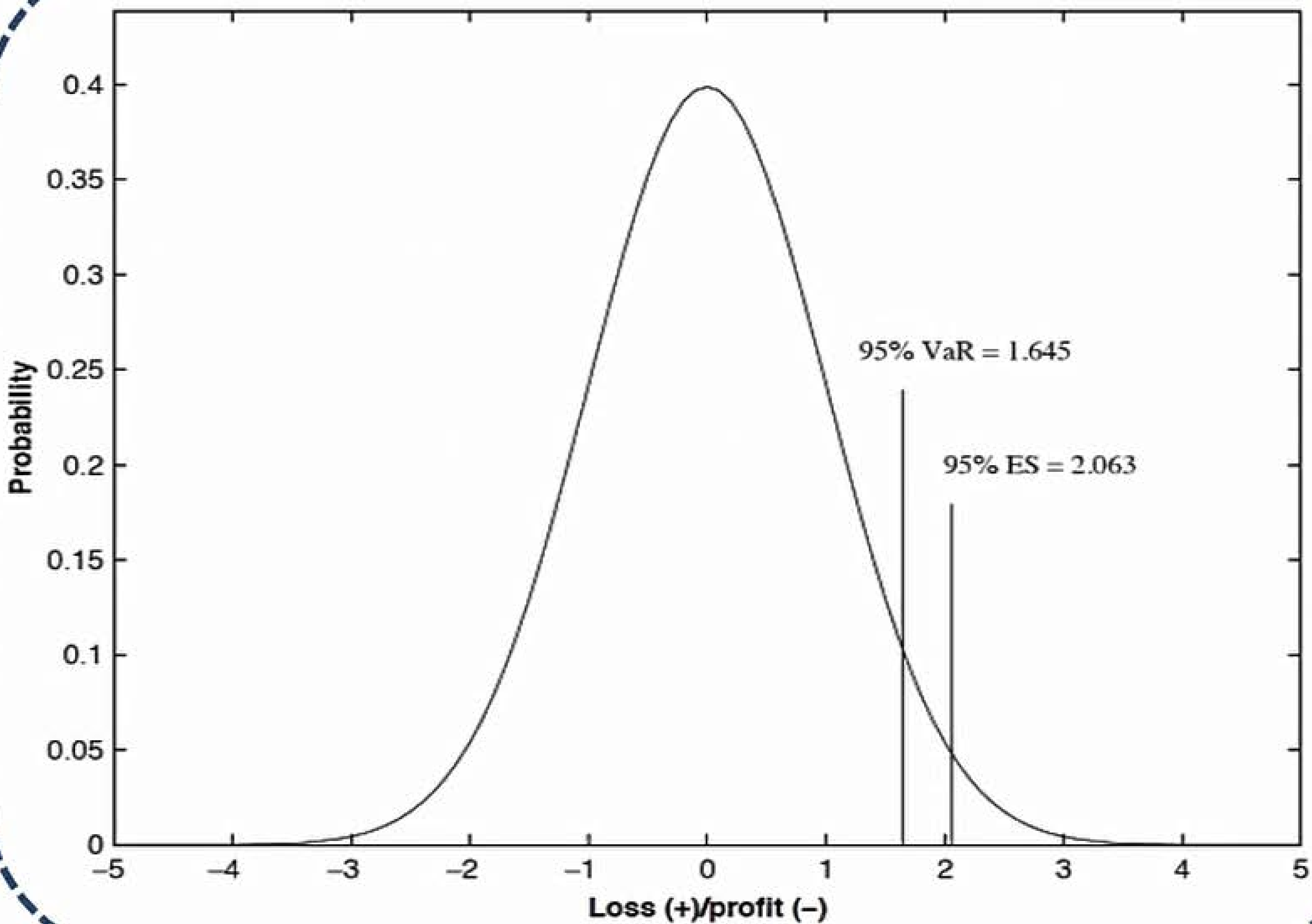
Estimates of exponential spectral coherent risk measure as a function of the number of tail slices	
Number of tail slices	Estimate of exponential spectral risk measure
10	0.4227
50	1.3739
100	1.5853
500	1.7896
1000	1.8197
5000	1.8461
10,000	1.8498
50,000	1.8529
100,000	1.8533
500,000	1.8536

GOLDEN FUTURE

Estimating Expected Shortfall

- **The Conditional VaR (expected shortfall)**
  - The expected value of the loss when it exceeds VaR.
  - Measures the average of the loss conditional on the fact that it is greater than VaR.
  - CVaR indicates the potential loss if the portfolio is "hit" beyond VaR. Because CVaR is an average of the tail loss, one can show that it qualifies as a subadditive risk measure.

Estimating Expected Shortfall

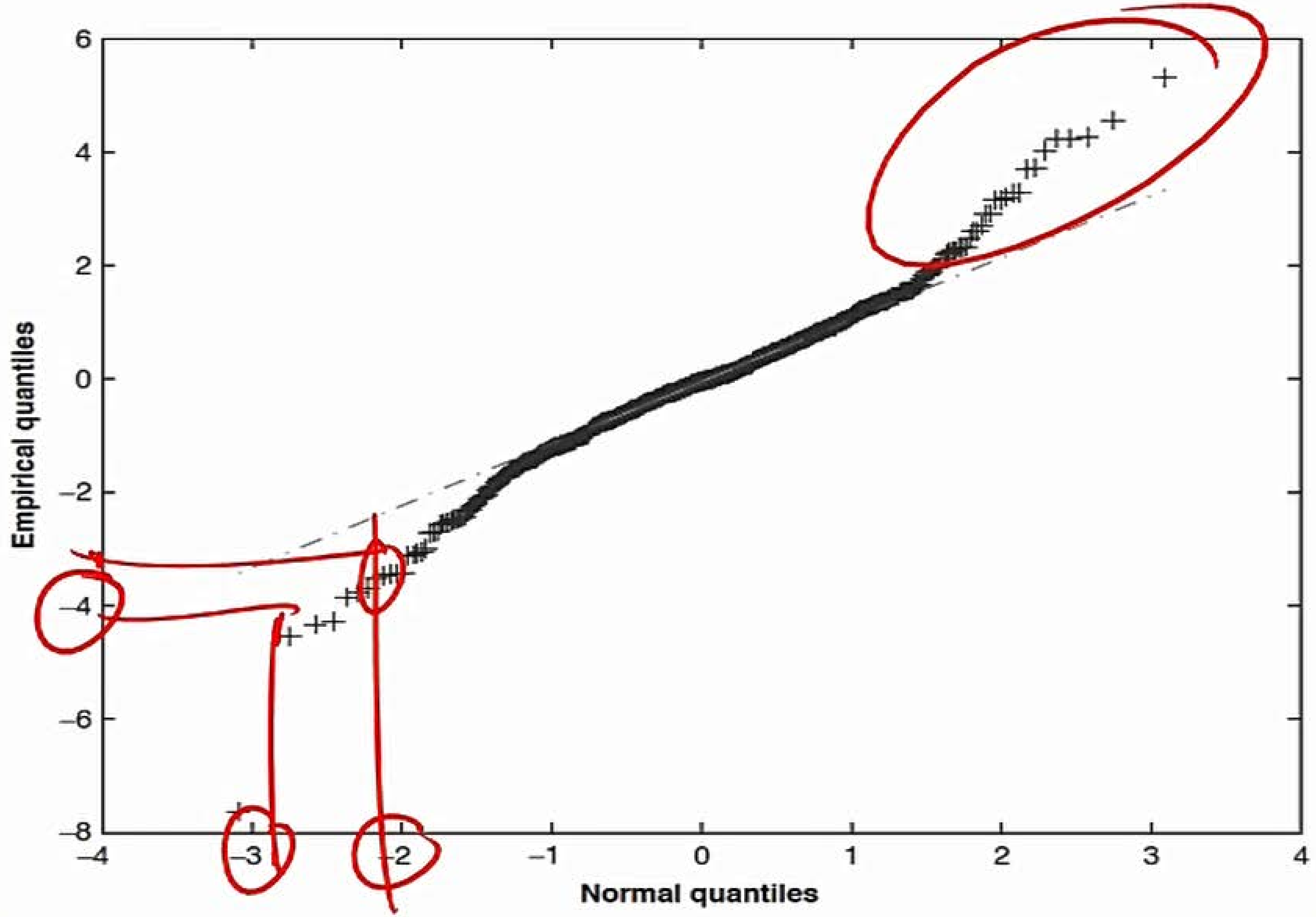


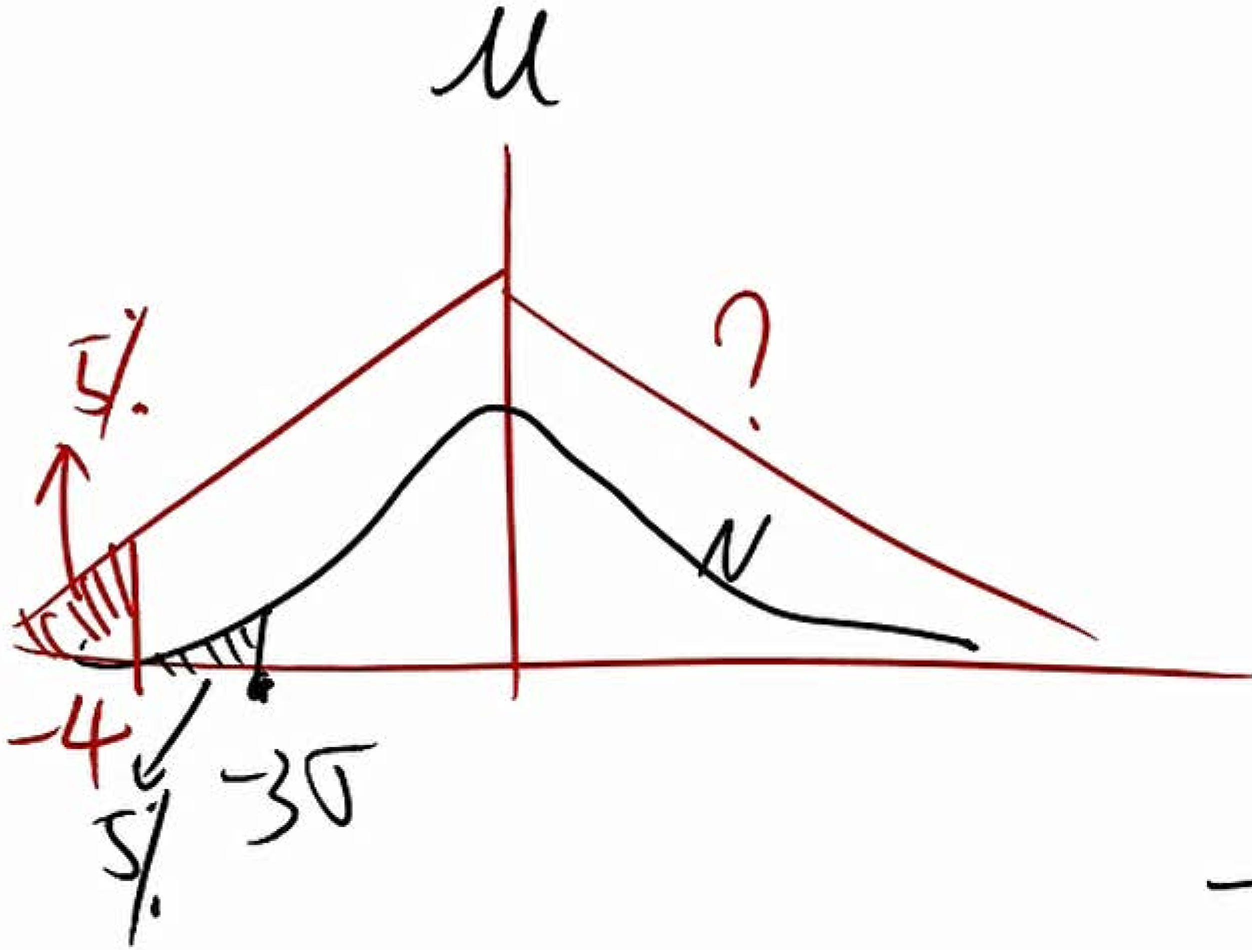
loss

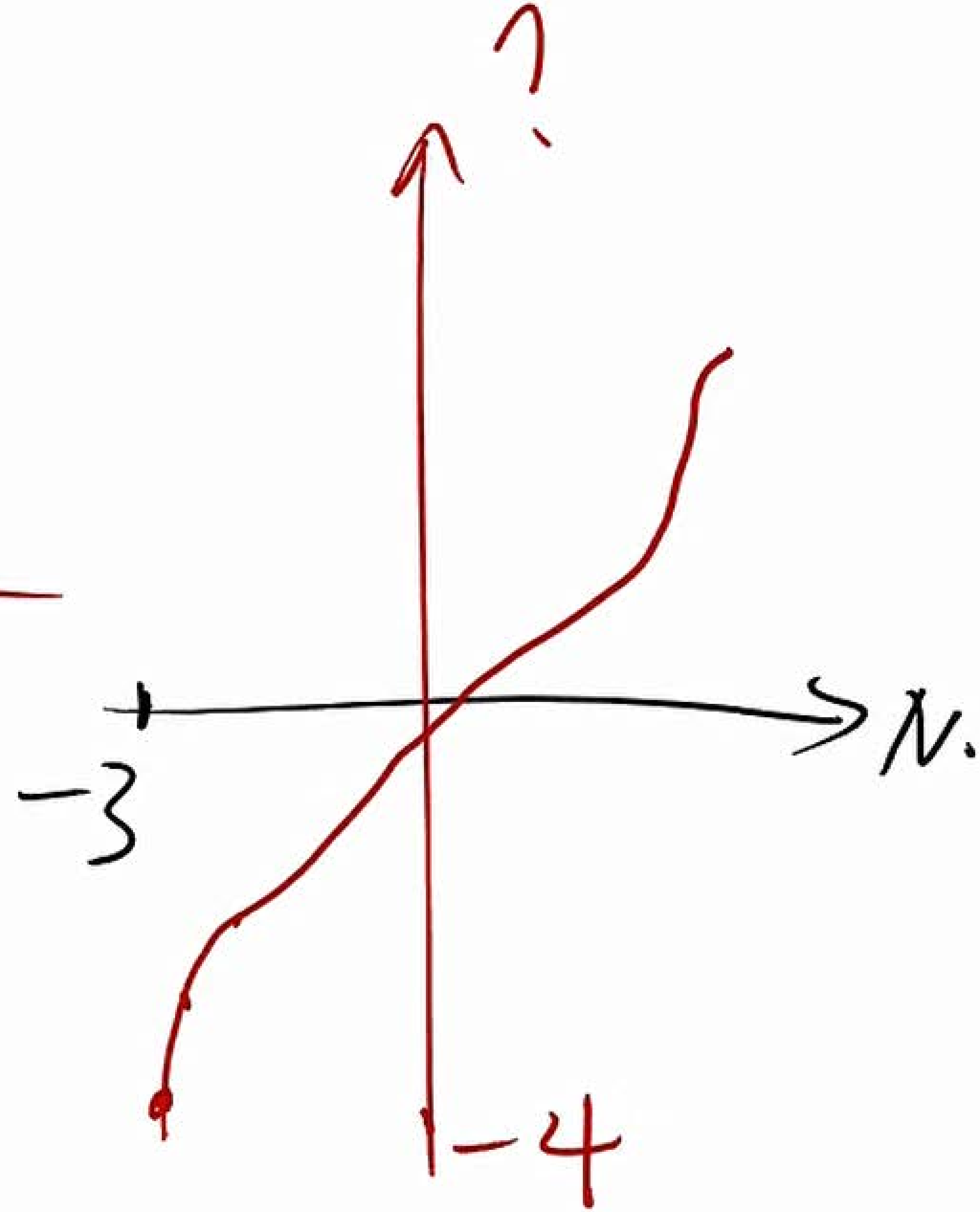
Quantile-Quantile Plots

- **We are interested in asking:**
  - If data looks right when we use parametric approach?
  - What we do is
    - ✓ Plot our data on a histogram and estimate the relevant summary statistics.
    - ✓ Consider what kind of distribution might fit our data.
- A plot of the quantiles of the empirical distribution against those of some specified distribution. The shape of the QQ plot tells us a lot about how the empirical distribution compares to the specified one.
- In particular, if the QQ plot is linear, then the specified distribution fits the data, and we have identified the distribution to which our data belong.

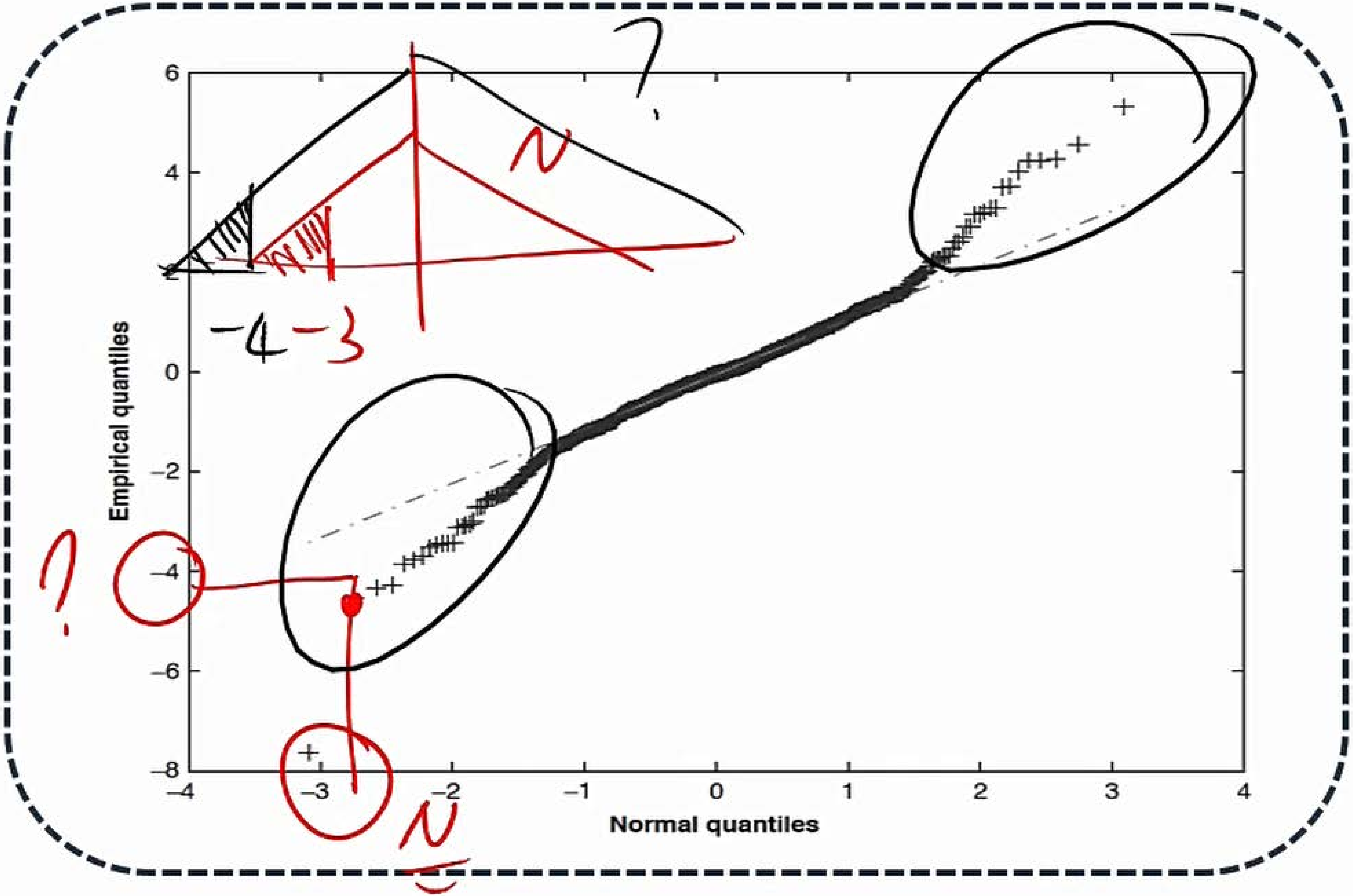
Quantile-Quantile Plots




μ
5%
?
N
-4
5%
-3σ


?
N.
-3
-4

Quantile-Quantile Plots



GOLDEN FUTURE

Estimating Expected Shortfall

15

VaR

- **Example:**
  - Given the following 30 ordered percentage returns of an asset:

    −16, −14, $\boxed{-10}$, $\boxed{-7}$, −7, −5, −4, −4, −4, −3, −1, −1, 0, 0, 0, 1, 2, 2, 4, 6, 7, 8, 9, 11, 12, 12, 14, 18, 21, 23.

    Calculate the VaR and expected shortfall at a 90% confidence level:

  - Solution:

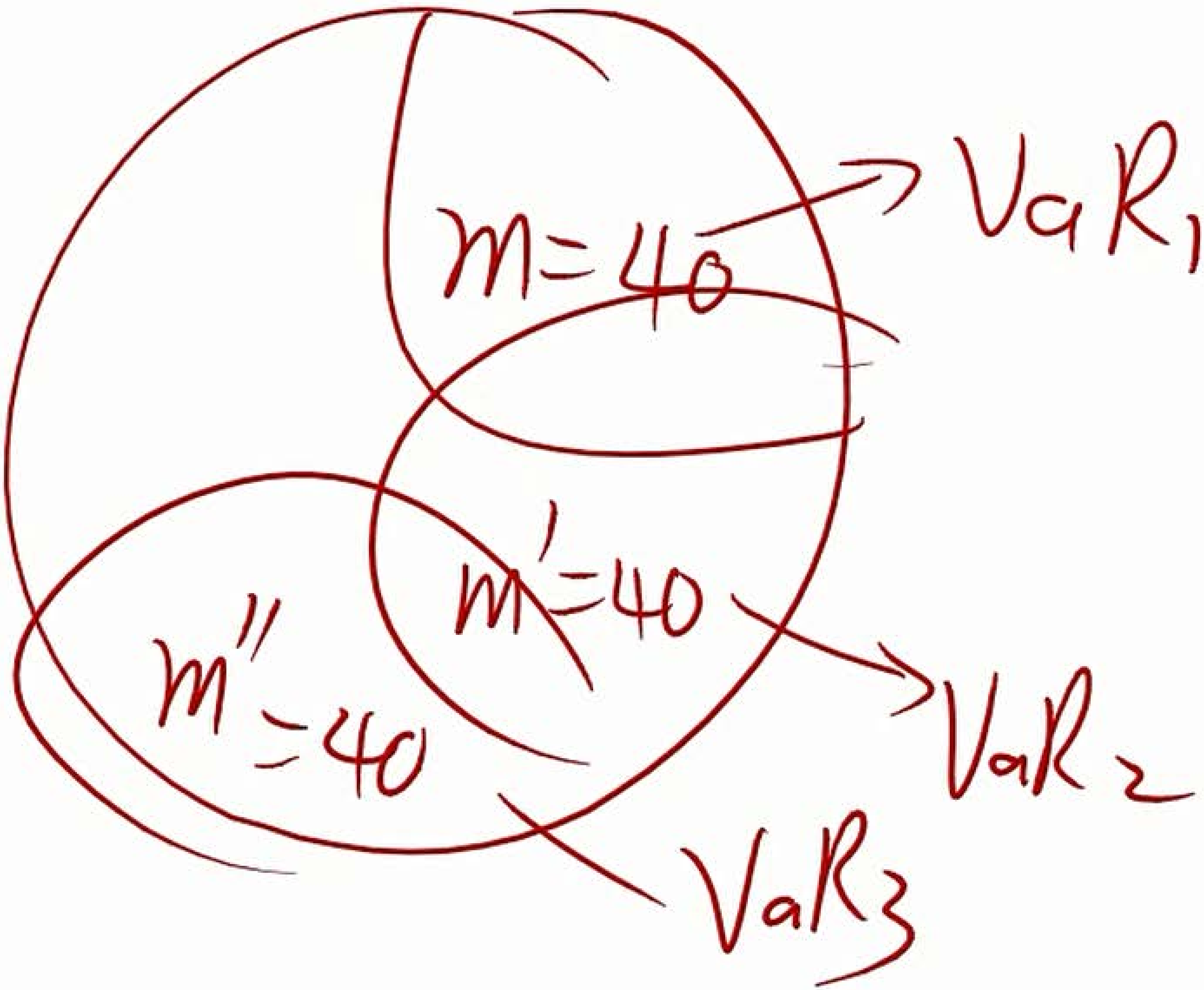
    VaR (90%) = 7, Expected Shortfall = 13.3

$30 \times 10\% = 3$

Bootstrapped Historical Simulation

- **The bootstrap is very intuitive and easy to apply.**
  - We create a large number of new samples, each observation of which is obtained by drawing at random from our original sample and replacing the observation after it has been drawn.
  - Each new 'resampled' sample gives us a new VaR estimate, and we can take our 'best' estimate to be the mean of these resample-based estimates. The same approach can also be used to produce resample-based ES estimates—each one of which would be the average of the losses in each resample exceeding the resample VaR—and our 'best' ES estimate would be the mean of these estimates.
- A bootstrapped estimate will often **be more accurate** than a 'raw' sample estimate, and bootstraps are also useful for gauging the precision of our estimates.

$n = 100$ 95% VaR



sample
put back
re-sample

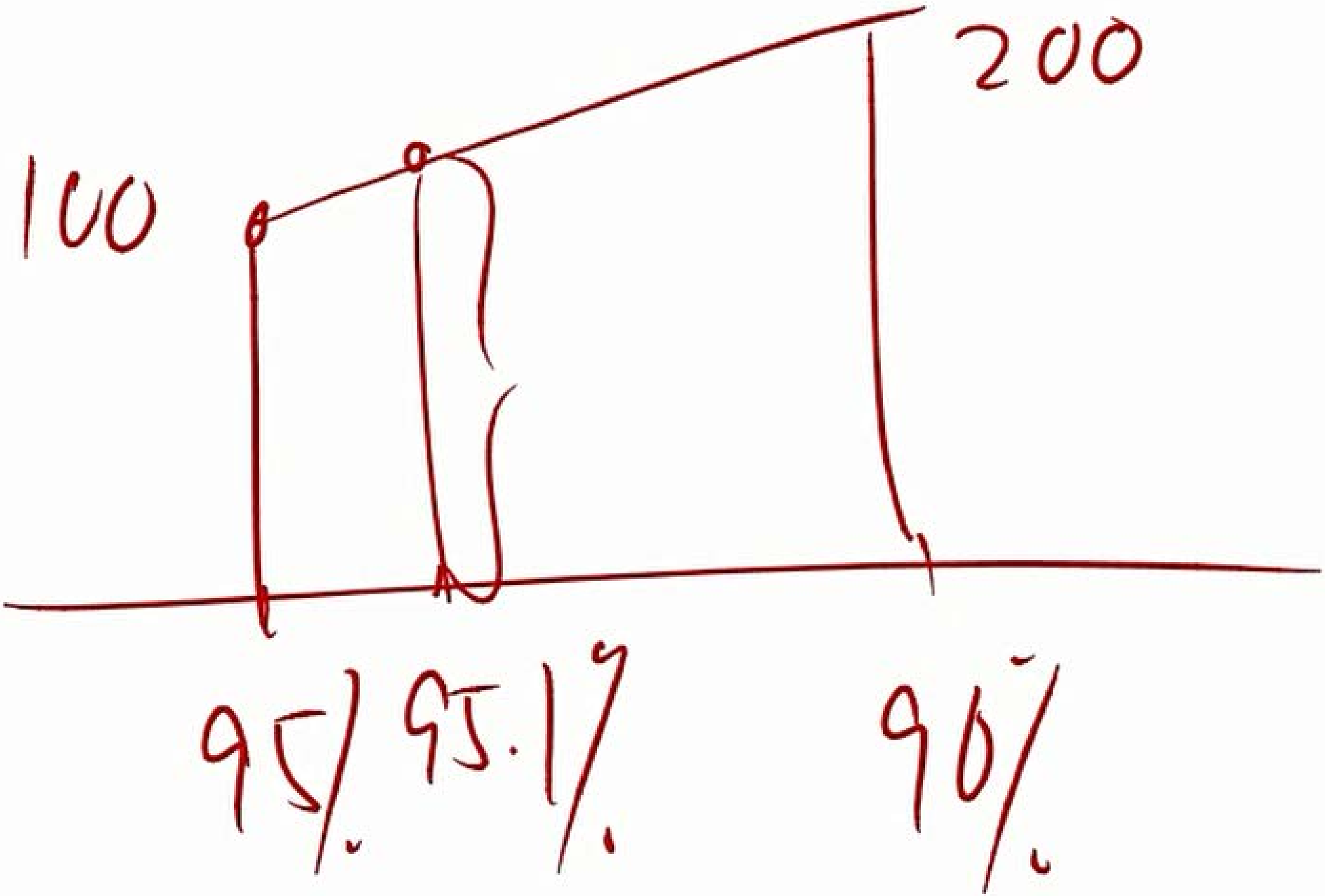
Drawbacks of HS

- Basic HS has the practical drawback that it only allows us to estimate VaRs at discrete confidence intervals determined by the size of our data set.
  - For instance, the VaR at the 95.1% confidence level is a problem because there is no corresponding loss observation to go with it.
  - With n observations, basic HS only allows us to estimate the VaRs associated with, at best, n different confidence levels.

n=100.

4.9%

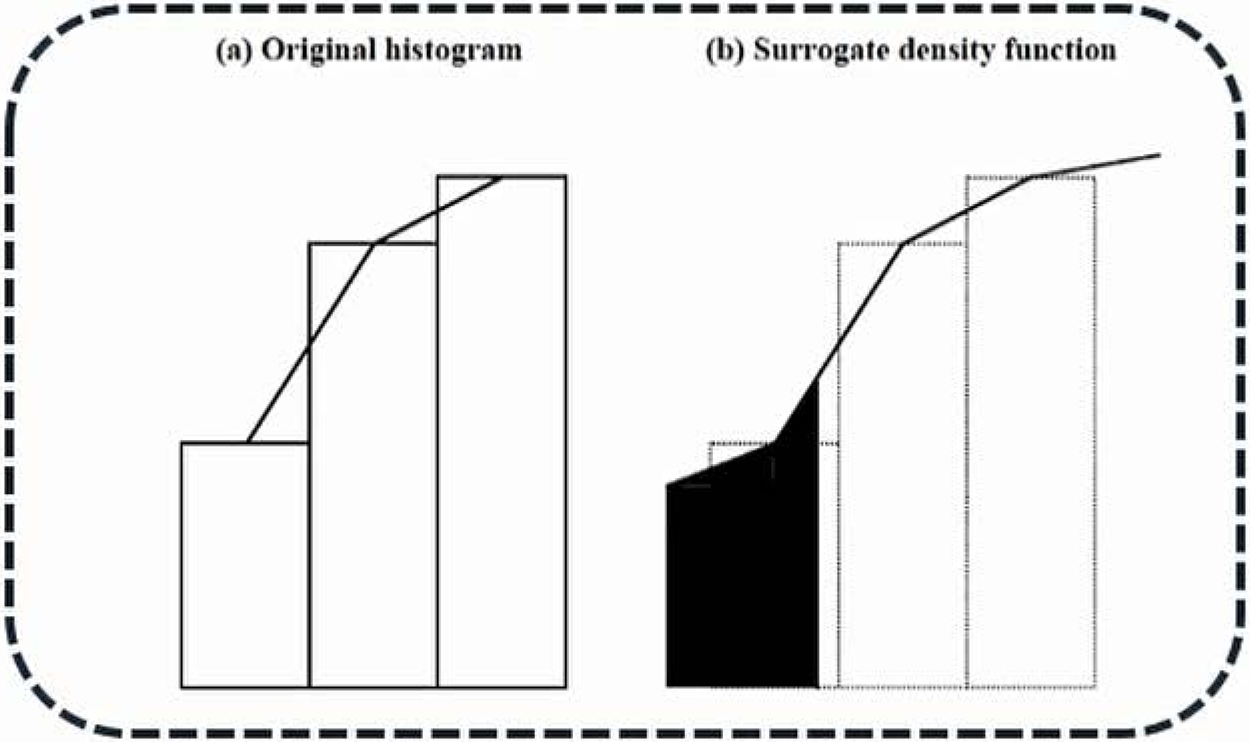
95% 90%



100
200
95%
95.1%
90%

Non-parametric Density Estimation

- Non-parametric density estimation offers a potential solution.
  - Draw in straight lines connecting the mid-points at the top of each histogram bar(Polygon).
  - Treating the area under the lines as a pdf then enables us to estimate VaRs at any confidence level.

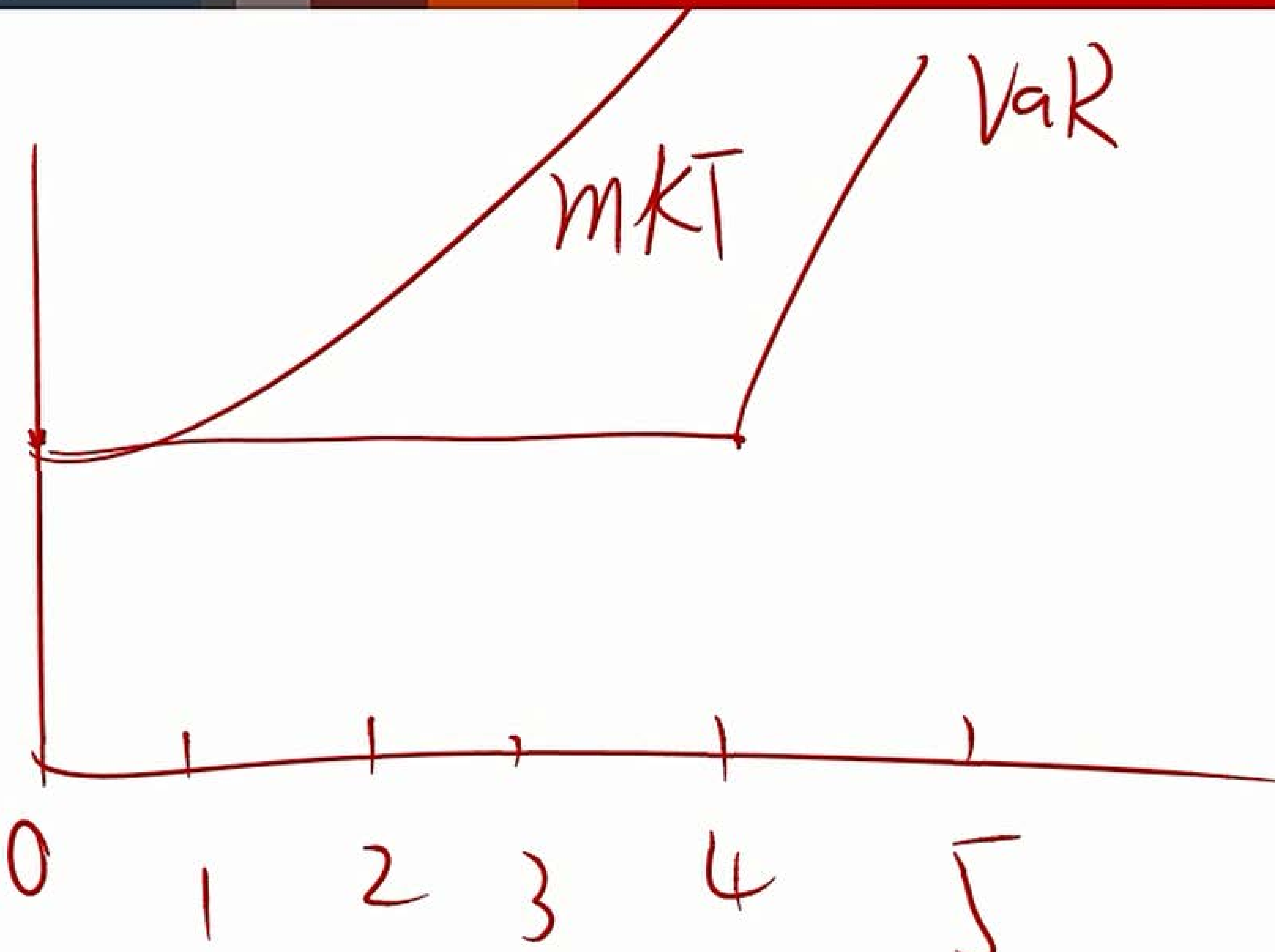


Blames for Traditional HS

Ghost Effect

- **Short survivorship:**
  - One return observation will affect each of the next n observations in our P/L series.
- **Discrete and not continues:**
  - But after n periods have passed, the observation will fall out of the data set used to calculate the current HS P/L series, and will thereafter have no effect on P/L.
  - Any observation is given the same weight on P/L provided it is less than n periods old, and no weight if it is older than that.
  - It is also hard to justify why an observation should have a weight that suddenly goes to zero when it reaches age n. Why is it that an observation of age n−1 is regarded as having a lot of value?

									VaR
0	5	5	5	5	5	4	3	2	5
1	20	5	5	5	5	5			5
2	30	20	5	5	5				5
3	40	30	20	5	5				5
4	50	40	30	20	5				5
5	60	50	40	30	20				20


MKT
VaR
0
1
2
3
4
5

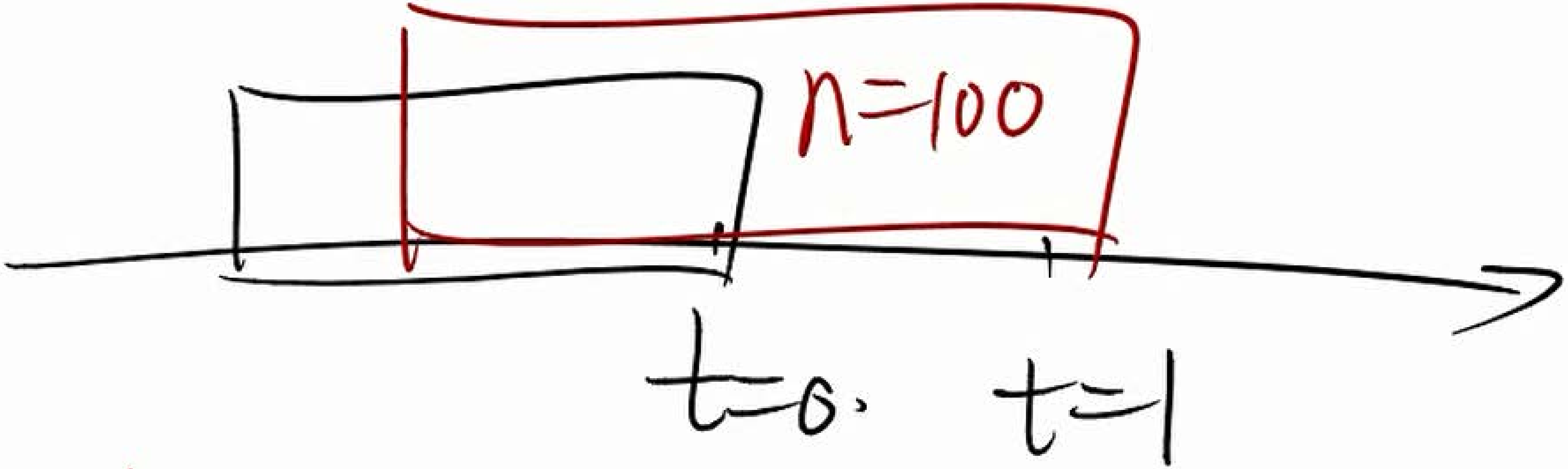
										ES	VaR
0	5	5	5	5	5	4	3	2		5	5
1	20	5	5	5	5	5					5
2	30	20	5	5	5						5
3	40	30	20	5	5						5
4	50	40	30	20	5						5
5	60	50	40	30	20						20

Blames for Traditional HS

- **A new loss appears slowly and suddenly**
  - The increase in risk would only show up later in VaR estimates if the stock market continued to fall in subsequent days.
  - It seems that risk had suddenly increased, and yet that increase in risk would be missed by most HS VaR estimates.
- **Solution:**
  - A stock market crash might have no effect on VaRs except at a very high confidence level.
  - The increase in risk would show up in **ES estimates** just after the first shock occurred.

Blames for Traditional HS

- **A new loss appears slowly and suddenly**
  - The increase in risk would only show up later in VaR estimates if the stock market continued to fall in subsequent days.
  - It seems that risk had suddenly increased, and yet that increase in risk would be missed by most HS VaR estimates.
- **Solution:**
  - A stock market crash might have no effect on VaRs except at a very high confidence level. 95% → 99%
  - The increase in risk would show up in **ES estimates** just after the first shock occurred.



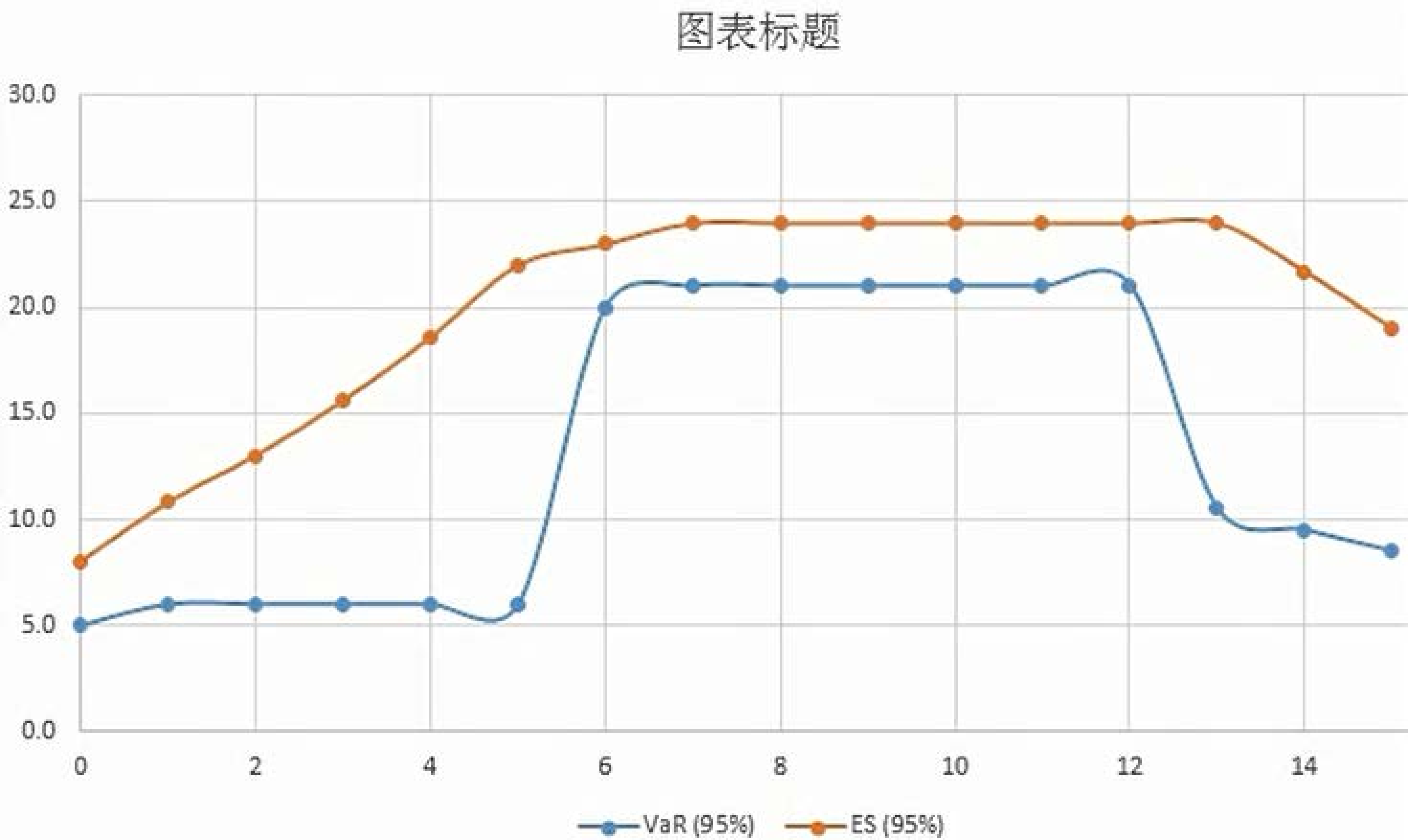
loss=100 ⟵ VaR VaR'=70

age = 100天前

Blames for Traditional HS

Ghost Effect

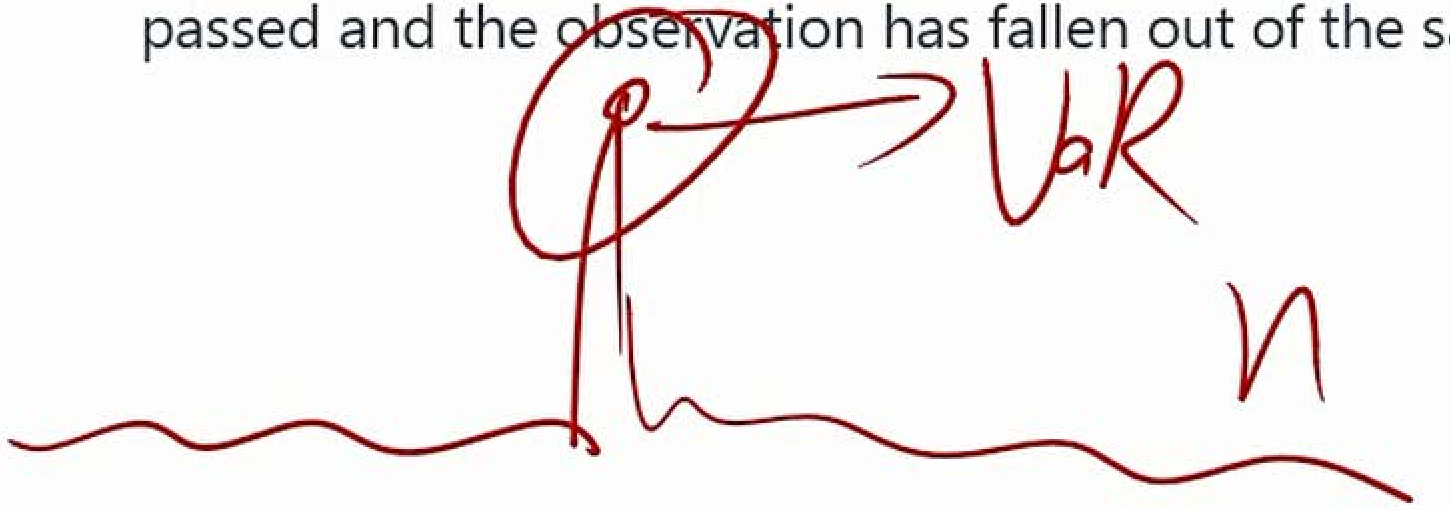
- **Short survivorship:**
  - One return observation will affect each of the next n observations in our P/L series.
- **Discrete and not continues:**
  - But after n periods have passed, the observation will fall out of the data set used to calculate the current HS P/L series, and will thereafter have no effect on P/L.
  - Any observation is given the same weight on P/L provided it is less than n periods old, and no weight if it is older than that.
  - It is also hard to justify why an observation should have a weight that suddenly goes to zero when it reaches age n. Why is it that an observation of age n−1 is regarded as having a lot of value?



图表标题
30.0
25.0
20.0
15.0
10.0
5.0
0.0
0
2
4
6
8
10
12
14
VaR (95%)
ES (95%)

Blames for Traditional HS

- **Ghost effects**
  - we can have a VaR that is unduly high (or low) because of a small cluster of high loss observations, or even just a single high loss, and the measured VaR will continue to be high (or low) until n days or so have passed and the observation has fallen out of the sample period.



Problems from Long Window

- **The longer the window:** $n=1000$
  - The greater the problems with aged data;
  - The longer the period over which results will be distorted by unlikely-to-recur past events, and the longer we will have to wait for ghost effects to disappear;
  - The more the news in current market observations is likely to be drowned out by older observations;
  - The greater the potential for data-collection problems.

Estimating Curves and Surfaces for VaR and ES

- **The longer the window, the sparser the VaR curve.**
- **The VaR curve is fairly unsteady**, as it directly reflects the randomness of individual loss observations, **but the ES curve is smoother**, because each ES is an average of tail losses.
- **As the holding period rises, the number of observations rapidly falls, and we soon find that we don't have enough data.**
- Even if we had a **very long run of data, the older observations might have very little relevance for current market conditions.**

horizon

1年=250天 m=12

daily: 250 week=52

A/D of Non-parametric Methods

- **Advantages**
  - Intuitive and conceptually simple;
  - Do not depend on parametric assumptions;
  - Accommodate any type of position;
  - No need for covariance matrices, no curses of dimensionality;
  - Use data that are (often) readily available;
  - Are capable of considerable refinement and potential improvement if we combine them with parametric "add-ons" to make them semi-parametric.

A/D of Non-parametric Methods

- **Disadvantages**
  - Very dependent on the historical data set;
  - Subject to ghost effect;
  - If our data period was unusually quiet, non-parametric methods will often produce VaR or ES estimates that are too low, vice versa;
  - Have difficulty(act slowly) handling shifts(permanent risk change) that take place during our sample period;

Ghost effect:

1. React slowly to new market data

ES or high α-level VaR

2. Sudden fall - - - old VaR is retired.

hybrid VaR

Window is too long.

old data

ghost effect disappears slowly.

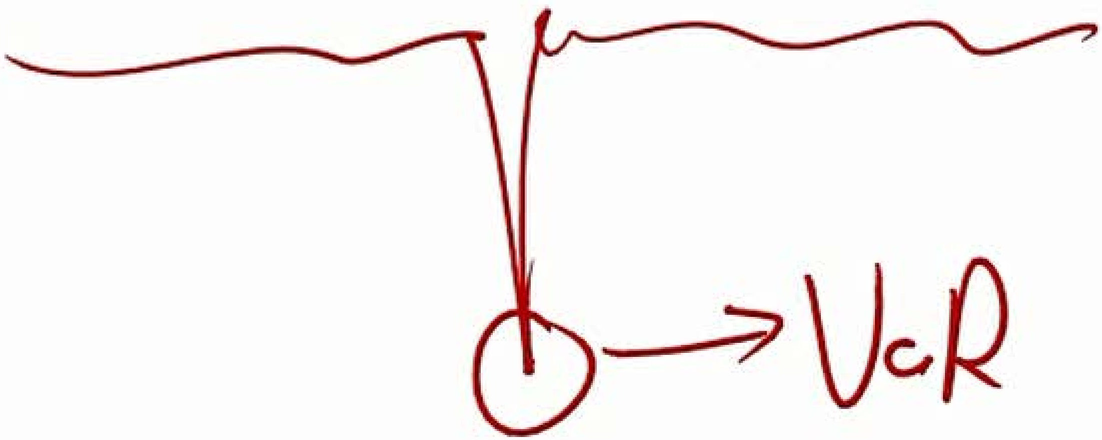
Horizon is too long: ⟶ daily VaR

Scare data / old data

A/D of Non-parametric Methods

one-off

- Have difficulty handling extreme value
  - ✓ If our data set incorporates extreme losses that are unlikely to recur, these losses can dominate non-parametric risk estimates even though we don't expect them to recur;
  - ✓ Make no allowance for plausible events that might occur, but did not actually occur, in our sample period.



Hybrid Approach

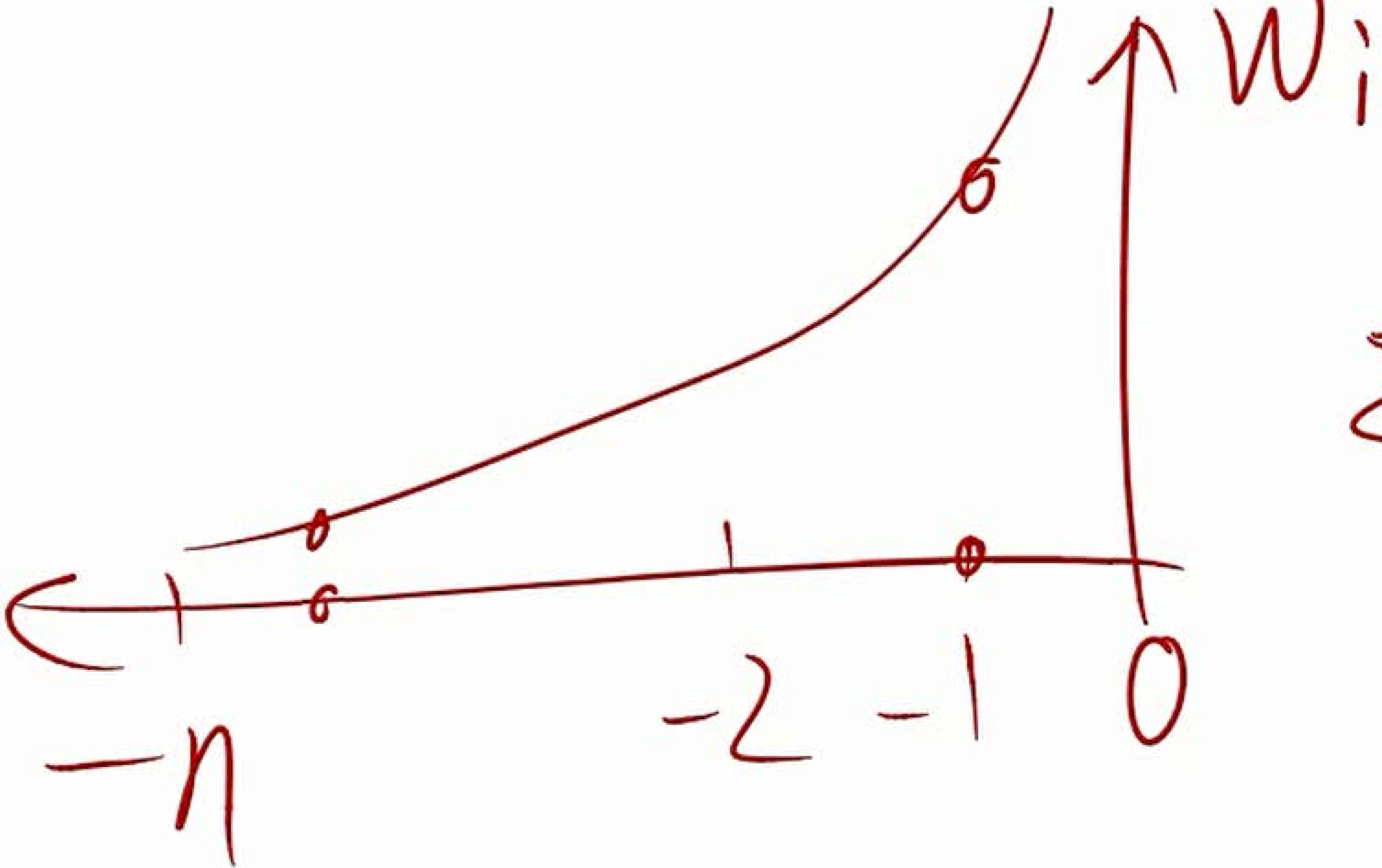
- **Parametric approach**
  - Normal VaR
  - Lognormal VaR
- **Non-Parametric approach(Historical Simulation)**
  - Bootstrap
  - Non-parametric density estimation
- **Hybrid approach(Semi-Parametric approach)**
  - Age-weighted(BRW)
  - Volatility-weighted(HW)
  - Correlation-weighted
  - Filtered historical simulation

Age-weighted Historical Simulation

➢ Boudoukh, Richardson and Whitelaw (BRW: 1998)

- $w_{(1)}$ is the probability weight given to an observation 1 day old.
- A λ close to 1 indicates a slow rate of decay, and a λ far away from 1 indicates a high rate of decay.

$$\omega_{(1)} + \lambda\omega_{(1)} + \ldots + \lambda^{n-1}\omega_{(1)} = 1 \rightarrow \omega_{(i)} = \frac{\lambda^{i-1}(1-\lambda)}{1-\lambda^n}$$



$$\sum w_i = 1$$

Age-weighted Historical Simulation

➢ **Boudoukh, Richardson and Whitelaw (BRW: 1998)**

- $w_{(1)}$ is the probability weight given to an observation 1 day old.
- A λ close to 1 indicates a slow rate of decay, and a λ far away from 1 indicates a high rate of decay.

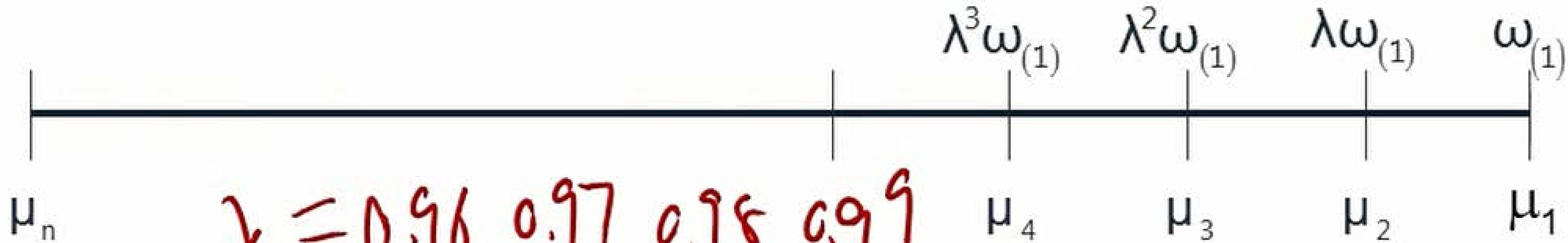
$$\omega_{(1)} + \lambda\omega_{(1)} + \dots + \lambda^{n-1}\omega_{(1)} = 1 \rightarrow \omega_{(i)}$$

$$\boxed{w_i = \frac{\lambda^{i-1}(1-\lambda)}{1-\lambda^n}}$$

Age-weighted Historical Simulation

➢ **Boudoukh, Richardson and Whitelaw (BRW: 1998)**

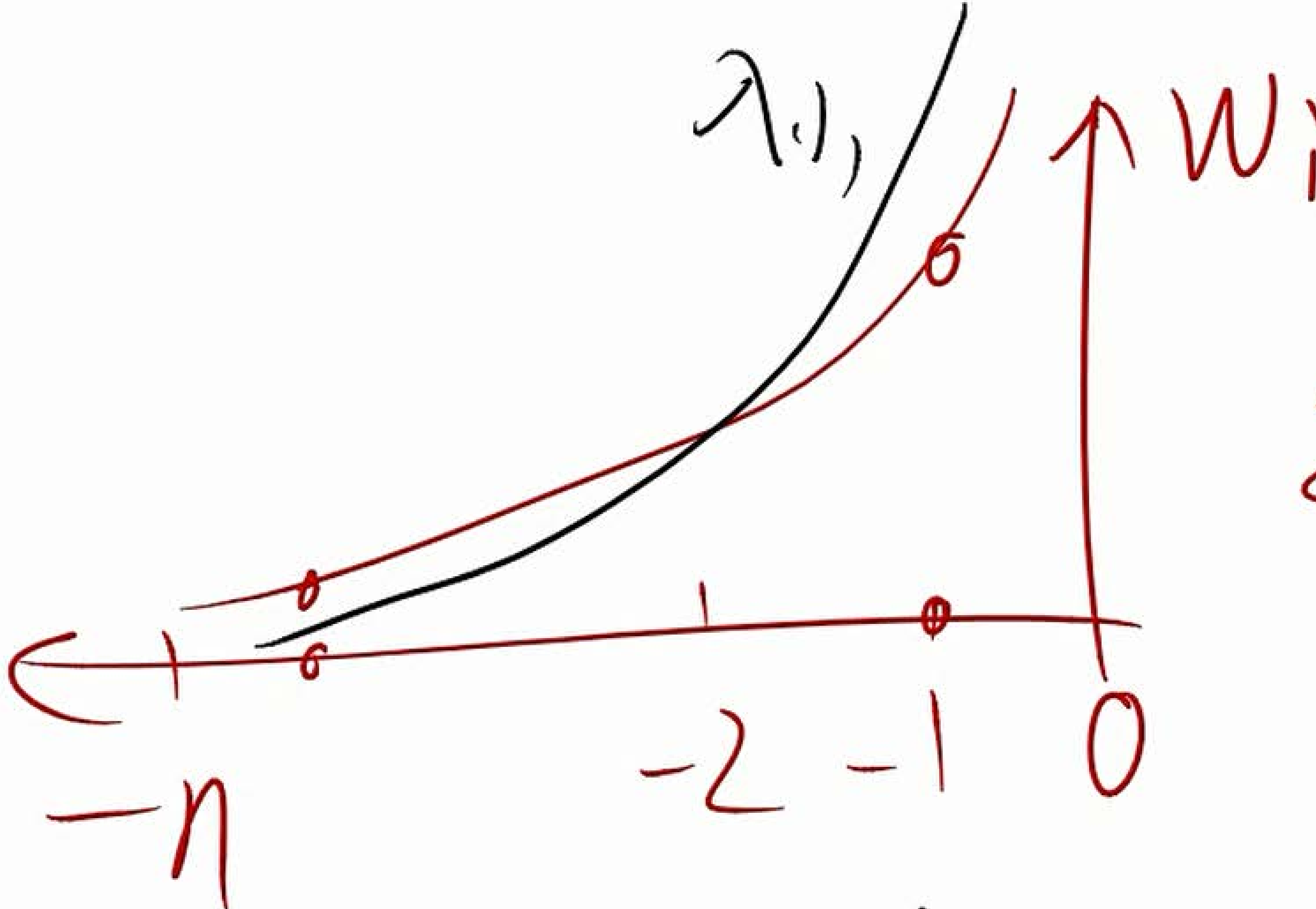
- $w_{(1)}$ is the probability weight given to an observation 1 day old.
- A λ close to 1 indicates a slow rate of decay, and a λ far away from 1 indicates a high rate of decay.



$\lambda = 0.96 \quad 0.97 \quad 0.98 \quad 0.99$

$$\omega_{(1)} + \lambda\omega_{(1)} + \dots + \lambda^{n-1}\omega_{(1)} = 1 \rightarrow \omega_{(i)} = \frac{\lambda^{i-1}(1-\lambda)}{1-\lambda^{n}}$$

$$\boxed{w_i = \frac{\lambda^{i-1}(1-\lambda)}{1-\lambda^{n}}}$$



$\sum \underline{\underline{w_i}} = 1$

EWMA

Age-weighted Historical Simulation

- **Major attractions**

  - It provides a nice generalization of traditional HS, because we can regard traditional HS as a special case with zero decay, or $\lambda \to 1$.
  - A large loss event will receive a higher weight than under traditional HS, and the resulting next-day VaR would be higher than it would otherwise have been.
  - Helps to reduce distortions caused by events that are unlikely to recur, and helps to reduce ghost effects.
    - ✓ As an observation ages, its probability weight gradually falls and its influence diminishes gradually over time. When it finally falls out of the sample period, its weight will fall from $\lambda^n \omega(1)$ to zero, instead of from $1/n$ to zero.

Age-weighted Historical Simulation

- **Major attractions**
  - Age-weighting allows us to let our sample period grow with each new observation, so we never throw potentially valuable information away. This would improve efficiency and eliminate ghost effects, because there would no longer be any "jumps" in our sample resulting from old observations being thrown away.
  - Age-weighting also reduces the effective sample size.

$$r_T \longrightarrow r^* \qquad \frac{r_T}{\sigma_T} = \frac{r^*}{\sigma_0}$$

T —————————————— 0

T: σ_T

熊 大

0: σ_0

牛 小

Volatility-weighted Historical Simulation

- **Hull and White (HW 1998)**
  - We adjust the historical returns to reflect how volatility tomorrow is believed to have changed from its past values.

$$\frac{r^*_{t,i}}{r_{t,i}} = \frac{\sigma_{T,i}}{\sigma_{t,i}}$$

$T \qquad t$

- ✓ $r_{t,i}$ = actual return for asset i on day t
- ✓ $\sigma_{t,i}$ = volatility forecast for asset i on day t
- ✓ $\sigma_{T,i}$ = current forecast of volatility for asset i

Volatility-weighted Historical Simulation

- **Major attractions**
  - It takes account of volatility changes in a natural and direct way.
  - It produces risk estimates that are appropriately sensitive to current volatility estimates.
  - It allows us to obtain VaR and ES estimates that can exceed the maximum loss in our historical data set.
    - ✓ In recent periods of high volatility, historical returns are scaled upwards, and the HS P/L series used in the HW procedure will have values that exceed actual historical losses.
  - Produces superior VaR estimates to the BRW one.

HW

✗	$\underline{HS} = \underline{1万}$	10天	$n=100$
✗	$BRW = \underline{1.2万}$	7天	
✓	$\boxed{HEW} = \underline{1.4万}$	5天	

➢ Hull and White (HW 1998)

- We adjust the historical returns to reflect how volatility tomorrow is believed to have changed from its past values.

$$\frac{r^*_{t,i}}{r_{t,i}} = \frac{\sigma_{T,i}}{\sigma_{t,i}}$$

$T \qquad t$

✓ $r_{t,i}$ = actual return for asset i on day t

✓ $\sigma_{t,i}$ = volatility forecast for asset i on day t

✓ $\sigma_{T,i}$ = current forecast of volatility for asset i

Correlation-weighted historical simulation

- **Correlation-weighted historical simulation**
  - Correlation-weighting is a little more involved than volatility-weighting.
  - To see the principles involved, suppose for the sake of argument that we have already made any volatility-based adjustments to our HS returns along Hull-White lines, but also wish to adjust those returns to reflect changes in correlations.

$r_T \rightarrow r^*$

$\rho_T \rightarrow 0$

$\rho_0 \rightarrow 1$

Filtered historical simulation

- **Filtered historical simulation(FHS)**
  - Combines historical simulation model with GARCH or AGARCH model.
- **The steps are as follows:**
  - Firstly, use the historical return to find any surprise and thus reproduce volatility with GARCH or AGARCH model.
  - Secondly, these volatility forecasts are then divided into the realized returns to produce a set of standardized returns, which is I.I.D..
  - The third stage involves bootstrapping from the set of standardized returns.
  - Finally, each of these simulated returns gives us a possible end-of-tomorrow portfolio value, and a corresponding possible loss, and we take the VaR to be the loss corresponding to our chosen confidence level.

Filtered historical simulation

- **Major attractions**
  - Combine the non-parametric attractions of HS with a sophisticated (e.g., GARCH) treatment of volatility, and so take account of changing market volatility conditions
  - It is fast, even for large portfolios
  - As with the earlier HW approach, FHS allows us to get VaR and ES estimates that can exceed the maximum historical loss in our data set.
  - It maintains the correlation structure in our return
  - It can be modified to take account of autocorrelations in asset returns
  - It can be modified to produce estimates of VaR or ES confidence intervals.
  - There is evidence that FHS works well.

VaR Backtest

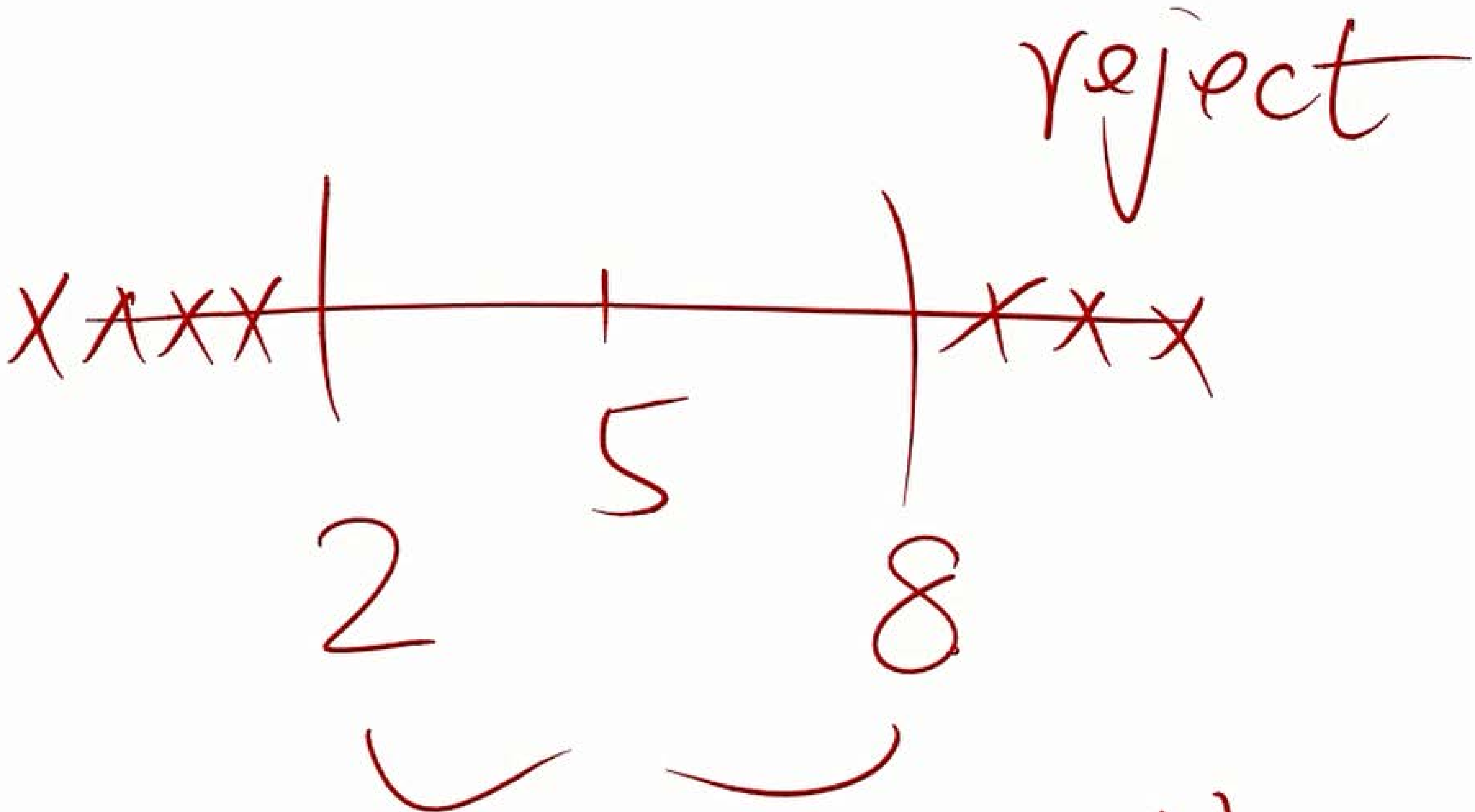
$n = 100$

$VaR = 1$万

$E = 6$

Exception: when loss > VaR

Exceedance



VaR is acceptable,

$\mu \pm 2.58\sigma$

?

?

$\mu = \underline{\underline{n \cdot p}}$

$VaR_{95\%}$ $p = 5\%$

$VaR_{99\%}$ $p = 1\%$

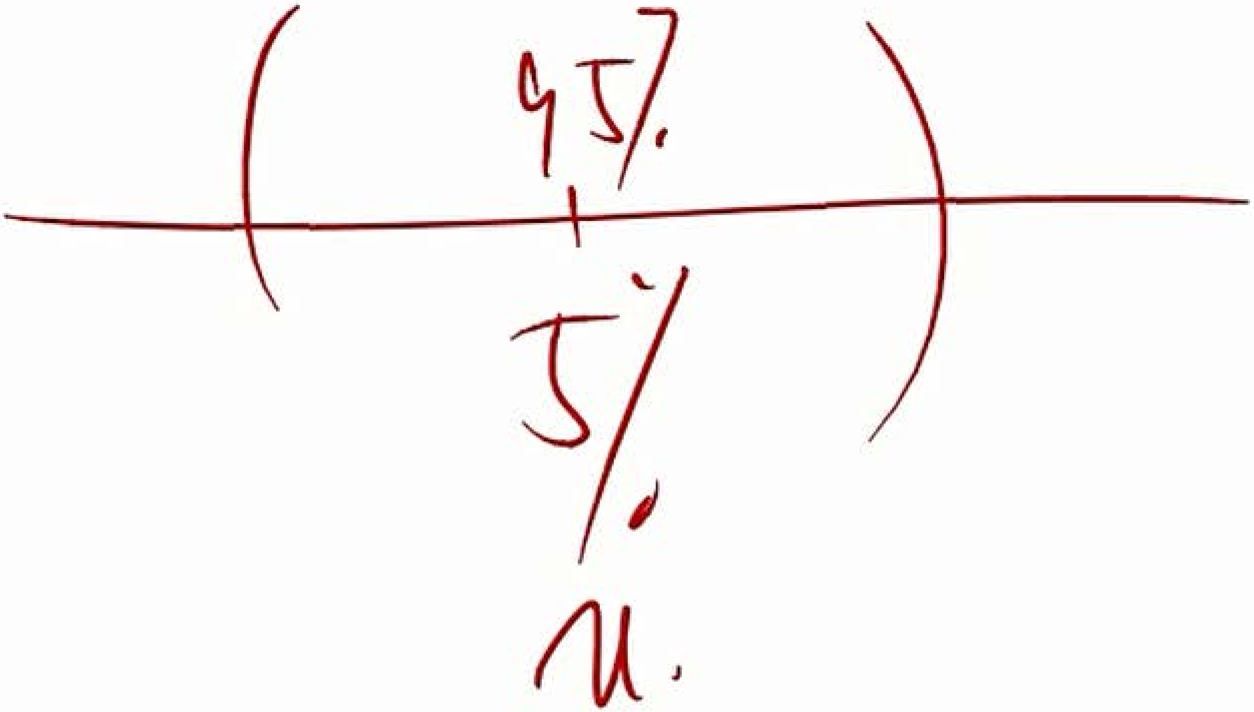
$$\mu = n \cdot p = 100 \times 5\% = 5\text{天}$$

$$\sigma^2 = np(1-p)$$

$$= 100 \times 5 \times 95\%$$

$$\mu \pm 1.65\sigma$$

$\mu \pm 1.96\sigma$ 双尾



VaR
- 95% 单 1.65
- 99% 单 2.33

天数 双尾
- 95% 1.96
- 99% 2.58

Model Verification Based on Failure Rates

- If the frequency of deviations becomes too large, the user must conclude that the problem lies with the model, not bad luck.
- **The failure rate**: the proportion of times VaR is exceeded in a given sample.
- The number of exceptions x follows a binomial probability distribution:

$$f(x) = C_T^x p^x (1-p)^{T-x}$$

- Approximate the binomial distribution by the normal distribution:

$$E(x) = pT \qquad V(x) = p(1-p)T$$

$$z = \frac{x - pT}{\sqrt{p(1-p)T}} \approx N(0,1)$$

Example

- In 1998, daily revenue of JP Morgan fell short of the downside (95% VaR) band on 20 days, or more than 5% of the time. Nine of these 20 occurrences fell within the August to October period.

$$z = \frac{x - pT}{\sqrt{p(1-p)T}} = \frac{20 - 0.05 * 252}{\sqrt{0.05(1-0.05)252}}$$
$$= 2.14 > 1.96$$

  - We reject the hypothesis that the VaR model is unbiased.

- What happens to test 99% VaR at 5% significance level and 95% VaR at 1% significance level?

Example

H_0: VaR correct H_a: VaR wrong

- In 1998, daily revenue of JP Morgan fell short of the downside (95% VaR) band on 20 days, or more than 5% of the time. Nine of these 20 occurrences fell within the August to October period.

up

$$z = \frac{20 - 250 \times 5\%}{\sqrt{np(1-p)}}$$

$$z = \frac{x - pT}{\sqrt{p(1-p)T}} = \frac{20 - 0.05 * 252}{\sqrt{0.05(1-0.05)252}} = 2.14 > 1.96$$

  - We reject the hypothesis that the VaR model is unbiased.

- What happens to test 99% VaR at 5% significance level and 95% VaR at 1% significance level?

Example

Ho: VaR correct   Ha: VaR wrong

- In 1998, daily revenue of JP Morgan fell short of the downside (95% VaR) band on 20 days, or more than 5% of the time. Nine of these 20 occurrences fell within the August to October period.

up

test: 95%

$$Z = \frac{20 - 250 \times 5\%}{\sqrt{np(1-p)}}$$

$$z = \frac{x - pT}{\sqrt{p(1-p)T}} = \frac{20 - 0.05 * 252}{\sqrt{0.05(1-0.05)252}} = 2.14 > 1.96$$

test: 99%

  - We reject the hypothesis that the VaR model is unbiased.

- What happens to test 99% VaR at 5% significance level and 95% VaR at 1% significance level?

$Z = 2.14 < 2.58$ 不拒绝

7					
8	95%				
9		252	days		
0	VaR c-level	lower	upper	mean	sigma
1	99.0%	-1	6	2.52	1.5794936
2	95.0%	6	19	12.6	3.4597688
3					

Nonrejection region

TABLE 3-2 Model Backtesting, 95 Percent Nonrejection Test Confidence Regions

Probability Level p	VaR Confidence Level c	Nonrejection Region for Number of Failures N		
		T = 252 Days	T = 510 Days	T = 1000 Days
0.01	99%	$N < 7$	$1 < N < 11$	$4 < N < 17$
0.025	97.5%	$2 < N < 12$	$6 < N < 21$	$15 < N < 36$
0.05	95%	$6 < N < 20$	$16 < N < 36$	$37 < N < 65$
0.075	92.5%	$11 < N < 28$	$27 < N < 51$	$59 < N < 92$
0.10	90%	$16 < N < 36$	$38 < N < 65$	$81 < N < 120$

➢ **Conclusion:**

- The longer the window, the easier the rejection.

Type I and Type II errors

- Users of VaR models need to balance type I errors against type II errors.
- Ideally, one would want to set a low type I error rate and then have a test that creates a very low type II error rate, in which case the test is said to be powerful. However, it is very difficult.

Decision	Correct	Incorrect
Accept	OK	Type II
Reject	Type I	Power of test

Kupiec VaR Backtest

- **Backtesting Exceptions**
  - Using Failure Rates in Model Verification
    - ✓ H_0: accurate model
    - ✓ H_a: inaccurate model
    - ✓ Test statistic:

$$LR_{uc} = -2\ln[(1-p)^{T-N}p^N] + 2\ln\{[1-(N/T)]^{T-N}(N/T)^N\}$$

      - p: the probability of exception, p = 1-c
      - N: the number of exceptions
      - T: the number of samples
  - We would reject the null hypothesis if LR > 3.841.

Kupiec VaR Backtest

VaR

LR$_{uc}$ Values for T=255

C level	N											
Ex	1	2	3	4	5	6	7	8	9	10	11	12
97.50%	~~7.16~~	~~4.19~~	2.27	1.04	0.33	0.02	0.06	0.39	0.98	1.81	2.84	~~4.06~~
98.00%	~~5.01~~	2.49	1.03	0.26	0	0.15	0.65	1.44	2.48	3.76	~~5.25~~	~~6.93~~
99.00%	1.24	0.13	0.08	0.71	1.86	$\boxed{3.42}$	~~5.32~~	~~7.51~~	~~9.97~~	~~12.65~~	~~15.55~~	~~18.63~~

接

- Conclusion:
  - The interval shrinks as the sample size extends.(Data should be larger)
  - Detection of systematic biases becomes increasingly difficult for high values of c because the exceptions in these cases are very rare events.(High confidence VaR should be avoided)

- Conclusion:
  - The interval shrinks as the sample size extends.(Data should be larger)
  - Detection of systematic biases becomes increasingly difficult for high values of c because the exceptions in these cases are very rare events.(High confidence VaR should be avoided)

Blame on Unconditional Coverage Models

- So far the framework focuses on unconditional coverage because it ignores conditioning, or time variation in the data. The observed exceptions, however, could **cluster** or " **bunch**" closely in time, which also should invalidate the model.

- In theory, these occurrences should be evenly spread over time. If, instead, we observed that 10 of these exceptions occurred over the last 2 weeks, this should raise a red flag.

Conditional Coverage Models

- **Christoffersen test**
  - Adjust $LR_{independence}$ to remove the feature of clustering.
    - ✓ If $LR_{independence} > 3.84$, we reject independence and conclude cluster.

Conclusions

- Ideally, one would want a framework that has very high power, or high probability of rejecting an incorrect model.
- The current framework could be improved by choosing a lower VaR confidence level or by increasing the number of data observations.
  - **The horizon should be as short as possible** in order to increase the number of observations and to mitigate the effect of changes in the portfolio composition.
  - **The confidence level should not be too high** because this decreases the effectiveness, or power, of the statistical tests.

Conditional Coverage Models

- **Christoffersen test**
  - Adjust $LR_{independence}$ to remove the feature of clustering.
    - ✓ If $LR_{independence} > 3.84$, we reject independence and conclude cluster.

Basel Committee Rules for Backtesting

- The Basel Committee requires that market VaR be calculated at the 99% confidence level and back testing over the past year. That is at the 99% confidence level, we would expect to have 2.5 exceptions (250×0.01) each year.
- Economic capital $= VaR \times (3 + k)$

Zone	Number of exceptions	Increase in K
Green	0-4	0
Yellow	5	0.4
	6	0.5
	7	0.65
	8	0.75
	9	0.85
Red	10+	1

$$\text{Capital} = \text{VaR} \times k$$

$$k \geq 3 \to 4$$

Basel Committee Rules for Backtesting

- **Four categories of causes for exceptions:**
  - Basic integrity of the model is lacking. Exceptions occurred because of incorrect data or errors in the programming. The penalty should apply.
  - Model accuracy needs improvement. The exceptions occurred because the model does not describe risks precisely. The penalty should apply.
  - Intraday trading. Positions changed during the day. The penalty should be considered.

    日内

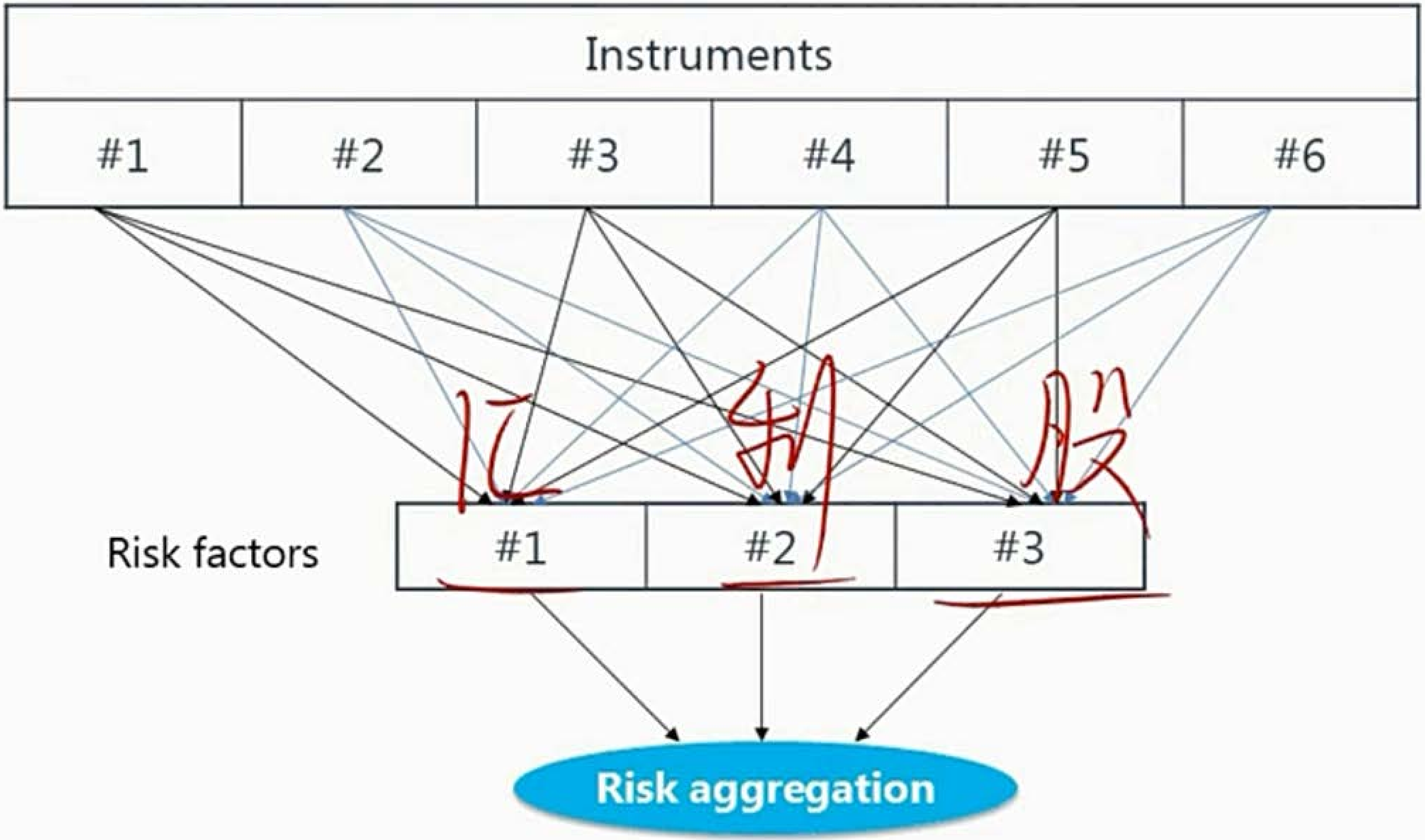
  - Bad luck. Markets were particularly volatile or correlations changed. These exceptions should be expected to occur at least some of the time.

Why Mapping?

- **Mapping is the process by which the current values of the portfolio positions are replaced by exposures on the risk factors.**
- **Mapping arises because of the fundamental nature of VaR.**
  - It would be too complex and time-consuming to model all positions individually as risk factors.
  - Furthermore, this is unnecessary because many positions are driven by the same set of risk factors and can be aggregated into a small set of exposures without loss of risk information.

Why Mapping?

- **For instance, a trader's desk with thousands of open dollar/euro forward contracts.** However the positions are exposed to a single major risk factor, which is the dollar/euro spot exchange rate.


Instruments
#1
#2
#3
#4
#5
#6
汇
利
股
Risk factors
#1
#2
#3
Risk aggregation

VaR Mapping

- **Three approaches for mapping**
  - **Principal mapping**. Only the risk associated with the return of principal at the maturity of the bond is mapped.
  - **Duration mapping**. With duration mapping, one risk factor is chosen that corresponds to the portfolio duration.
  - **Cash flow mapping**. With cash-flow mapping, the portfolio cash flows are grouped into maturity buckets.

principal mapping

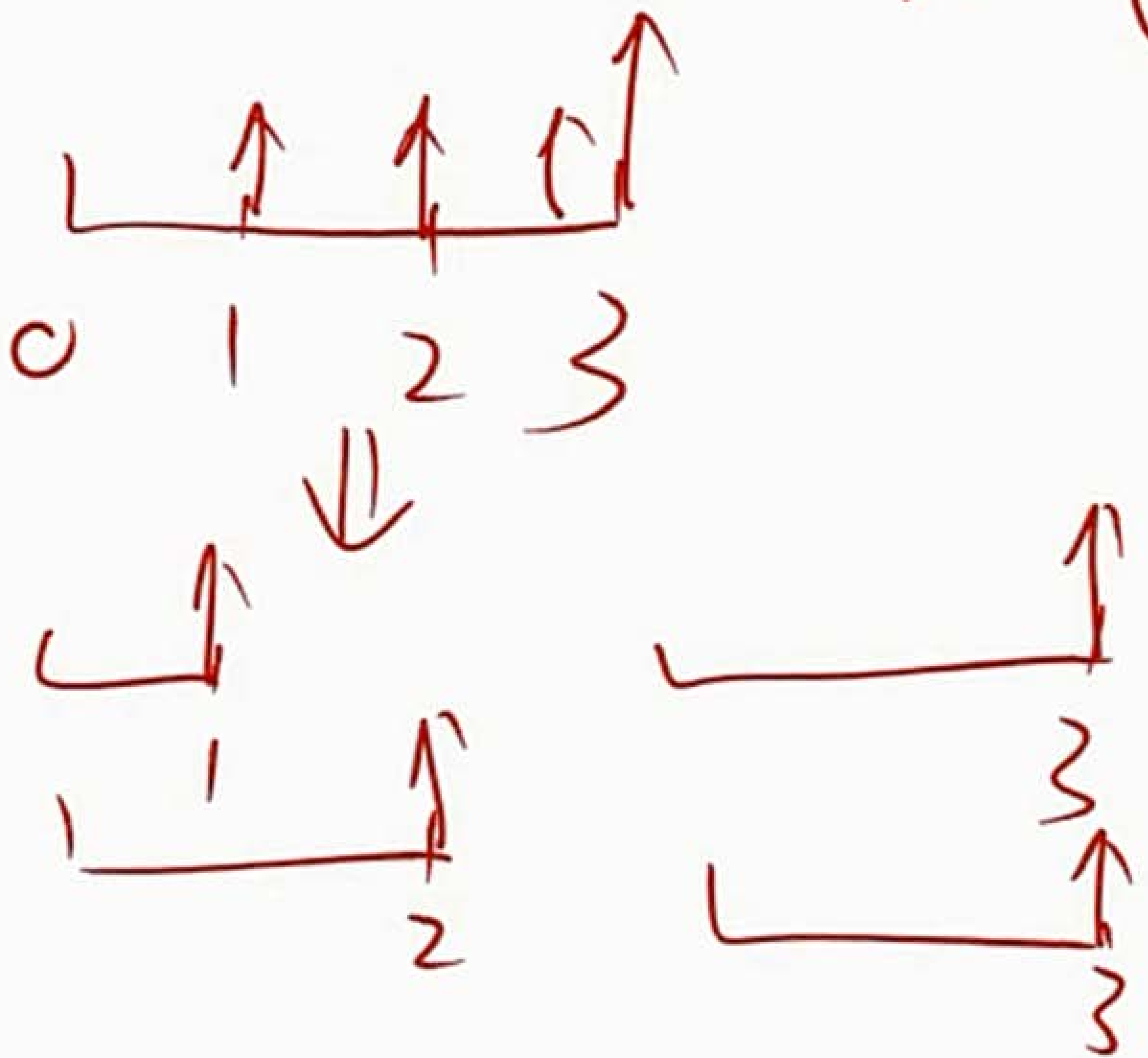
VaR

Duration Mapping

$D = 2.7$

Zero

Cash flow mapping



Case study

Jorion VaR

- A two-bond portfolio consisting of a $100 million 5-year 6% issue and a $100 million 1-year 4% issue.
- Both issues are selling at par, implying a market value of $200 million.
- The portfolio has an average maturity of 3 years and a duration of 2.733 years.
  - VaR(portfolio)=VaR(bond)*NP
  - VaR(portfolio)=Duration*VaR(int)*NP

TABLE 4-2 Mapping for a Bond Portfolio ($ millions)

	Cash Flows			Mapping (PV)		
Term (Year)	5-Year	1-Year	Spot Rate	Principal	Duration	Cash Flow
1	$6	$104	4.000%	0.00	0.00	$105.77
2	$6	0	4.618%	0.00	0.00	$5.48
2.733	—	—		—	$200.00	—
3	$6	0	5.192%	$200.00	0.00	$5.15
4	$6	0	5.716%	0.00	0.00	$4.80
5	$106	0	6.112%	0.00	0.00	$78.79
Total				$200.00	$200.00	$200.00

$$PV_1 \times VaR_1\% = \$VaR_1$$

$$PV_2 \times \boxed{VaR_2}\% \quad \$VaR_2$$

$$\vdots \qquad \vdots$$

$$PV_J \ \boxed{VaR_J}\% \quad \$VaR_J$$

$$\Sigma$$

$$VaR_p^2 = VaR_1^2 + VaR_2^2 + VaR_3^2$$

$$+ 2\rho_{12} VaR_1 VaR_2$$

$$+ 2\rho_{13} VaR_1 VaR_3$$

$$+ 2\rho_{23} VaR_2 VaR_3$$

$$\sigma_p^2 = (w_1\sigma_1)^2 + (w_2\sigma_2)^2 + 2\rho w_1 w_2 \sigma_1 \sigma_2$$

二次型

$$(w_1 \ w_2 \ w_3)\begin{pmatrix} \boxed{\sigma_{11}^2} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \boxed{\sigma_{22}^2} & \sigma_{23} \\ \sigma_{13} & \sigma_{32} & \boxed{\sigma_{33}^2} \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

$$\sigma_p^2 =$$

{=MMULT(MMULT(B1:D1,D4:F6),B3:B5)}

B	C	D	E	F
0.5	0.3	0.2		
0.5				
0.3		0.6	0.36	0.24
0.2		0.36	0.5	0.77
		0.24	0.77	0.4
0.4594				

VaR Mapping

➢ **Three approaches for mapping**

- ① **Principal mapping**. Only the risk associated with the return of principal at the maturity of the bond is mapped.
- ② **Duration mapping**. With duration mapping, one risk factor is chosen that corresponds to the portfolio duration.
- ③ **Cash flow mapping**. With cash-flow mapping, the portfolio cash flows are grouped into maturity buckets.

④ div VaR

Principal Mapping

VaR3 (给定)

- Principal mapping considers the timing of redemption payments only.
  - VaR(portfolio)=1.484% × $200 = $2.97million
  - VaR(portfolio)=3*VaR(3-yr int)*200M
- This method is simple but overstates the true risk because it ignores intervening coupon payments.

TABLE 4-3 Computing VaR from Change in Prices of Zeroes

Term (Year)	Cash Flows	Old Zero Value	Old PV of Flows	Risk (%)	New Zero Value	New PV of Flows
1	$110	0.9615	$105.77	0.4696	0.9570	$105.27
2	$6	0.9136	$5.48	0.9868	0.9046	$5.43
3	$6	0.8591	$5.15	1.4841	0.8463	$5.08
4	$6	0.8006	$4.80	1.9714	0.7848	$4.71
5	$106	0.7433	$78.79	2.4261	0.7252	$76.88
Total			$200.00			$197.37
Loss						$2.63

Duration Mapping

- We replace the portfolio by a zero coupon bond with maturity equal to the duration of the portfolio, which is 2.733 years.
- The duration-based VaR
  - VaR(portfolio)=$\left(0.987\% + (1.484\% - 0.987\%) \times (2.733 - 2)\right) \times \$200 =$ $\$2.70\text{million}$
  - VaR(portfolio)=2.733*VaR(2.733-yr int)*200M

VaR_2 VaR_3

$VaR_{2.733} \times 200 = VaR_{Duration}$

Cash Flow Mapping

- The cash-flow mapping method consists of grouping all cash flows on term structure "vertices" that correspond to maturities.

TABLE 4-4 Computing the VaR of a $200 Million Bond Portfolio (monthly VaR at 95 percent level)

	PV Cash Flows	Individual VaR	Correlation Matrix R					Component VaR
Term (Year)	x	$x \times V$	1Y	2Y	3Y	4Y	5Y	$x\Delta$VaR
1	$105.77	0.4966	1					$0.45
2	$5.48	0.0540	0.897	1				$0.05
3	$5.15	0.0765	0.886	0.991	1			$0.08
4	$4.80	0.0947	0.866	0.976	0.994	1		$0.09
5	$78.79	1.9115	0.855	0.966	0.988	0.998	1	$1.90
Total	$200.00	2.6335						
Undiversified VaR		$2.63						
Diversified VaR								$2.57

- Undiversified VaR=$2.63million
- Diversified VaR=$2.57million

Cash Flow Mapping

- Each cash flow is represented by the present value of the cash payment, discounted at the appropriate zero-coupon rate.
  - Undiversified VaR: considering interest rate VaRs with each maturities
  - Diversified VaR: also considering interest rate correlations
- Var(Prn)>VaR(duration)> VaR(Diversified)

Mapping Linear Derivatives

- **Forward and Futures**
  - are the simplest types of derivatives. Since their value is linear in the underlying spot rates, their risk can be constructed easily from basic building blocks.
- **Currency forward contracts**

$$F_t = (S_t e^{-yt})e^{rt}$$

  - S_t = spot price of the underlying cash asset
  - r = domestic risk-free rate
  - y = income flow on the asset
  - t = time to maturity
- **Long currency forward contract = long foreign currency spot + long foreign currency bill + short U.S. dollar bill**

Mapping Linear Derivatives

- **Forward and Futures**
  - are the simplest types of derivatives. Since their value is linear in the underlying spot rates, their risk can be constructed easily from basic building blocks.
- **Currency forward contracts**

$$F_t = (S_t e^{-yt})e^{rt}$$

$$F_t = S_0 e^{(r_D - r_F)t}$$

$$= (S_0 e^{-r_F t}) e^{r_D}$$

  - S_t = spot price of the underlying cash asset
  - r = domestic risk-free rate
  - y = income flow on the asset
  - t = time to maturity
- **Long currency forward contract = long foreign currency spot + long foreign currency bill + short U.S. dollar bill**

$S_0 = 7\,RMB/\$$

$$
\begin{tikzcd}
\text{借}\ \$1 \arrow[r, "\times 7"] & 7\,RMB\,(\text{外币}) \arrow[d, "\text{买入外国bill}\,(r_f)"] \\
USD\,(\text{本}) \arrow[u, "\text{还}\ \$1\times(1+r_D)"] & 7\times(1+r_f) \arrow[l, "\div"']
\end{tikzcd}
$$

- t = time to maturity

7X(CHF) (RMB

➢ **Long currency forward contract = long foreign currency spot + long foreign currency bill + short U.S. dollar bill**

$-\$1\times(1+r_0)$

Mapping Linear Derivatives

- Let us examine the risk of a 1-year forward contract to purchase 100 million euros in exchange for $130.086 million.

TABLE 4-6 Risk and Correlations for Forward Contract Risk Factors (monthly VaR at 95 percent level)

			Correlations		
Risk Factor	**Price or Rate**	**VaR (%)**	**EUR Spot**	**EUR 1Y**	**USD 1Y**
EUR spot	$1.2877	4.5381	1	0.1289	0.0400
Long EUR bill	2.2810%	0.1396	0.1289	1	−0.0583
Short USD bill	3.3304%	0.2121	0.0400	−0.0583	1
EUR forward	$1.3009				

- Sources of Risk
  - **The volatility of the spot contract** is the highest by far, with a 4.54 percent VaR (95 % confidence level).
  - This is much greater than the 0.14 percent VaR for **the EUR 1-year bill** or even the 0.21 percent VaR for **the USD bill**. Thus most of the risk of the forward contract is driven by the cash EUR position.
- Risk is also affected by **correlations**.

Mapping Linear Derivatives

TABLE 4-7 Computing VaR for a EUR 100 Million Forward Contract (monthly VaR at 95 percent level)

Position	Present-Value Factor	Cash Flows (CF)	PV of Flows, x	Individual VaR, $\|x\|V$	Component VaR, $x\Delta VaR$
EUR spot			\$125.89	\$5.713	\$5.704
Long EUR bill	0.977698	EUR100.00	\$125.89	\$0.176	\$0.029
Short USD bill	0.967769	−\$130.09	−\$125.89	\$0.267	\$0.002
Undiversified VaR				\$6.156	
Diversified VaR					\$5.735

- Undiversified VaR=\$5.713 million
- Diversified VaR=\$5.735 million

Mapping Linear Derivatives

- **Forward Rate Agreements**
  - Are forward contracts that allow users to lock in an interest rate at some future date.
  - The buyer of an FRA locks in a borrowing rate; the seller locks in a lending rate. In other words, the " long" receives a payment if the spot rate is above the forward rate.
  - Suppose that you sold a 6×12 FRA on $100 million. This is equivalent to borrowing $100 million for 6 months and investing the proceeds for 12 months.
  - **Long 6×12 FRA = long 6-month bill + short 12-month bill**

$(\tilde{F} - \bar{K}) \times NP$

for long

$-\$100$ $\quad +100 \times \left(1+\frac{S_6}{2}\right)$

$$\boxed{100 \times \left(1+\frac{S_6}{2}\right)\left(1+\frac{F}{2}\right) - 100 \times (1+S_{12})}$$

$+100$

$$\Downarrow$$

$$-100 \times \left(1+\frac{S_6}{2}\right)\left(1+\frac{K}{2}\right)$$

$$S_6 \rightarrow k \leftarrow S_{12}$$

$$\left(1+\frac{S_6}{2}\right)\left(1+\frac{k}{2}\right) = 1 + S_{12}$$

$(\tilde{F}-\bar{K})\times NP$

for long

$-\$100$

$+100\times(1+\frac{S_6}{2})$

$100\times(1+\frac{S_6}{2})(1+\frac{F}{2})$

$-100\times(1+S_{12})$

$\Downarrow$

$-100\times(1+\frac{S_6}{2})(1+\frac{K}{2})$

$+100$

$\underbrace{100\times(1+\frac{S_6}{2})}_{NP}\left[\frac{K}{2}-\frac{F}{2}\right]$

Mapping Linear Derivatives

- If the 360-day spot rate is 5.8125 percent and the 180-day rate is 5.6250 percent, the forward rate must be such that $1+\frac{f_{1,2}}{2}=\frac{1+5.8125\%}{1+5.625\%} \Rightarrow f_{1,2}=$ 5.836%.
- The present value of the notional $100 million in 6 months is x = $100/(1 + 5.625/200) =$97,264 million

TABLE 4-9 Computing the VaR of a $100 Million FRA (monthly VaR at 95 percent level)

Position	PV of Flows, x	Risk (%), V	Correlation Matrix, R		Individual VaR, $\lvert x\rvert V$	Component VaR, $x\Delta$VaR
180 days	−$97.264	0.1629	1	0.8738	$0.158	−$0.116
360 days	$97.264	0.4696	0.8738	1	$0.457	$0.444
Undiversified VaR					$0.615	
Diversified VaR						$0.327

Mapping Linear Derivatives

收固一付浮

- **Interest Rate Swaps**
  - A payer swap is equivalent to buying a floating rate bond and simultaneously shorting a fixed rate bond.
  - **Or** a series of forward rate agreement with same forward rate.
- Let us compute the VaR of a $100 million 5-year interest-rate swap. Pays 6.195 percent annually for 5 years in exchange for floating-rate payments indexed to LIBOR.

VaR − VaR

  - We consider a situation where the floating-rate note is about to be reset.

+fixed bond −float bond.

NP

Mapping Options

- **BSM Model:**

$$c = Se^{-yt}N(d_1) - Ke^{-rt}N(d_2)$$

- **Long call option = long $N(d_1)$*asset + short $N(d_2)$ bill**
- **Long put option = long $N(-d_2)$*bill+ short $N(-d_1)$*asset**
- Linear approximations may be acceptable for options with long maturities when the risk horizon is short.

$$+\underline{\underline{S}}\,\underbrace{N(d_1)}_{\substack{<1\\0.7}} - \left(Ke^{-rt}\right)\underbrace{N(d_2)}_{\substack{<1\\0.6}}$$

Mapping Options

Ke^{-rt} K

0

- **BSM Model:**

$$c = Se^{-yt}N(d_1) - Ke^{-rt}N(d_2)$$

- **Long call option = long $N(d_1)$*asset + short $N(d_2)$ bill**
- **Long put option = long $N(-d_2)$*bill+ short $N(-d_1)$*asset**
- Linear approximations may be acceptable for options with long maturities when the risk horizon is short.

$$+ S\underbrace{N(d_1)}_{<1} - (Ke^{-rt})\underbrace{N(d_2)}_{<1}$$

0.7 0.6

VaR Mapping

> **Which of these statements regarding risk factor mapping approaches is/are correct?**

I. Under the cash flow mapping approach, only the risk associated with the average maturity of a fixed-income portfolio is mapped.

II. Cash flow mapping is the least precise method of risk mapping for a fixed-income portfolio.

III. Under the duration mapping approach, the risk of a bond is mapped to a zero-coupon bond of the same duration.

IV. Using more risk factors generally leads to better risk measurement but also requires more time to be devoted to the modeling process and risk computation.

A. I and II

B. I, III, and IV

C. III and IV

D. IV only

> **Answer: C**

Trading Book	Banking Book
held for trading	held to maturity
mark to market	historical cost
market risk	credit risk

VaR Implementation

- **The three categories of implementation issues reviewed are:**
  - (1) time horizon over which VaR is estimated;
    - ✓ The appropriate VaR horizon varies across positions and **depends on the position's nature and liquidity**.
    - ✓ For regulatory capital purposes, the horizon should be long, and yet the common square root of time scaling approach for short horizon VaR (e.g., one-day VaR) may generate biased long horizon VaR (e.g., ten-day VaR) estimates.

$$VaR_t = VaR_1\sqrt{t}$$

VaR Implementation

- **The three categories of implementation issues reviewed are:**
  - (2) the recognition of time-varying volatility in VaR risk factors;
    - ✓ The implication is that capturing time-varying volatility **may not be as important when the VaR horizon is long**, compared to when the VaR horizon is relatively short.
    - ✓ In contrast, volatility generated by **stochastic jumps** will diminish the accuracy of long-horizon VaR measures unless properly account for the jump features of the data.
    - ✓ using VaR with time-varying volatility for regulatory capital raises the concerns of volatile and potentially pro-cyclical regulatory standards.

VaR Implementation

- **The three categories of implementation issues reviewed are:**
  - (3) VaR backtesting
    - ✓ Backtesting is not effective when the number of VaR exceptions is small. In addition, backtesting is less effective over longer time horizons due to portfolio instability.

Integrating Liquidity into VaR Models

外

- **Exogenous liquidity** refers to the transaction cost for trades of average size
  - easily integrated into a VaR framework

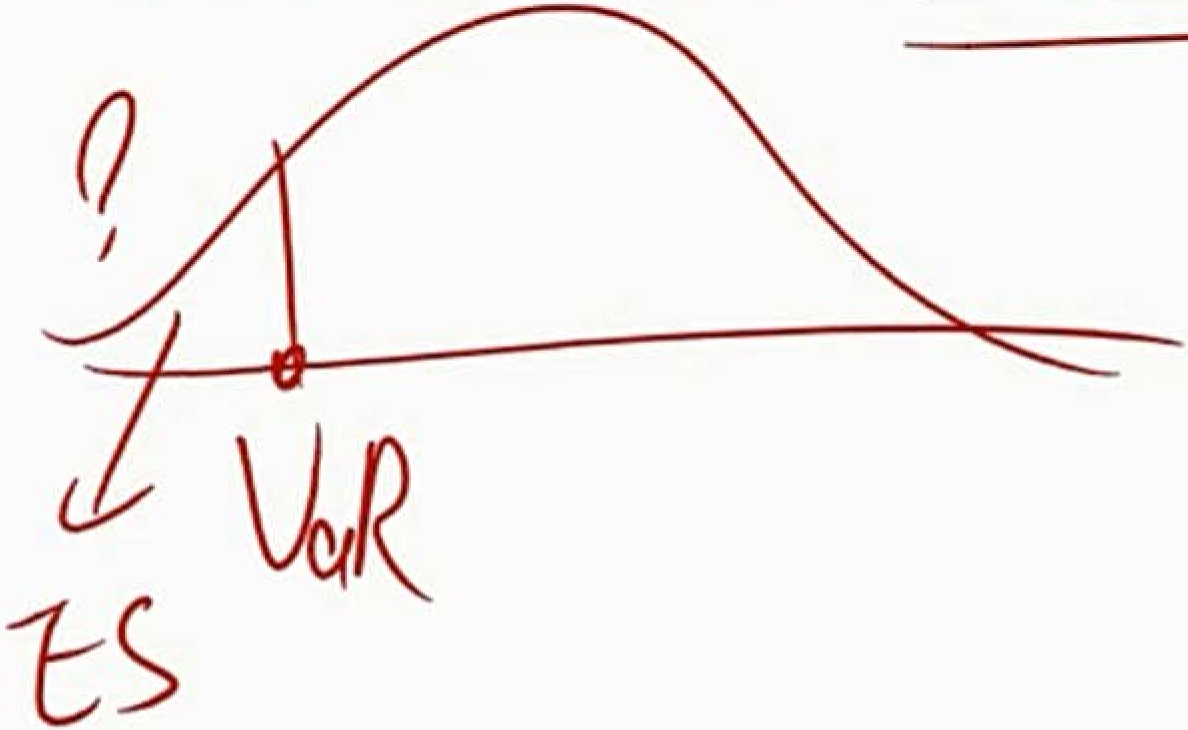
内

- **Endogenous liquidity** is related to the cost of unwinding portfolios large enough that the bid-ask spread cannot be taken as given, but is affected by the trades themselves.
  - more difficult to include in a VaR computation

$VaR_{日} = 1万$ 99%

Risk Measures

- **Value at Risk (VaR)** has become a standard risk measure in finance.
  - VaR measures only quantiles of losses, and thus disregards any loss beyond.
  - It has been criticized for its lack of subadditivity.



Risk Measures

➢ **Expected Shortfall (ES)** is the most well-known risk measure following VaR.

- ES corrects three shortcomings of VaR.
  - ✓ **First**, ES does account for the severity of losses beyond the confidence threshold. This property is especially important for regulators, who are concerned about exactly these losses. **Second**, it is always subadditive and coherent. **Third**, it mitigates the impact that the particular choice of a single confidence level may have on risk management decisions, while there is seldom an objective reason for this choice.
- Expected shortfall is more complex and computationally intensive.

Risk Measures

- **Spectral risk measures**
  - are a promising generalization of expected shortfall . **They have advantages over expected shortfall**, including **favorable smoothness properties** and the possibility of adapting the risk measure directly to the risk aversion of investors.

谱

W_i

ES

Integrated Risk Measurement

- There are two ways of modelling the aggregate risks present across a bank's trading and banking books
  - a **compartmentalised approach**—namely, the sum of risks measured separately
  - or **a unified approach** that considers the interaction between these risks explicitly.

Risk Aggregation

- **Top-down approach**
  - Reference an institution as a whole. With respect to market and credit risk, the top-down approach explicitly assumes that the risks are separable and can be aggregated in some way.
  - Risk diversification is present and ignored by the separate approach.
- **Bottom-up approach**
  - A common assumption of most current risk measurement models is that market and credit risks are separable and can be addressed independently. Bottom-up approaches can range from the portfolio level up to an institutional level.
  - Risk diversification should be questioned.

Balance Sheet Management

- A literature stream on the systemic consequences of individual risk management systems as the basis of regulatory capital charges has found that the mechanical link between measured **risks** derived from risk models and historical data and **regulatory capital charges** can work **as a systemic amplifier of boom and bust cycles.**

Asset = 100 $\xrightarrow{8\%}$ Capital = 8 72

900 1000 80

专业 · 创新 · 增值

Balance Sheet Management

➢ When a balance sheet is actively managed, the amount of leverage on the balance sheet becomes $\boxed{\text{procyclical}}$. 顺周期性

- An **economic boom** will relax this VaR constraint since a bank's level of equity is expanding. Thus, this expansion allows financial institutions to take on more risk and further increase debt.
- In contrast, **an economic bust** will tighten the VaR constraint and force investors to reduce leverage by selling assets when market prices and liquidity are declining.

Some Correlation Basics

Modeling Dependence: Correlations And Copulas

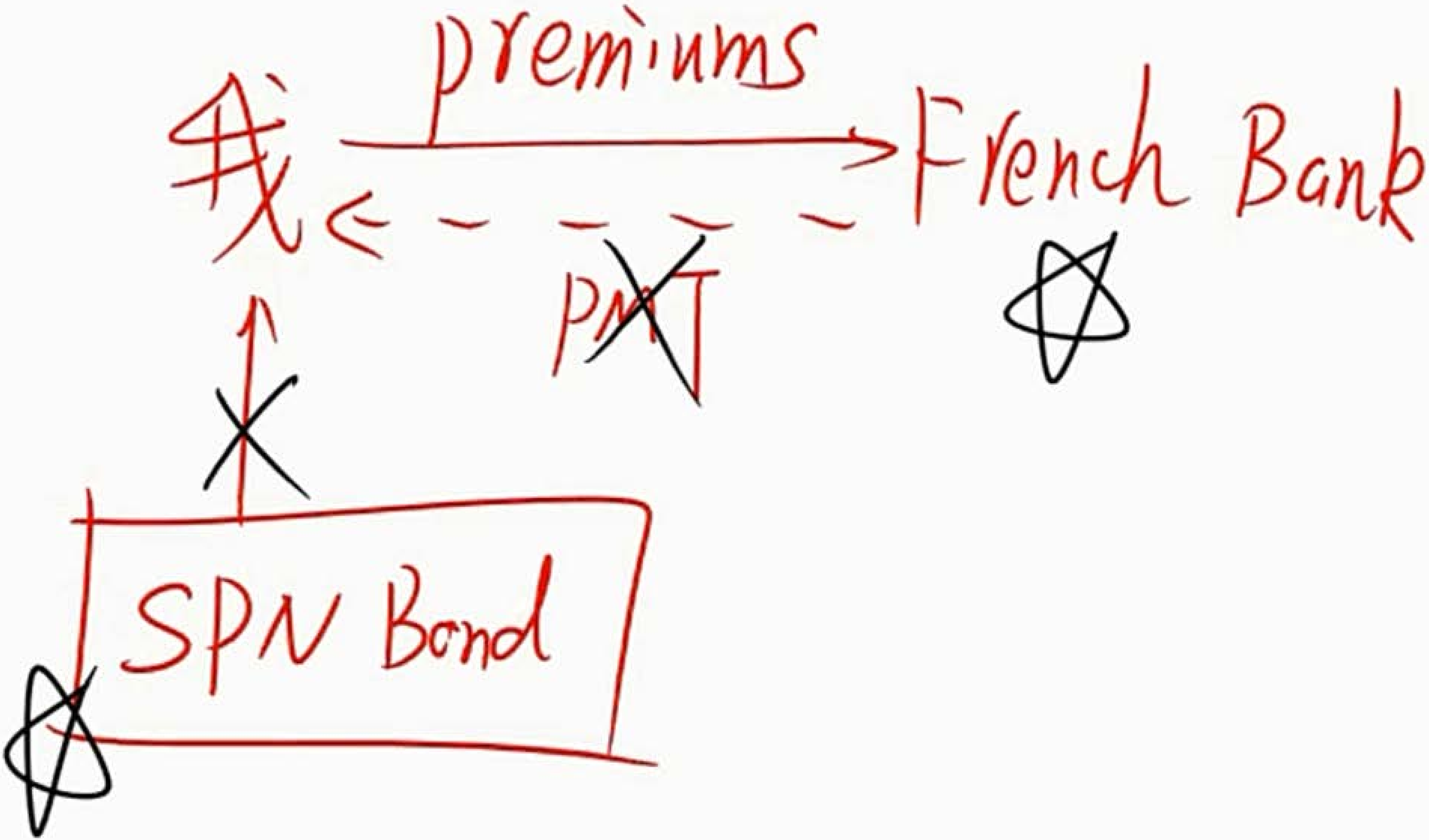
What are Financial Correlation risk?

- Financial correlation risk is the risk of financial loss due to adverse movements in correlation between two or more variables.
- Dynamic financial correlations measure how two or more financial assets move together in time.

Example 1

CDS

- **Invest in a bond from Spain, while purchase a credit default swap from a French bank:**
  - The value of the CDS is determined by the default probability of the reference entity Spain and by the joint default correlation of the French bank and Spain.
  - If the correlation between Spain and BNP Paribas increases, the present value of the CDS for the investor will decrease.
  - **Wrong way risk**: when the exposure to Spain increases, it is more unlikely that the counterparty French bank can pay the default insurance.


premiums
我
French Bank
PMT
SPN Bond

Example 1

CDS

- **Invest in a bond from Spain, while purchase a credit default swap from a French bank:**
  - The value of the CDS is determined by the default probability of the reference entity Spain and by the joint default correlation of the French bank and Spain.
  - If the correlation between Spain and BNP Paribas increases, the present value of the CDS for the investor will decrease.
  - **Wrong way risk**: when the exposure to Spain increases, it is more unlikely that the counterparty French bank can pay the default insurance.

CR

Example 2

- **A quanto option**: allows a domestic investor to exchange his potential option payoff in a foreign currency back into his home currency at a fixed exchange rate.
  - **For example**, an American believes the Nikkei will increase, but she is worried about a decreasing yen. The investor can buy a quanto call on the Nikkei, with the yen payoff being converted into dollars at a fixed exchange rate.

Example 2

- If the correlation is positive, an increasing Nikkei will also mean an increasing yen. That is favorable for the call seller. She has to settle the payoff, but only needs a small yen amount to achieve the dollar payment.
  - The **more positive** the correlation coefficient, the lower the price for the quanto option.
  - The **lower the correlation** coefficient, the more expensive the quanto option.

investor $\xrightarrow{¥}$ Nikkei ↑

P: Nikkei/¥ → >0

Nikkei → 1年后 → profit(¥) ↑

USD $\xleftarrow{\text{Quanto?}}$ profit(¥)

K 1日 ⇒ 2美

MKT 1日 ⇒ 3美

investor $\xrightarrow{¥}$ Nikkei ↑

ρ: Nikkei/¥ → >0

作后

profit(¥) ↑

~~Quanto~~ ↓

USD ← profit(¥)

MKT rate

GOLDEN FUTURE

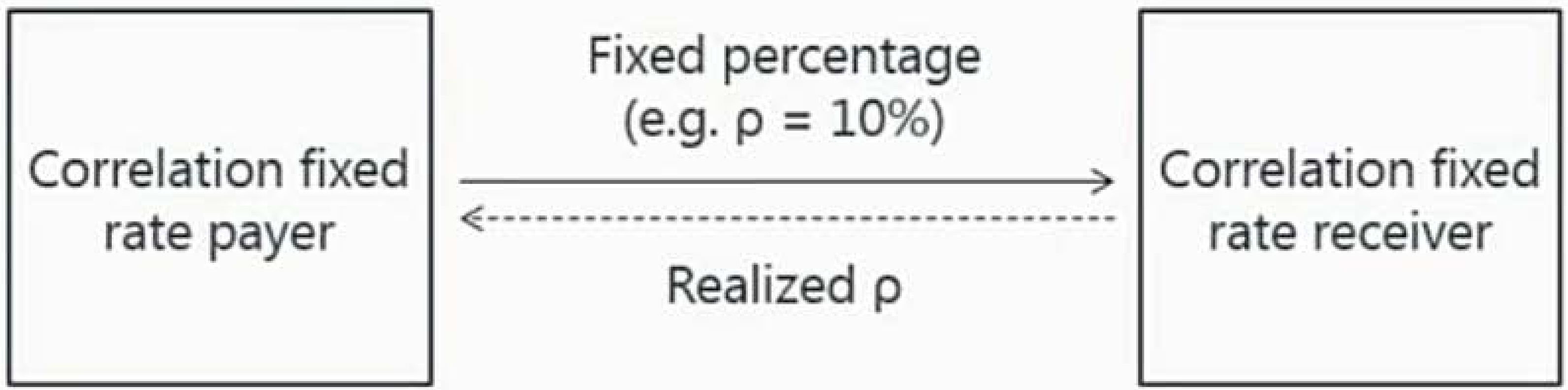
Example 2

- If the correlation is positive, an increasing Nikkei will also mean an increasing yen. That is favorable for the call seller. She has to settle the payoff, but only needs a small yen amount to achieve the dollar payment.
  - The **more positive** the correlation coefficient, the lower the price for the quanto option.

    Nikkei vs ¥
  - The **lower the correlation** coefficient, the more expensive the quanto option.

Correlation Swap

- Paying a fixed rate in a correlation swap is also called buying correlation. The payoff = NP * (realized ρ – fixed ρ).
- This is because the present value of the correlation swap will increase for the correlation buyer if the realized correlation increases.
- The fixed rate receiver is selling correlation.

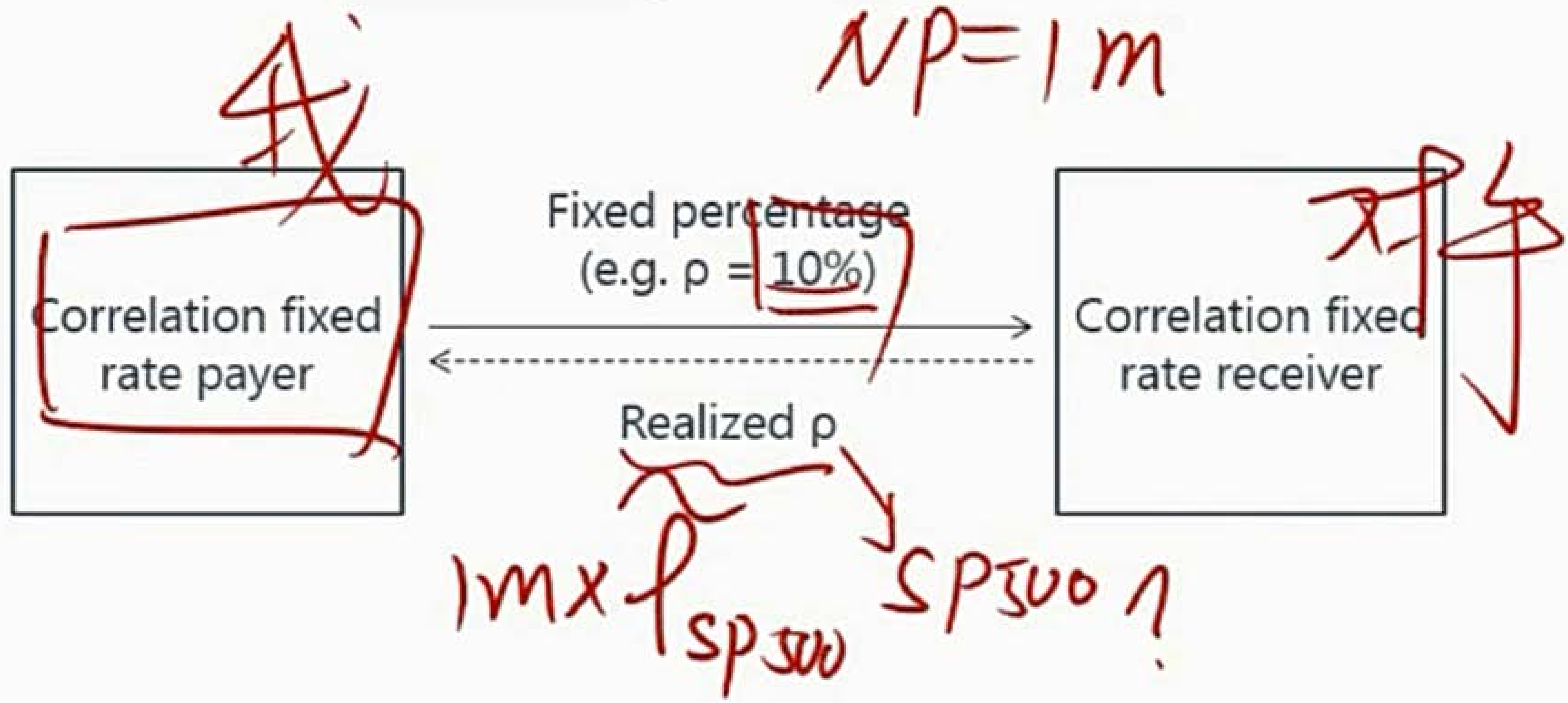


Correlation Swap

- Paying a fixed rate in a correlation swap is also called buying correlation. The payoff = NP * (realized ρ – fixed ρ).

SP500 → ρ

- This is because the present value of the correlation swap will increase for the correlation buyer if the realized correlation increases.
- The fixed rate receiver is selling correlation.



$$CF = \bar{\rho}_{fix} \times NP = 0.1\,m$$

10% $1m$

$$O \underset{NP \times \hat{\rho} = 100 \times ?}{\overset{NP \times \bar{\rho} = 100 \times 0.1}{\rightleftarrows}} O$$

Correlation Swap

- There is a positive relationship between correlation and volatility.
- Another way of buying correlation (i.e., benefiting from an increase in correlation) is **to buy call options on an index such as the Dow Jones Industrial Average and sell call options on individual stocks of the Dow.**
  - Therefore, if correlation between the stocks of the Dow increases, so will the implied volatility of the call on the Dow.
  - This increase is expected to outperform the potential loss from the increase in the short call positions on the individual stocks.

call

S.
K
T

$\rho\uparrow \rightarrow \sigma_{DJIA}\uparrow$

买入 $\uparrow$ call on DJIA

我

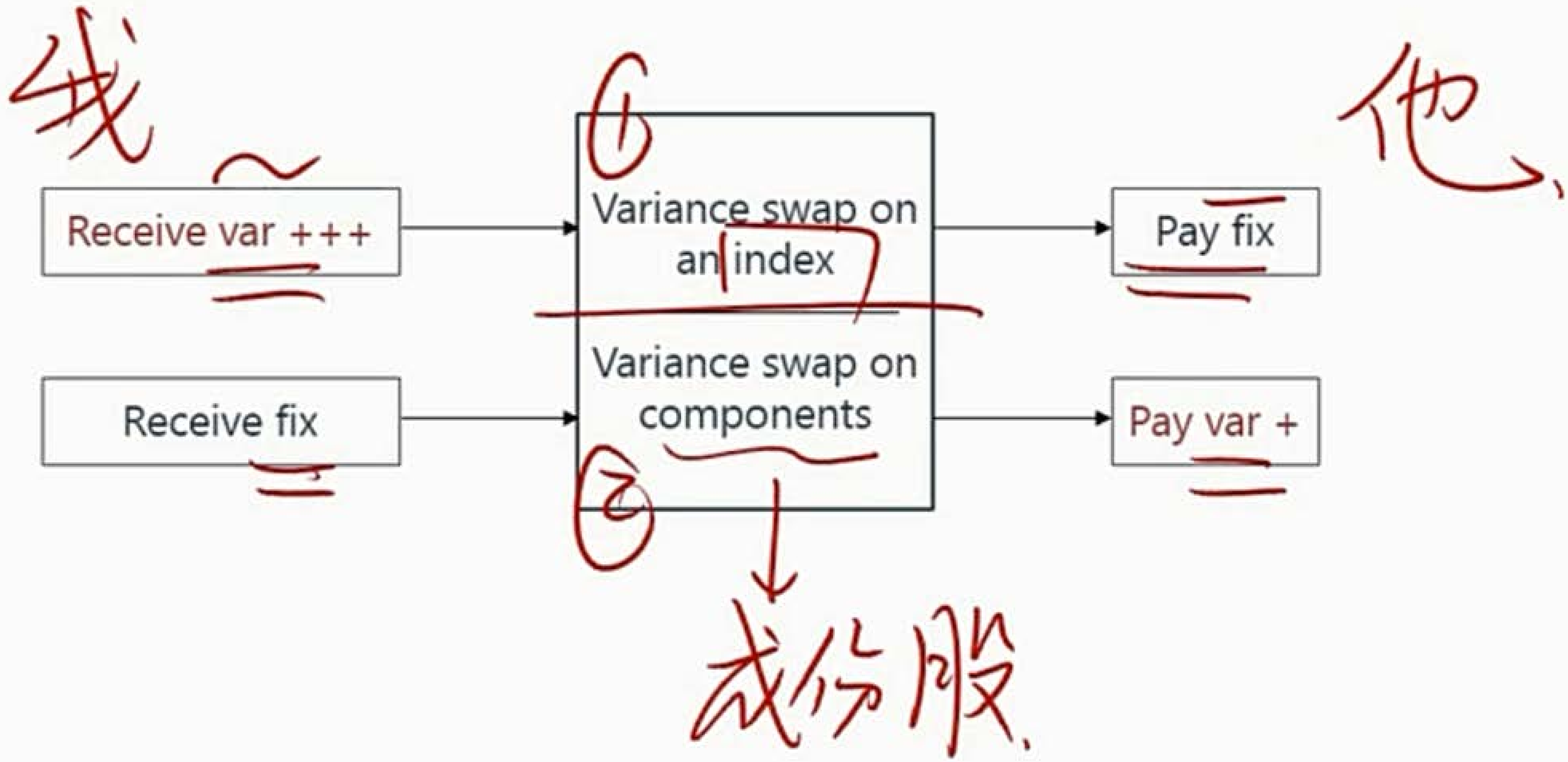
卖出 call on coke $\uparrow$

$\rho\uparrow \rightarrow \sigma_{coke}\uparrow$

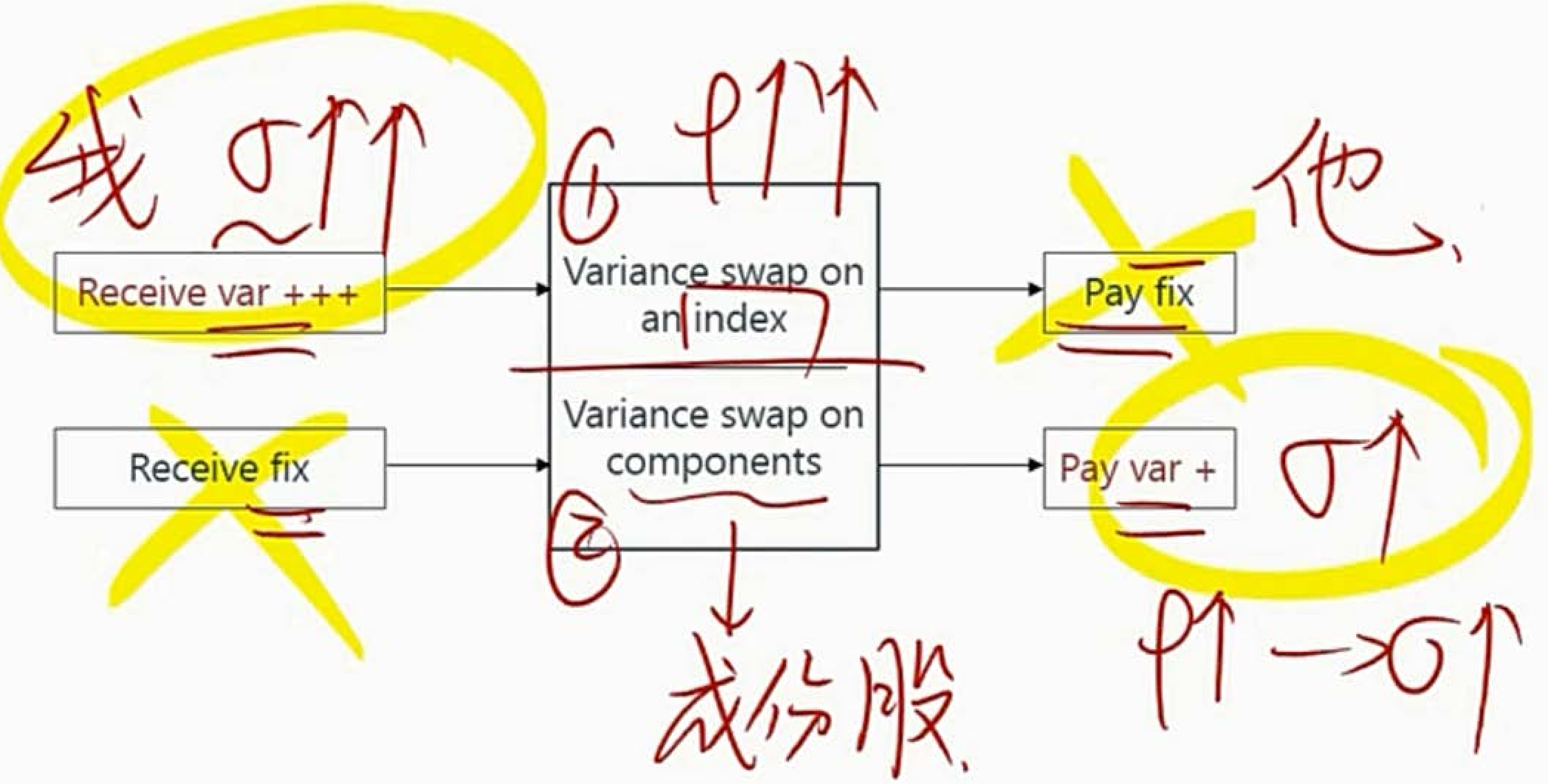
Variance Swap

- If correlation increases, so will the variance.
- A further way to buy correlation is to **pay fixed in a variance swap on an index and to receive fixed in variance swaps on individual components of the index.**
  - As a consequence, the present value for the variance swap buyer, the fixed variance swap payer, will increase.
  - This increase is expected to outperform the potential losses from the short variance swap positions on the individual components.

Correlation Swap – When Correlation Increases



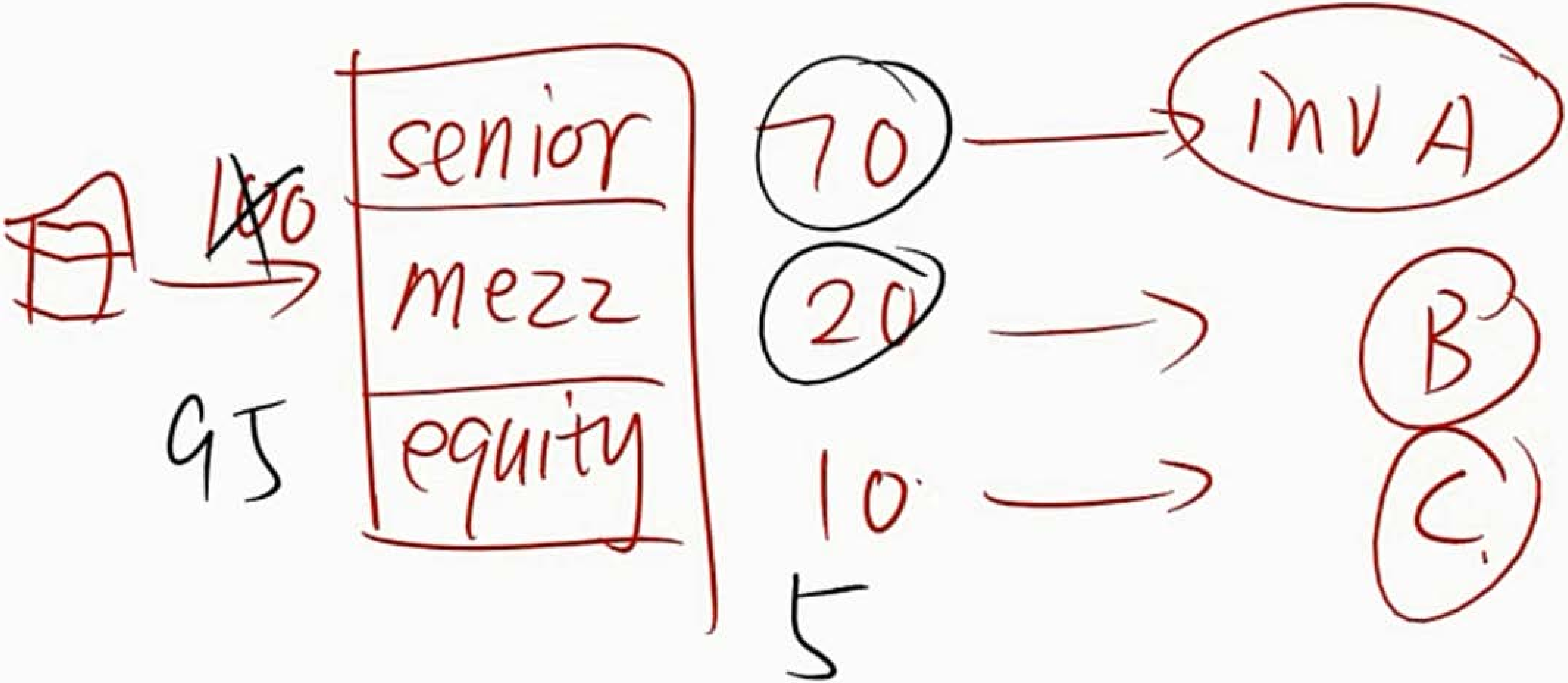
Correlation Swap – When Correlation Increases



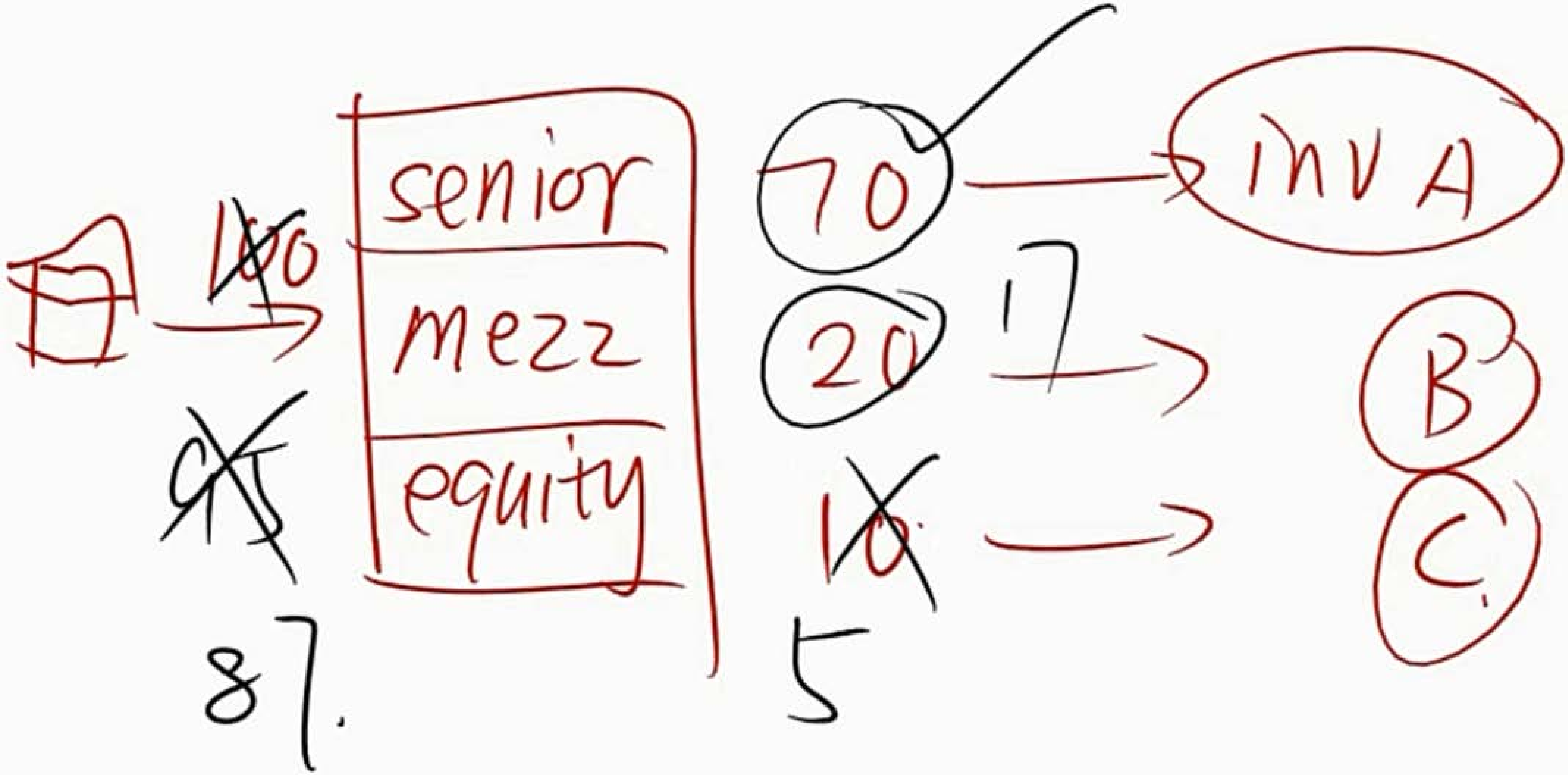
Credit Crisis Resulting from CDOs

- **In May 2005**: the equity tranche holder is exposed to the first 3% of mortgage defaults, the mezzanine tranche holder is exposed to the 3% to 7% of defaults, and so on.
- **Their strategies: short the equity tranche of the CDO and long the mezzanine tranche.**
  - Shorting the equity tranche means receiving the (high) equity tranche contract spread.
  - Going long the mezzanine tranche means paying the (fairly low) mezzanine tranche contract spread.

CDO



CDO



Credit Crisis Resulting from CDOs

- **In May 2005**: the equity tranche holder is exposed to the first 3% of mortgage defaults, the mezzanine tranche holder is exposed to the 3% to 7% of defaults, and so on.
- **Their strategies: short the equity tranche of the CDO and long the mezzanine tranche.**
  - Shorting the equity tranche means receiving the (high) equity tranche contract spread. 卖出equity CDS，收保费
  - Going long the mezzanine tranche means paying the (fairly low) mezzanine tranche contract spread. 买入mezz保险 付保费

Equity $\longrightarrow$ HF $\longrightarrow$ Mezz

保费 $\uparrow\uparrow$ 　　　　 保费、$\uparrow$

Credit Crisis Resulting from CDOs

- **When the correlations of the assets in the CDO decreased, the hedge funds lost on both positions.**
  - **The equity tranche spread increased sharply**, resulting in a paper loss.
  - The hedge funds lost on their long mezzanine tranche positions, since a lower correlation **lowers the mezzanine tranche spread**, resulting in another paper loss.

$\rho \downarrow$

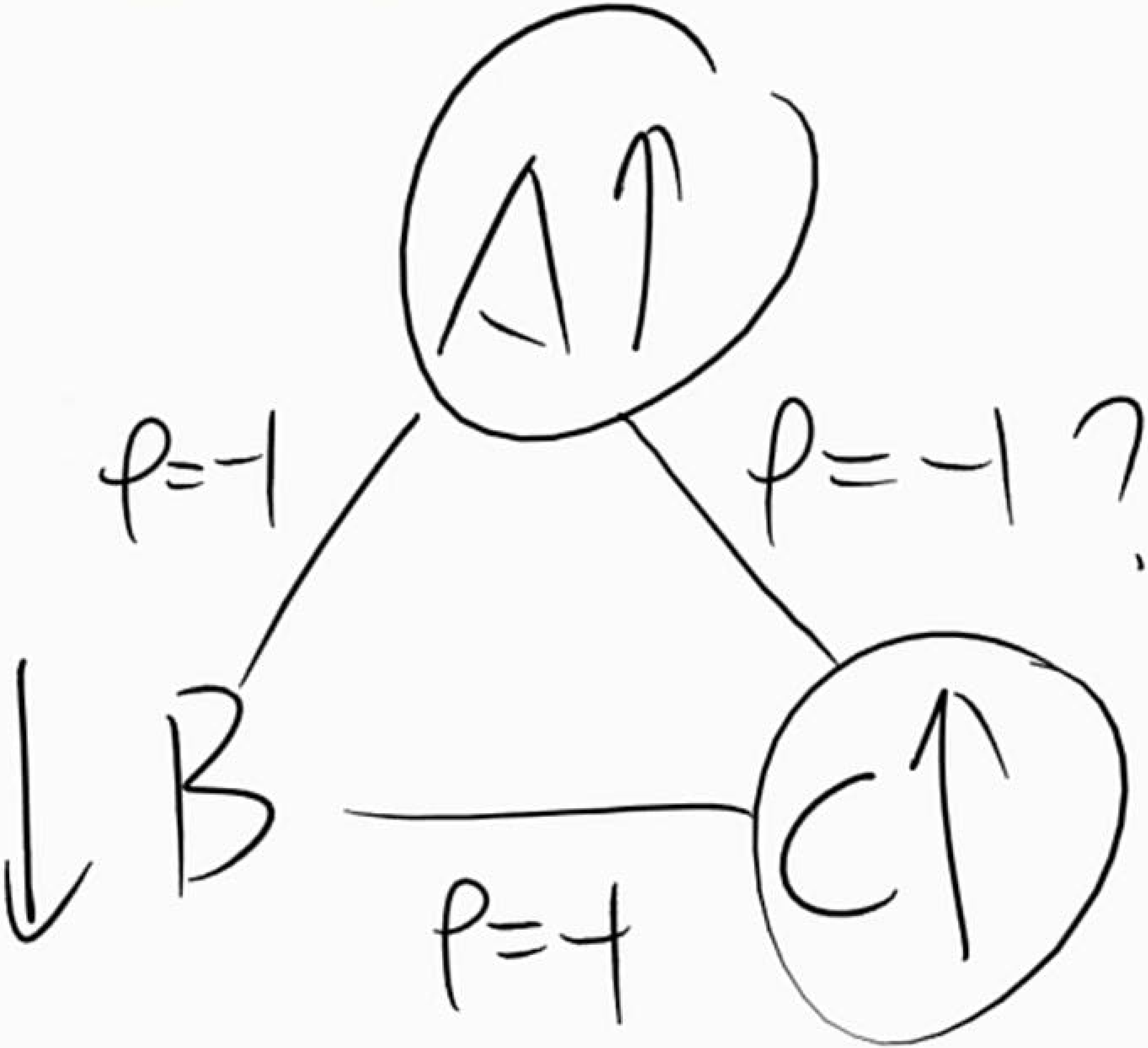
$\rho \uparrow = 1$ $\rho \downarrow$

$V_m \uparrow \downarrow$

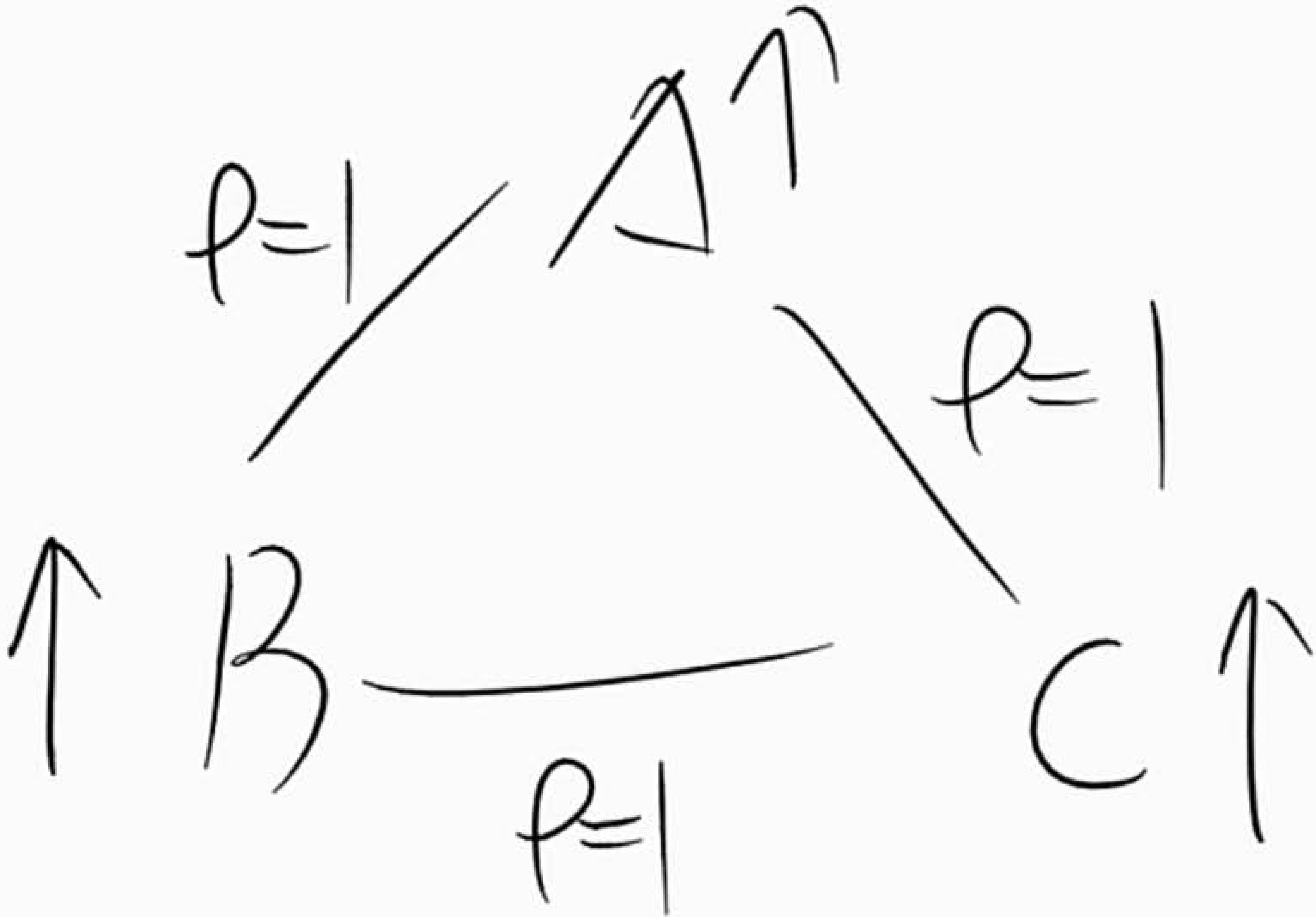
Mezz → 被保护对象

equity → 防弹衣

$V_e \downarrow \uparrow$



A↑
ρ=−1
ρ=−1?
↓B
C↑
ρ=+1


A ↑
ρ=1
ρ=1
↑ B
C ↑
ρ=1

$\rho\downarrow$:	$V_M\uparrow$	$V_e\downarrow$
risk:	$\downarrow$	$\uparrow$
premium:	$\downarrow$	$\uparrow$

Equity → HF → Mezz

保费 ↑　　保费、↓

Credit Crisis Resulting from CDOs

- **When the correlations of the assets in the CDO decreased, the hedge funds lost on both positions.**
  - **The equity tranche spread increased sharply**, resulting in a paper loss.
  - The hedge funds lost on their long mezzanine tranche positions, since a lower correlation **lowers the mezzanine tranche spread**, resulting in another paper loss.

$\rho_1 \downarrow$

Credit Crisis Resulting from CDOs

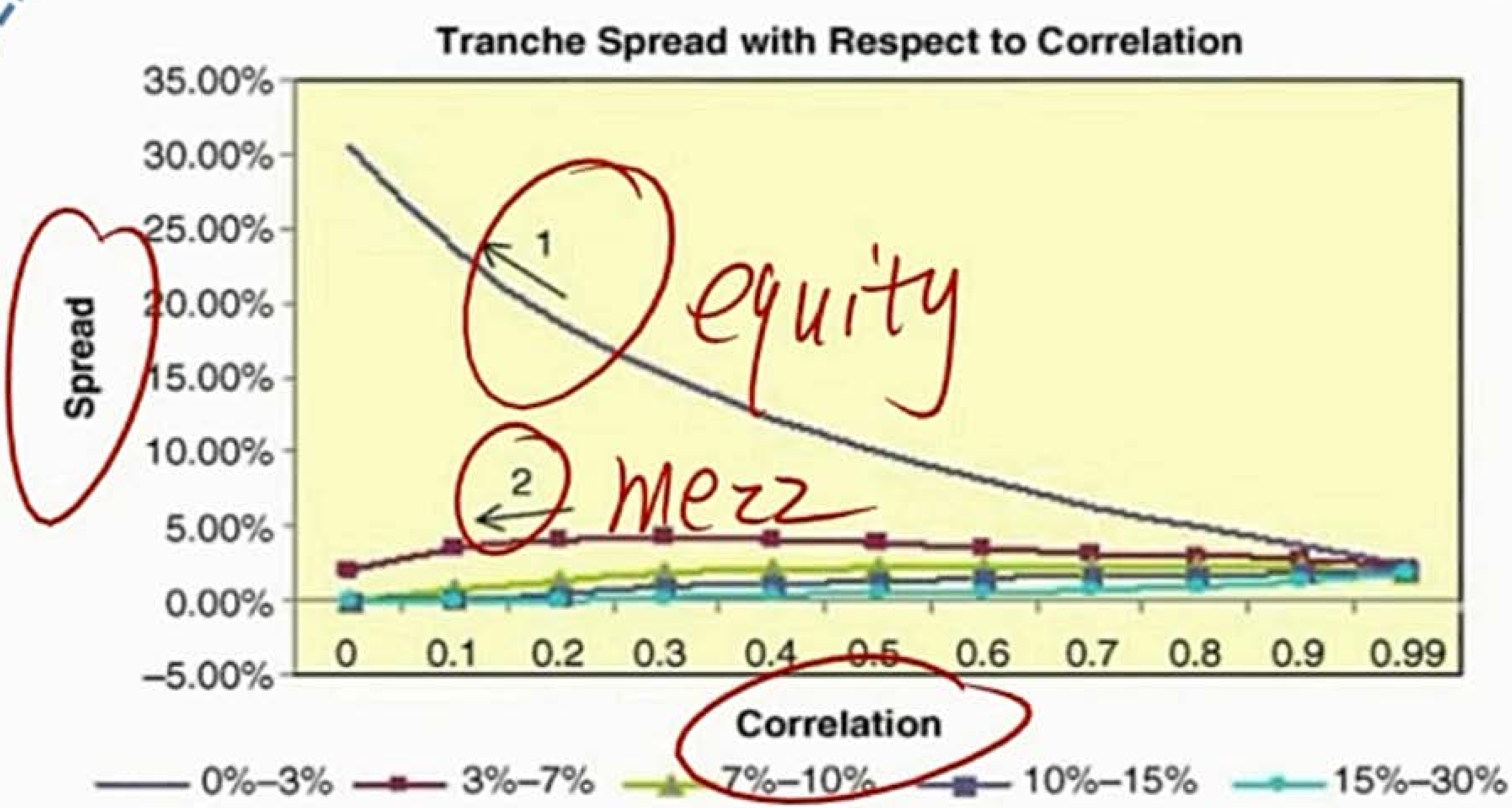


FIGURE 1.7 CDO Tranche Spread with Respect to Correlation

Credit Crisis Resulting from CDOs

- **From 2007 to 2009, default correlations of the mortgages in the CDOs increased**. This actually helped equity tranche investors. $\rho\uparrow$
  - If default correlations increase, the equity tranche spread decreases, leading to an increase in the value of the equity tranche. $V_e\uparrow$
  - However, **this increase was overcompensated by a strong increase in default probability of the mortgages**. $\uparrow PD \rightarrow V_e\downarrow$
- As a consequence, tranche spreads increased sharply, resulting in huge losses for the equity tranche investors as well as investors in the other tranches.

买入 equity tranch $\downarrow\downarrow$

Credit Crisis Resulting from CDOs

- **What is more, correlations between the tranches of the CDOs also increased during the crisis.** This had a devastating effect on the super-senior tranches.

Senior

- In normal times, these tranches were considered extremely safe since:
  - They were AAA rated.
  - They were protected by the lower tranches.
- But with the increased tranche correlation and the generally deteriorating credit market, **these super-senior tranches were suddenly considered risky and lost up to 20% of their value.**

Credit Default Swaps

AIG

- **CDSs can also be used as speculative instruments.**
  - For example, the CDS seller (i.e., the insurance seller) hopes that the insured event (e.g., default of a company or credit deterioration of the company) will not occur.
  - In this case the CDS seller keeps the CDS spread (i.e., the insurance premium) as income.
- A CDS buyer who does not own the underlying asset is speculating on the credit deterioration of the underlying asset.
- The entire global financial crisis can be summed up in one word: **Greed!**

Correlation Risk and Credit Risk

- **The default correlation within sectors is higher than between sectors.**
  - If General Motors defaults, it is more likely that Ford will default, rather than Ford benefiting from the default of its rival.
- Since the intrasector default correlations are higher than intersector default correlations, **a lender is advised to have a sector-diversified loan portfolio to reduce default correlation risk.**

行业

Correlation Risk and Credit Risk

- For most investment grade bonds, the term structure of default probabilities increases in time. Because **the longer the time horizon, the higher the probability of adverse events.**
- For a **distressed company** the immediate future is critical. If the company survives the coming problematic years, **the probability of default decreases.**

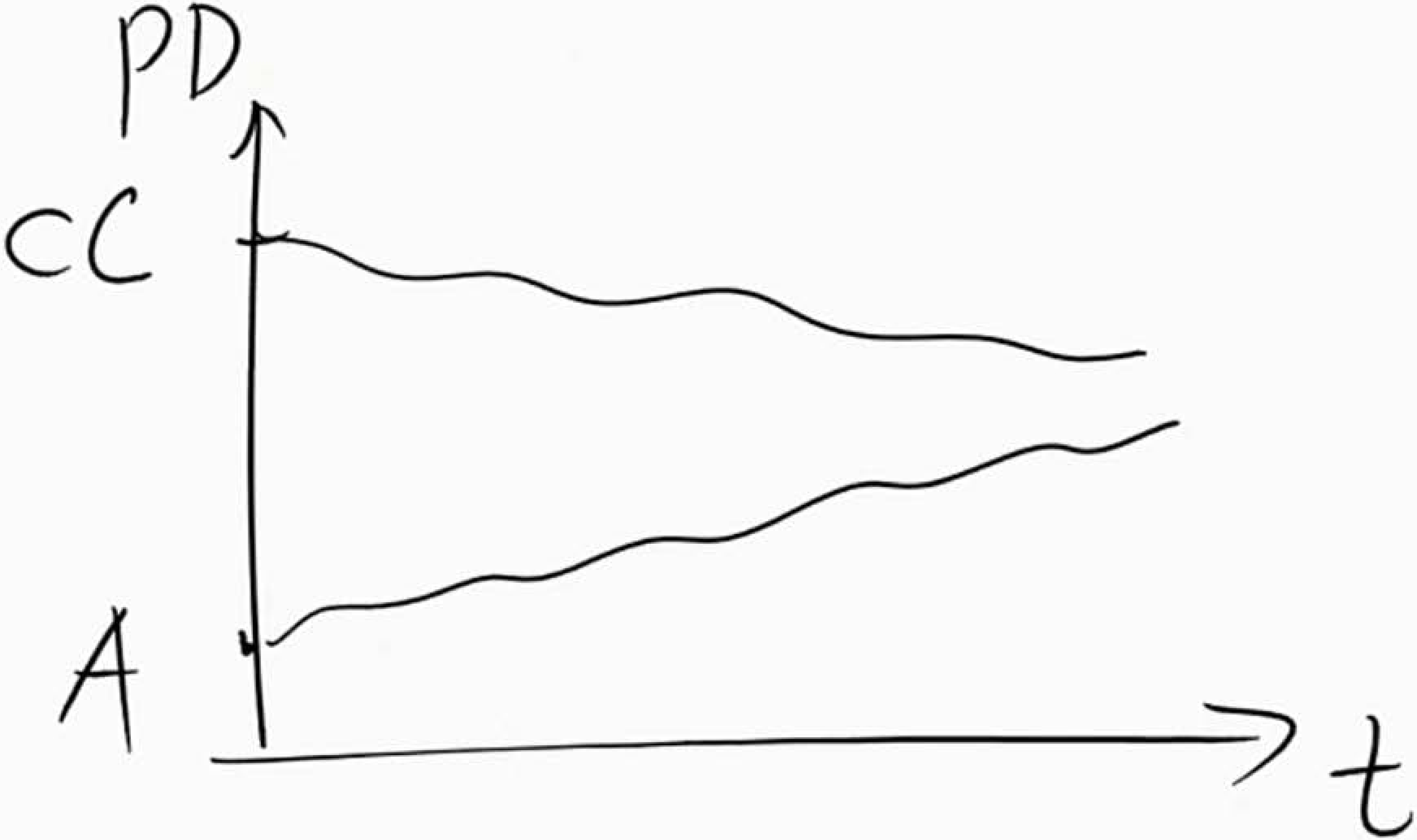
Term Structure of Default Probabilities for an A-Rated Bond and a CC-Rated Bond on 2002

PD↑

	Year									
	1	2	3	4	5	6	7	8	9	10
A	0.02%	0.07%	0.13%	0.14%	0.15%	0.17%	0.18%	0.21%	0.24%	0.25%
CC	23.83%	13.29%	10.31%	7.62%	5.04%	5.13%	4.04%	4.62%	2.62%	2.04%

Source: Moody's

PD↓↓


PD
CC
A
t

Correlation Risk and Systemic Risk

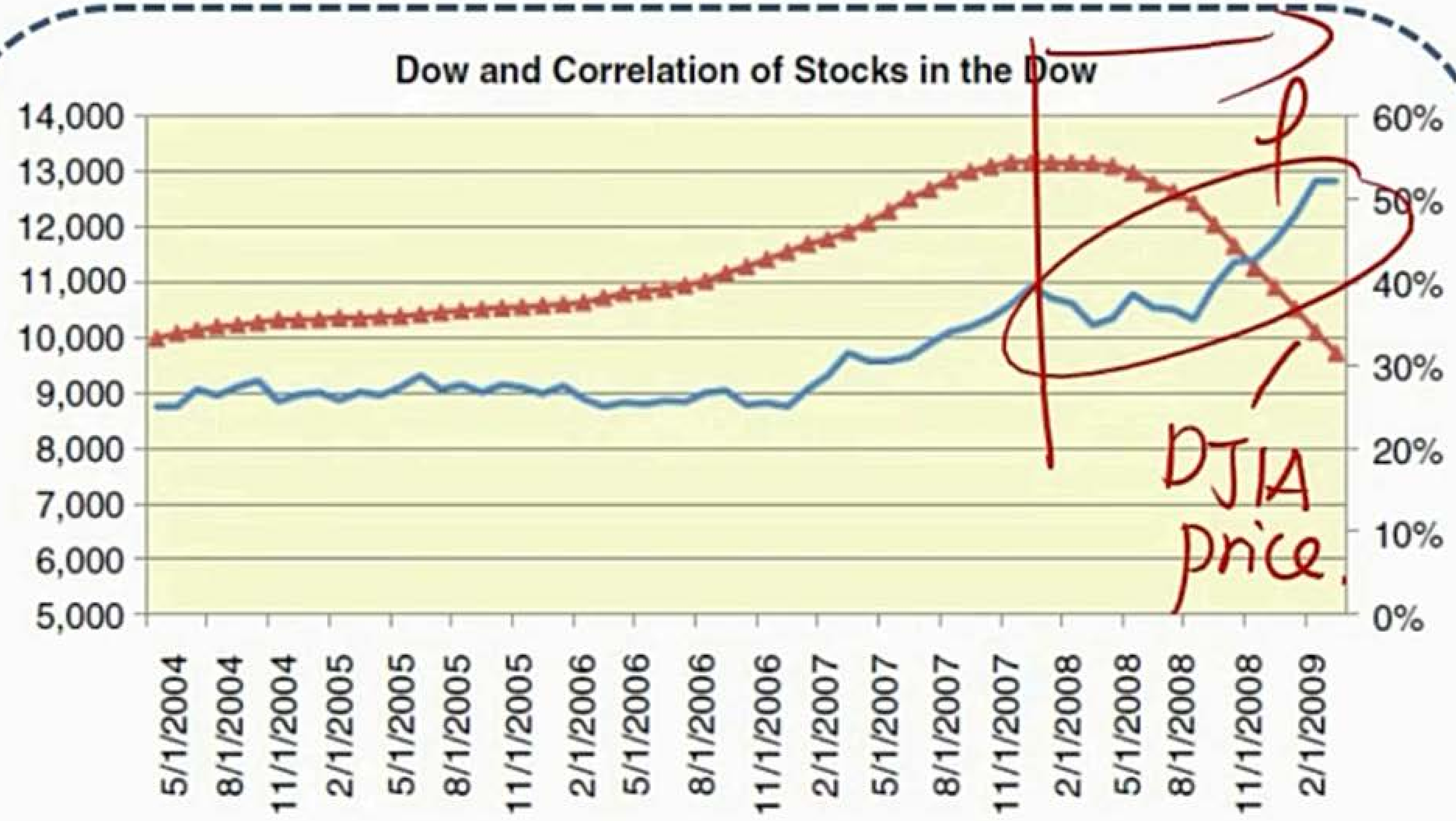


FIGURE 1.8 Relationship between the Dow (graph with triangles, numerical values on left axis) and Correlation between the Stocks in the Dow (numerical values on right axis)

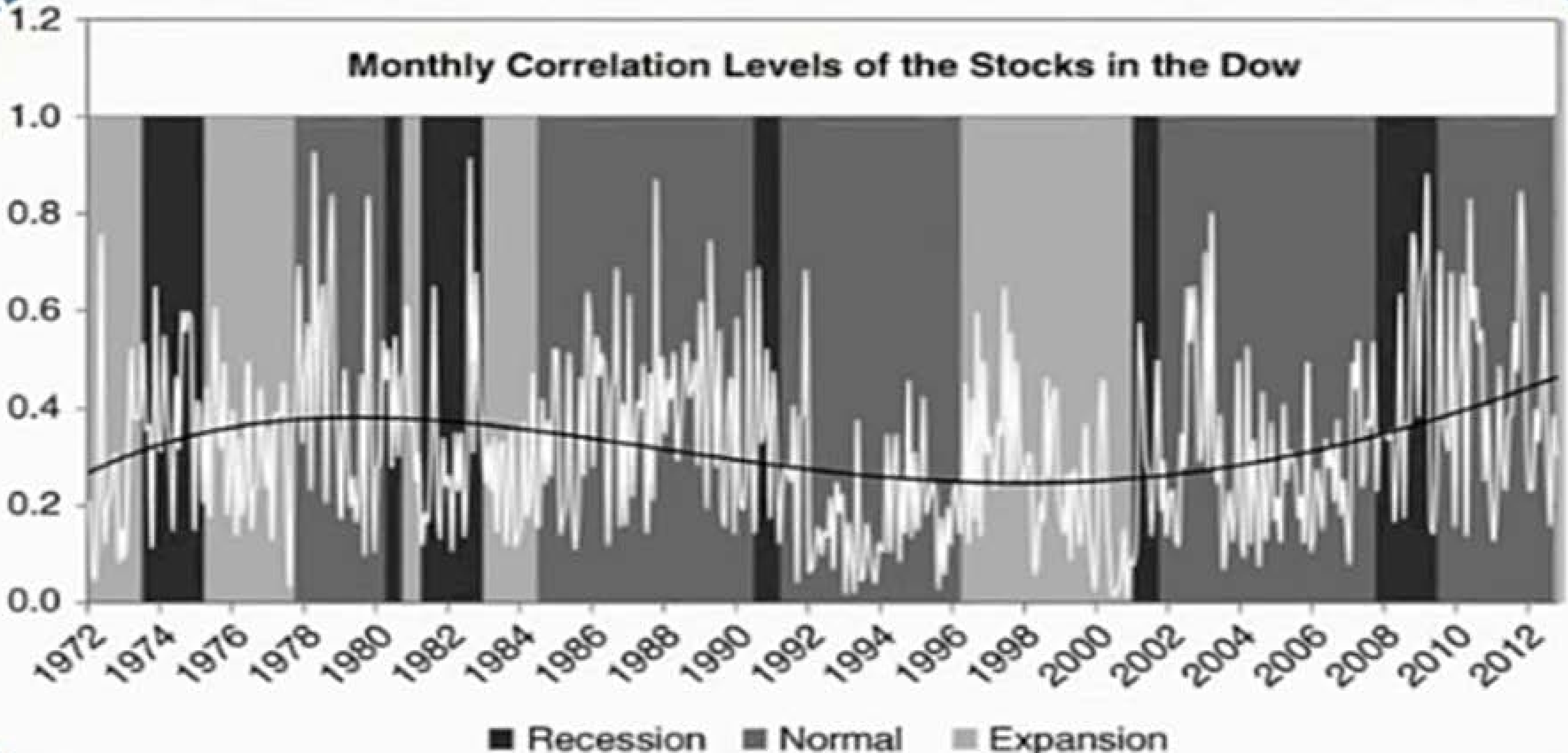
Correlation Risk and Systemic Risk

- Systemic risk and correlation risk are highly dependent.
  - The correlation in the Dow increases when the Dow increases more strongly.
  - In the time of the severe decline of the Dow from August 2008 to March 2009, a sharp increase in the correlation from non-crisis levels of on average 27% to over 50%.
- Portfolios that were considered well diversified in benign times experienced a sharp increase in correlation and hence unexpected losses due.

$P\downarrow \quad \rho\uparrow$

Equity Correlation Behaviors

- Correlation levels are lowest in strong economic growth times.
- In recessions, correlation levels typically increase.

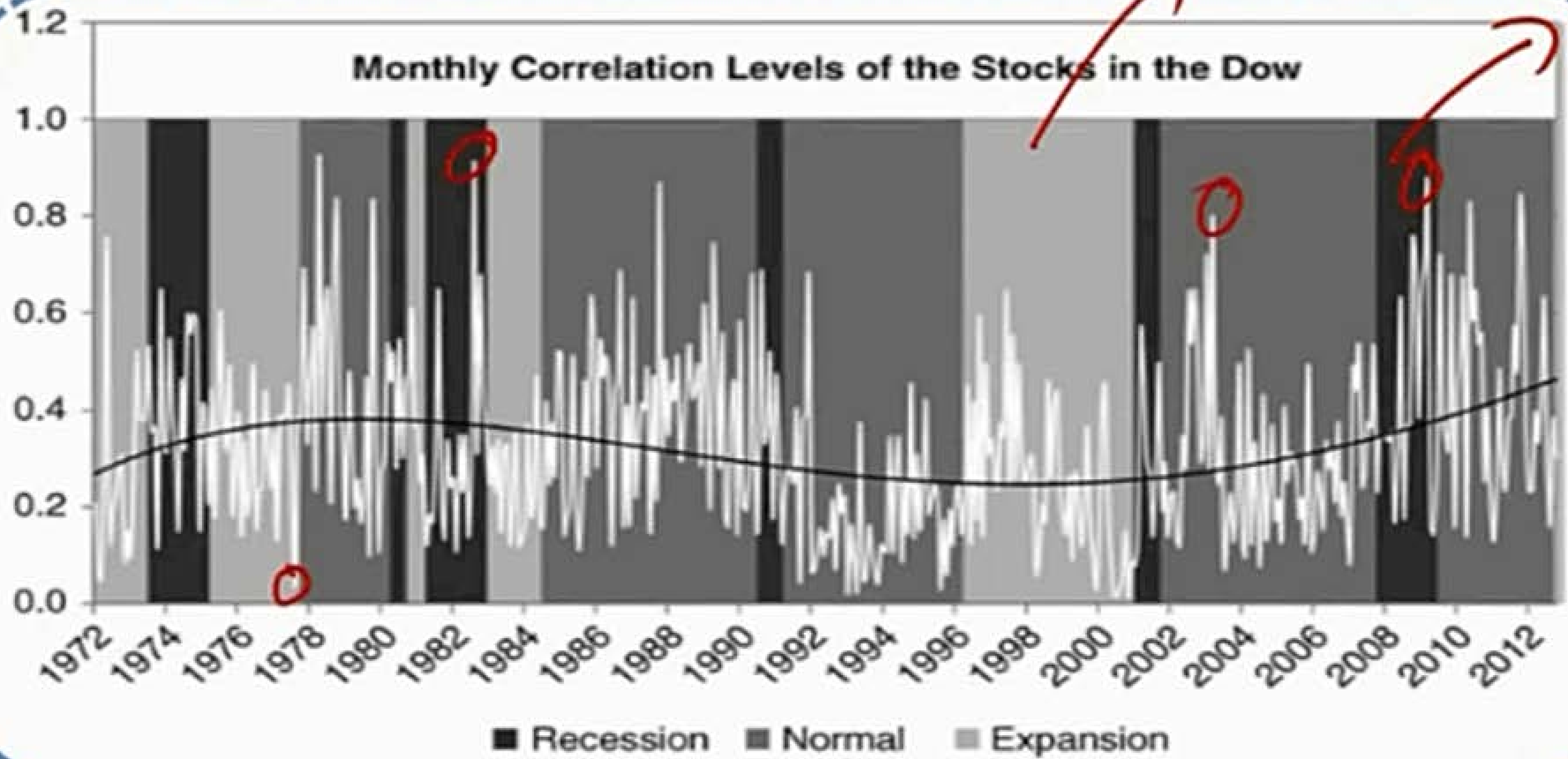


Equity Correlation Behaviors

- Correlation levels are lowest in strong economic growth times.
- In recessions, correlation levels typically increase.

Boom

recession



Equity Correlation Behaviors

- Correlation volatility is lowest in an economic expansion and highest in worse economic states.

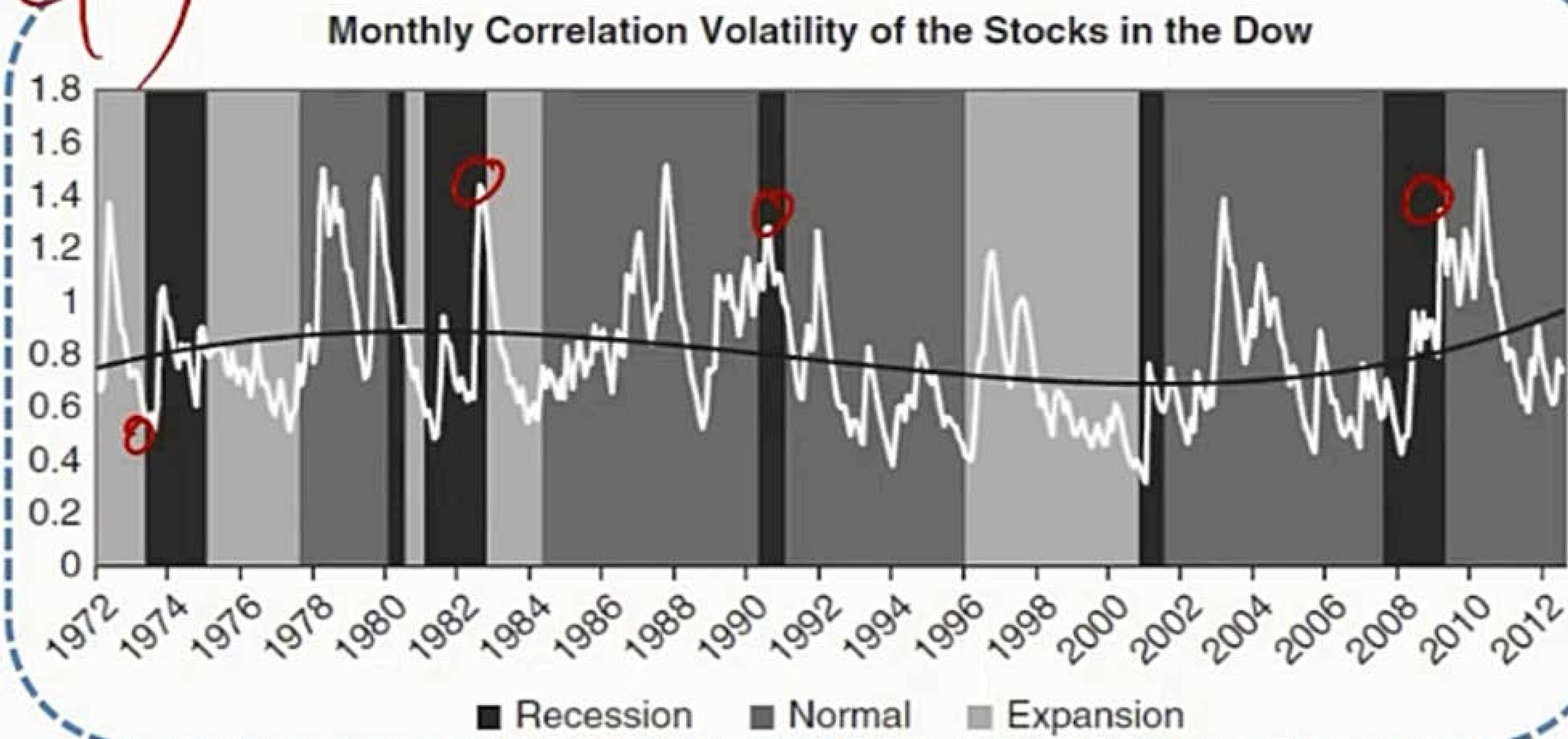
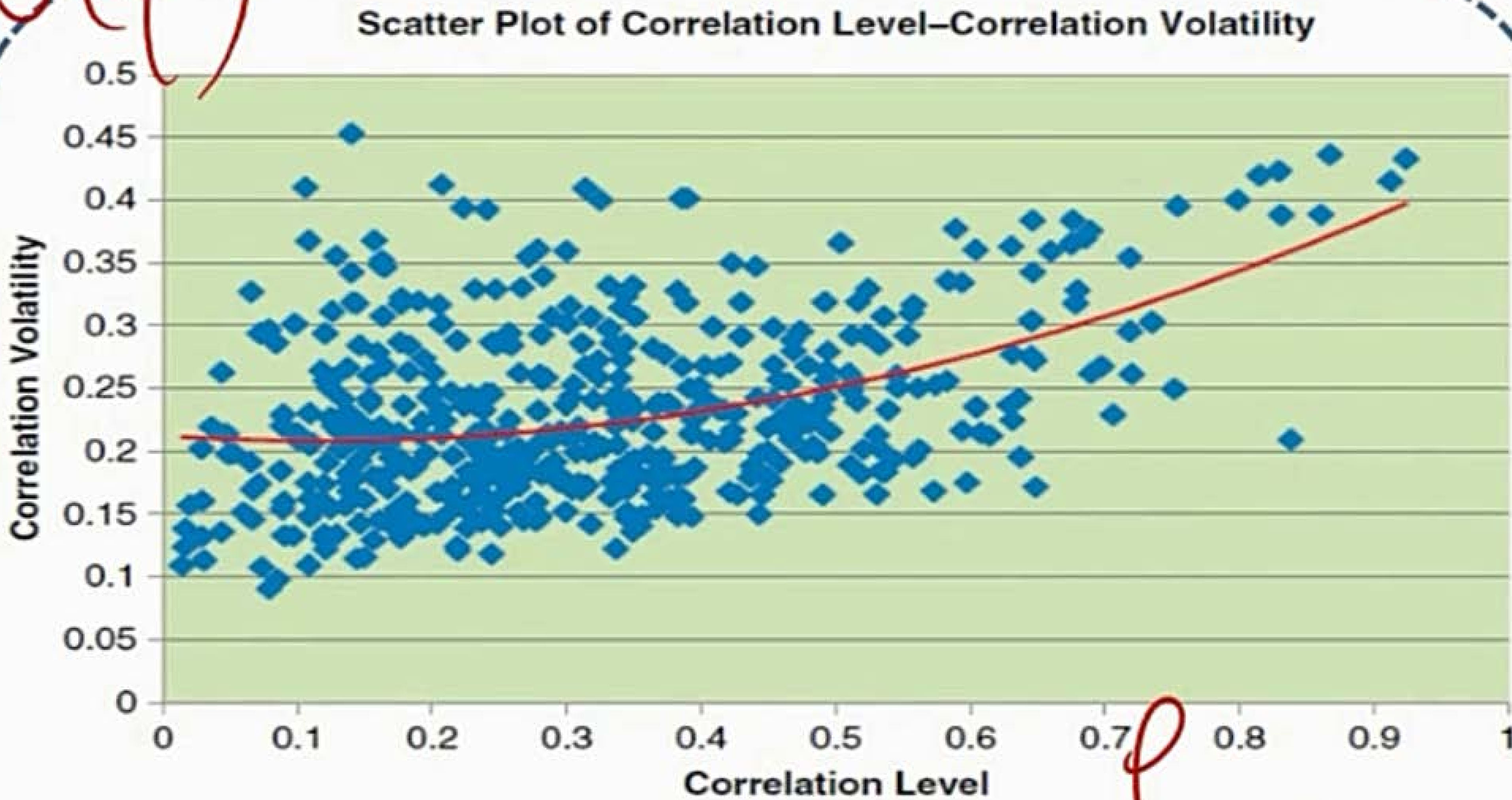


TABLE 7-1 Correlation Level and Correlation Volatility with Respect to the State of the Economy

	Correlation Level	Correlation Volatility
Expansionary period	27.46% 低	71.17% 低
Normal economic period	32.73%	83.40%
Recession	36.96% 较高	80.48% 较高

Equity Correlation Behaviors

$\sigma(\rho)$

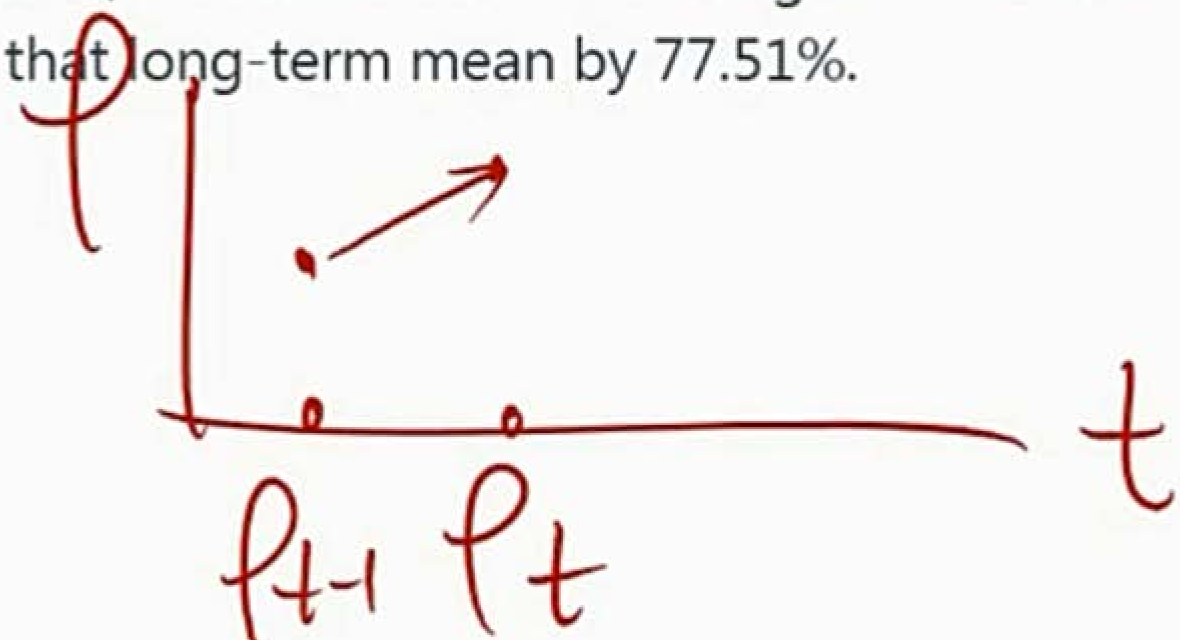


ρ

FIGURE 2.3 Positive Relationship between Correlation Level and Correlation Volatility with a Polynomial Trend Line of Order 2

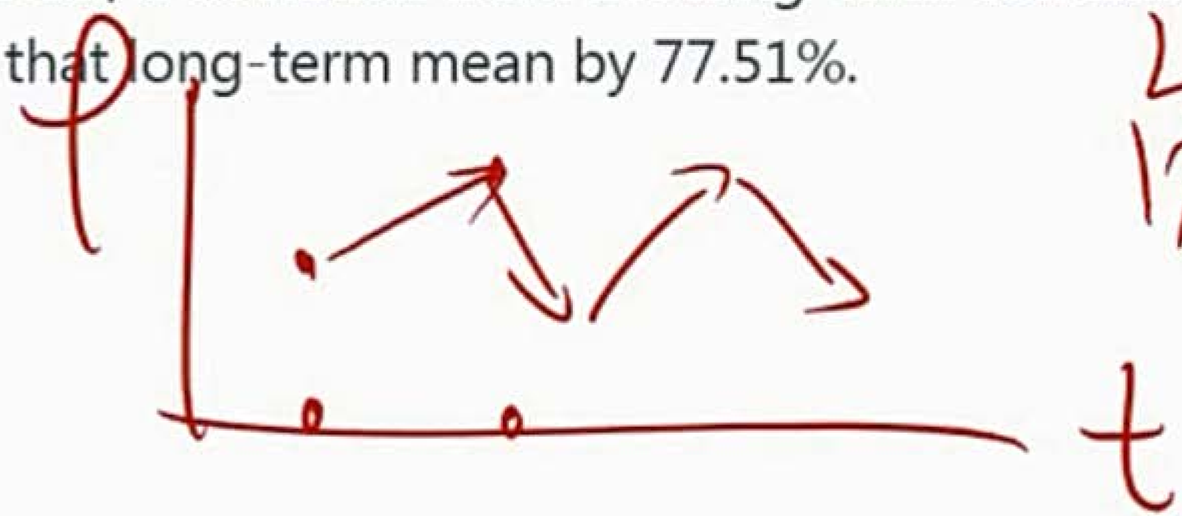
Mean reversion in correlations

- **Mean reversion** is present if there is a negative relationship between the change of a variable, $S_t - S_{t-1}$, and the variable S_{t-1}.
  - $S_t = a(\mu_s - S_{t-1}) + S_{t-1}$
  - In that case, a is called the mean reversion coefficient, or simply, mean reversion.
  - A strong mean reversion of 77.51% means that on average in every month, a deviation from the long-term correlation mean is pulled back to that long-term mean by 77.51%.



Mean reversion in correlations

- **Mean reversion** is present if there is a negative relationship between the change of a variable, $S_t - S_{t-1}$, and the variable S_{t-1}.
  - $S_t = a(\mu_s - S_{t-1}) + S_{t-1}$
  - In that case, a is called the mean reversion coefficient, or simply, mean reversion.
  - A strong mean reversion of 77.51% means that on average in every month, a deviation from the long-term correlation mean is pulled back to that long-term mean by 77.51%.



负相关

Autocorrelation in correlations

- Autocorrelation is the degree to which a variable is correlated to its past values.
  - In finance, positive autocorrelation is also termed **persistence**. In mutual fund or hedge fund performance analysis, an investor typically wants to know if an above-market performance of a fund has persisted for some time.

正相关

- **Autocorrelation is the opposite property of mean reversion.**
- Therefore, the mean reversion of 77.51% indicates the autocorrelation of 22.49%, **which add up to 1**.

Mean reversion and autocorrelation

➢ **Interpretation mean reversion and autocorrelation with AR model**

- $Y_t = b_1(\mu - Y_{t-1}) + Y_{t-1}$
  - ✓ The mean reversion in the regression is b_1.
- $Y_t = (1 - b_1)Y_{t-1} + b_1\mu$
  - ✓ The one-period lag autocorrelation of the correlation is $1-b_1$.

Is Equity Correlation an Indicator for Future Recession?

- **Before every recession a downturn in correlation volatility occurred.**
- However, the relationship between a decline in volatility and the severity of the recession is statistically non-significant.

Decrease in Correlation Volatility Preceding a Recession		
	% Change in Correlation Volatility before Recession	Severity of Recession (% Change of GDP)
1973-1974	-7.22%	-11.93%
1980	-10.12%	-6.53%
1981-1982	-4.65%	-12.00%
1990-1991	0.06%	-4.05%
2001	-5.55%	-1.80%
2007-2009	-2.64%	-14.75%

The decrease in correlation volatility is measured as a six months change of six-month moving average correlation volatility. The severity of the recession is measured as the total GDP decline during the recession.

Bond Correlations and PD Correlations

- The default probability correlation distribution for bonds is similar to the equity correlation distribution and can be replicated best with **the Johnson SB distribution**.
- However, the bond correlation distribution **shows a more normal shape** and can be best fitted with the generalized extreme value distribution.

Statistical Correlation Models

Modeling Dependence: Correlations And Copulas

Pearson Correlation Approach

$$Cov(X,Y) = \frac{1}{n-1}\sum_{t=1}^{n}(X_t - \mu_X)(Y_t - \mu_Y)$$

$$\rho_1(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - \left(E(X)\right)^2}\sqrt{E(Y^2) - \left(E(Y)\right)^2}}$$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Limitations of the Pearson Correlation Approach

1. **Linear dependencies**: do not appear often in finance.
2. **Zero correlation does not necessarily mean independence.** E.g. $Y=X^2$ (correlation=0).
3. **Linear correlation measures are natural dependence measures only if the joint distribution of the variables is elliptical**: Normal or t.
4. **The variances of the sets X and Y have to be finite.** $\sigma_x \to 0$
5. **Not invariant to transformations**: For example, the Pearson correlation between pairs X and Y is in general different from the Pearson correlation between the pairs ln(X) and ln(Y).

$\sigma_y \to \infty$

$X \overset{\rho}{—} Y \qquad \ln x \overset{\rho'}{—} \ln y$

Spearman's Rank Correlation

Performance of a Portfolio with Two Assets				
	Asset X	Asset Y	Return of Asset X	Return of Asset Y
2008	100	200		
2009	120	230	20.00%	15.00%
2010	108	460	-10.00%	100.00%
2011	190	410	75.93%	-10.87%
2012	160	480	-15.79%	17.07%
2013	280	380	75.00%	-20.83%
		Average	29.03%	20.07%

Spearman's Rank Correlation

- Order the return set pairs of X and Y with respect to the set X.
- Then derive the ranks of X_i and Y_i.
- Derive the difference of the ranks and square the difference.

Ranked Asset Returns to Derive the Spearman Correlation Coefficient

	Ranked Return of X_i	Assigned (same year) Return of Y_i	Rank of X_i	Rank of Y_i	d_i	d_i^2
2012	1 -15.79%	17.07%	1	4	-3	9
2010	2 -10.00%	100.00%	2	5	-3	9
2009	3 20.00%	15.00%	3	3	0	0
2013	4 75.00%	-20.83%	4	1	3	9
2011	5 75.93%	-10.87%	5	2	3	9
						Sum = 36

step 2

GOLDEN FUTURE

Spearman's Rank Correlation

- Order the return set pairs of X and Y with respect to the set X.
- Then derive the ranks of X_i and Y_i.
- Derive the difference of the ranks and square the difference.

Ranked Asset Returns to Derive the Spearman Correlation Coefficient

	Ranked Return of X_i	Assigned (same year) Return of Y_i	Rank of X_i	Rank of Y_i	d_i	d_i^2
2012	1 -15.79%	17.07% 4	1	4	-3	9
2010	2 -10.00%	100.00% 5	2	5	-3	9
2009	3 20.00%	15.00% 3	3	3	0	0
2013	4 75.00%	-20.83% 1	4	1	3	9
2011	5 75.93%	-10.87% 2	5	2	3	9
						Sum = 36

step2

step3

Spearman's Rank Correlation

- Order the return set pairs of X and Y with respect to the set X.
- Then derive the ranks of X_i and Y_i.
- Derive the difference of the ranks and square the difference.

Ranked Asset Returns to Derive the Spearman Correlation Coefficient

	Ranked Return of X_i	Assigned (same year) Return of Y_i	Rank of X_i −	Rank of Y_i	= d_i	^{2} = d_i^2
2012	1 -15.79%	17.07% 4	1	4	-3	9
2010	2 -10.00%	100.00% 5	2	5	-3	9
2009	3 20.00%	15.00% 3	3	3	0	0
2013	4 75.00%	-20.83% 1	4	1	3	9
2011	5 75.93%	-10.87% 2	5	2	3	9
	step 2		step3		step4	Sum = 36 step5

Spearman's Rank Correlation

- Spearman correlation coefficient

$$\rho_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$$

Step 6

- We derive:

$$\rho_S = 1 - \frac{6 \times 36}{5 \times (5^2 - 1)} = -0.8$$

$n = 5$

- Since the Spearman correlation coefficient is defined between −1 and +1, we find that the returns of assets X and Y are highly negatively correlated.

Kendall's τ

- A **concordant pair** is a pair of observations, each on two variables, $\{X_1, Y_1\}$ and $\{X_2, Y_2\}$, having the property that:
  - $\text{sgn}(X_2 - X_1) = \text{sgn}(Y_2 - Y_1)$

- A **discordant pair** is a pair of two-variable observations such that:
  - $\text{sgn}(X_2 - X_1) = -\text{sgn}(Y_2 - Y_1)$

GOLDEN FUTURE

Kendall's τ

$(1,4)(4,1): 1<4, 4>1$ 逆

- **Recall the previous example**
  - 2 concordant pairs: {(1,4),(2,5)}, {(4,1),(5,2)}
  - 8 discordant pairs: {(1,4),(4,1)}, {(1,4),(5,2)}, {(2,5),(4,1)}, {(2,5),(5,2)}, {(1,4),(3,3)}, {(2,5),(3,3)}, {(3,3),(4,1)}, {(3,3),(5,2)}
  - 0 neither concordant nor discordant pairs

1	4
2	5
3	3
4	1
5	2

$(1,4)(2,5): 1<2 \& 4<5$ 顺

$(4,1)(5,2)$ $4<5 \& 1<2$ 顺

Kendall's τ

- **Kendall's τ:**

$$\tau = \frac{\overset{2}{n_c} - \overset{8}{n_d}}{\underset{5}{n}(n-1)/2}$$

  - n_c: the number of concordant data pairs.
  - n_d: the number of discordant pairs.

- **We derive:**

$$\tau = \frac{2-8}{5 \times (5-1)/2} = -0.6$$

The Problems in Application

- **The problem with applying ordinal rank correlations to cardinal observations is that ordinal correlation are less sensitive to outliers.**
  - Let's double the outliers of the returns of asset X in Table 3.2.
  - That result in an increase of the Pearson correlation coefficient from −0.7402 to −0.6108.
- **A special problem with the Kendall t is when many non-concordant and many non-discordant pairs occur, which are omitted in the calculation.**
  - if the 10 observation pairs, four are neither concordant nor discordant, leaving just six pairs to be evaluated.

ρ

- **A special problem with the Kendall t is when many non-concordant and many non-discordant pairs occur, which are omitted in the calculation.**
    - if the 10 observation pairs, four are neither concordant nor discordant, leaving just six pairs to be evaluated.

Financial Correlation Modeling

Modeling Dependence: Correlations And Copulas

Aware of The Limitation of Financial Models

- Like Monte Carlo Simulation, we often use random models in financial modeling, since we believe they can better replicate random human behavior.
- Models need to be stress-tested. This means that extreme scenarios such as economic recessions and systemic market crashes are simulated.
- These limitations were ignored in the crisis of 2007 to 2009, when many traders and risk managers **blindly trusted the new copula correlation model**.
- When real estate prices declined sharply in 2007 to 2009, and structured products such as CDOs, which referenced mortgages, declined by 50% or more.

Copulas

- **Copula functions are designed to simplify statistical problems.** They allow the joining of multiple univariate distributions to a single multivariate distribution.
- **There exists a copula function C such that:**

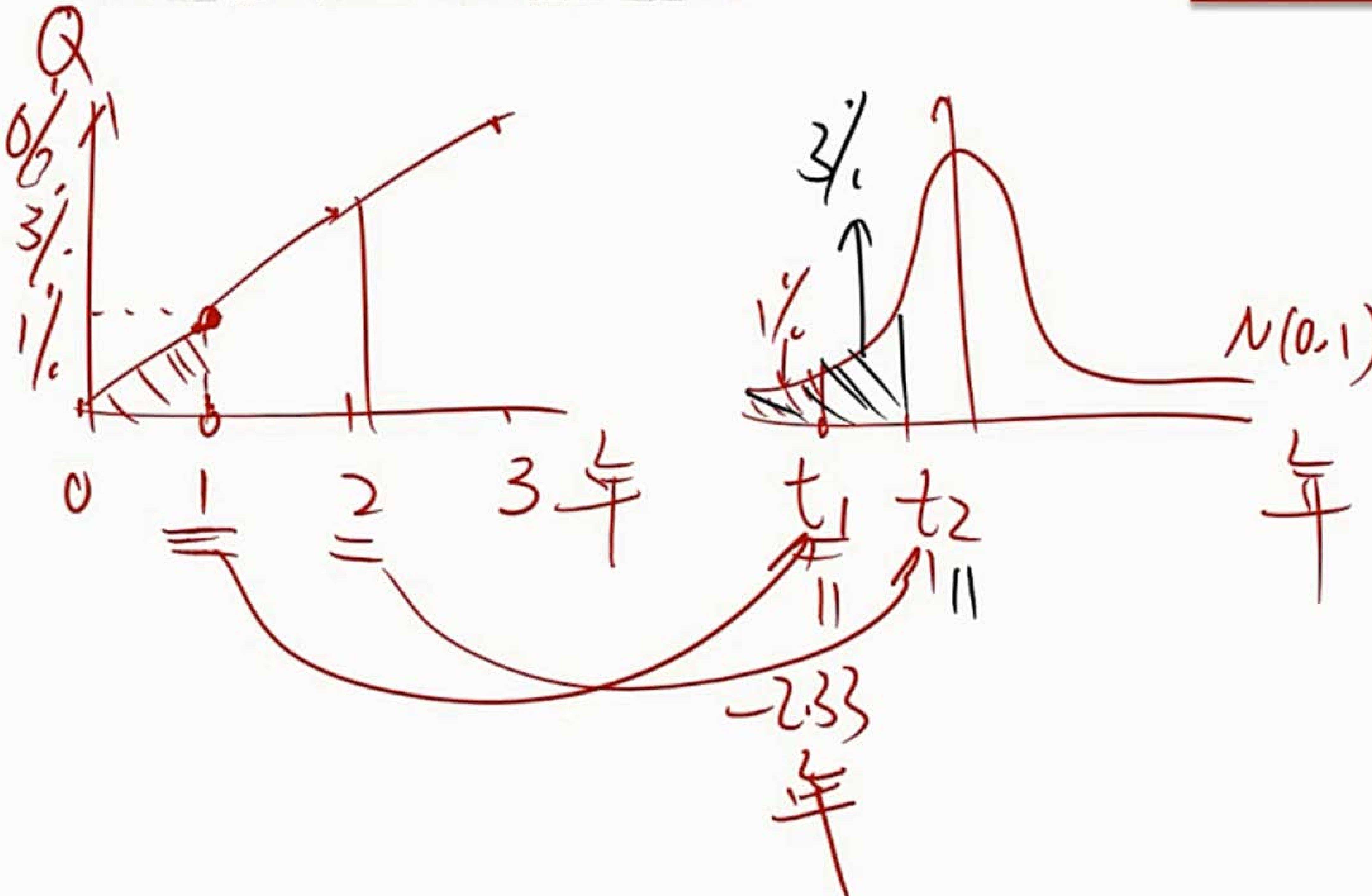
$$C[G_1(u_1),\ldots,G_n(u_n)] = \underbrace{F_n}_{1}[\underbrace{F_1^{-1}}_{2}(\underbrace{G_1(u_1)}_{3}),\ldots,F_n^{-1}(G_n(u_n));\rho_F]$$

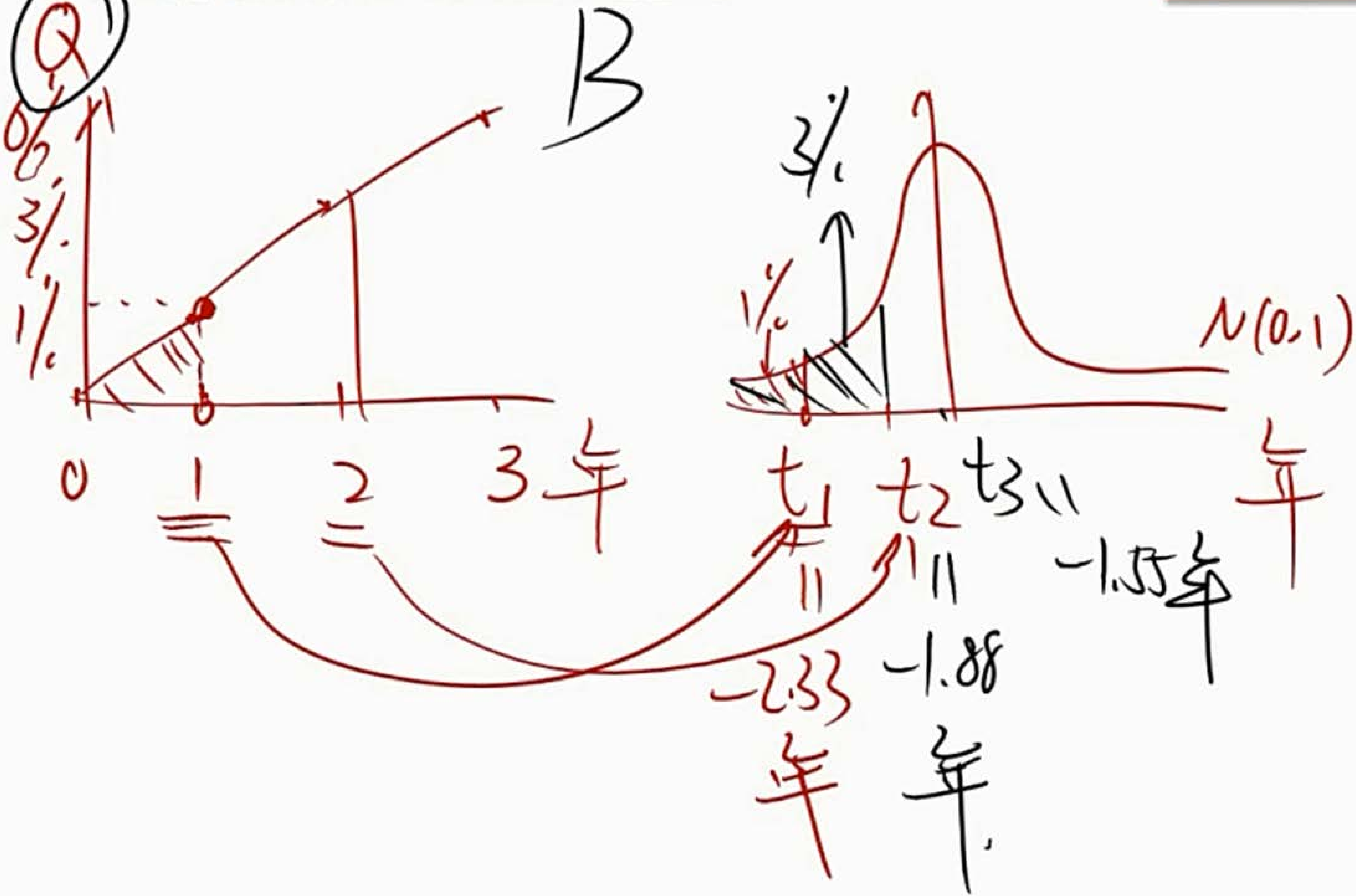
- $G_i(u_i)$: marginal uniform distributions
- F_n: the joint cumulative distribution function
- $F_i^{-1}(G_i(u_i))$: the inverse of F_i, standard normal distribution
- ρ_F: the correlation structure of F_n

David Li's Copula

Default Time t	Company B Default Probability	Company B Cumulative Default Probability $Q_B(t)$	Company Caa Default Probability	Company Caa Cumulative Default Probability $Q_{Caa}(t)$
1	6.51%	6.51%	23.83%	23.83%
2	7.65%	14.16%	13.29%	37.12%
3	6.87%	21.03%	10.31%	47.43%
4	6.01%	27.04%	7.62%	55.05%
5	5.27%	32.31%	5.04%	60.09%
6	4.42%	36.73%	5.13%	65.22%
7	4.24%	40.97%	4.04%	69.26%
8	3.36%	44.33%	4.62%	73.88%
9	2.84%	47.17%	2.62%	76.50%
10	2.84%	50.01%	2.04%	78.54%

Table title: Default Probability and Cumulative Default Probability of Company B and Caa


Q
0%
3%
1%
0
1
2
3 年
一
二
3%
1%
N(0,1)
t1
t2
年
-2.33 年


Q
B
0%
3%
1%
0
1
2
3 年
一
二
3%
1%
N(0,1)
t1
t2
t3
年
-1.55年
-1.88
-2.33
年
年

David Li's Copula

Default Probability and Cumulative Default Probability of Company B and Caa				
Default Time t	Company B Default Probability	Company B Cumulative Default Probability $Q_B(t)$	Company Caa Default Probability	Company Caa Cumulative Default Probability $Q_{Caa}(t)$
1	6.51%	6.51%	23.83%	23.83%
2	7.65%	14.16%	13.29%	37.12%
3	6.87%	21.03%	10.31%	47.43%
4	6.01%	27.04%	7.62%	55.05%
5	5.27%	32.31%	5.04%	60.09%
6	4.42%	36.73%	5.13%	65.22%
7	4.24%	40.97%	4.04%	69.26%
8	3.36%	44.33%	4.62%	73.88%
9	2.84%	47.17%	2.62%	76.50%
10	2.84%	50.01%	2.04%	78.54%

PD

Q_B

PD

Q_C

David Li's Copula

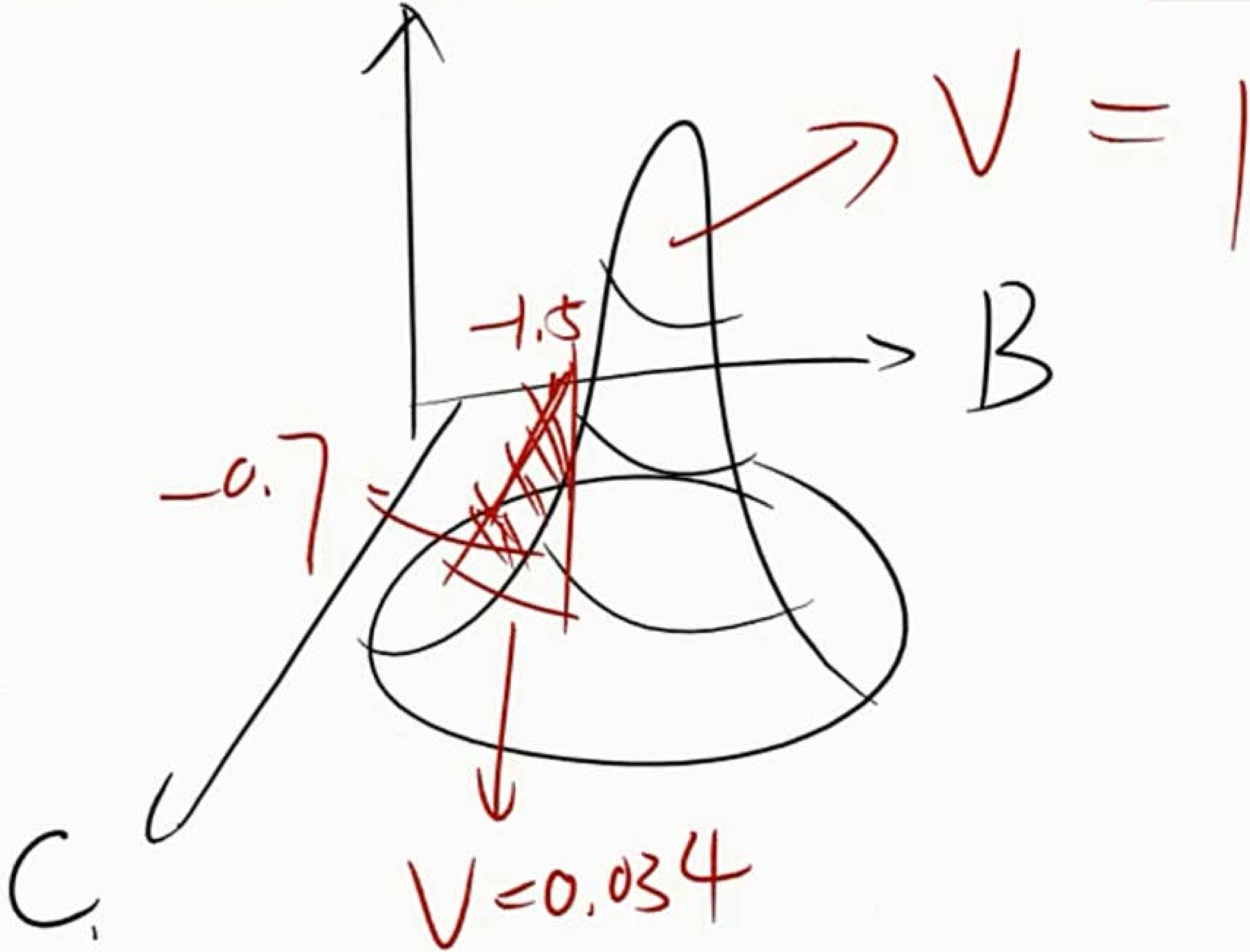
Cumulative Default Probabilities and Corresponding Standard Normal Percentiles of Company B and Caa

Default Time t	Company B Cumulative Default Probability $Q_B(t)$	Company B Cumulative Standard Normal Percentiles $N^{-1}(Q_B(t))$	Company Caa Cumulative Default Probability $Q_{Caa}(t)$	Company Caa Cumulative Standard Normal Percentiles $N^{-1}(Q_{Caa}(t))$
1	6.51%	-1.5133	23.83%	-0.7118
2	14.16%	-1.0732	37.12%	-0.3287
3	21.03%	-0.8054	47.43%	-0.0645
4	27.04%	-0.6116	55.05%	0.1269
5	32.31%	-0.4590	60.09%	0.2557
6	36.73%	-0.3390	65.22%	0.3913
7	40.97%	-0.2283	69.26%	0.5032
8	44.33%	-0.1426	73.88%	0.6397
9	47.17%	-0.0710	76.50%	0.7225
10	50.01%	0.0003	78.54%	0.7906

Q_B N_B^{-1} Q_C N_C^{-1}

$$P(B \le \underline{\underline{1}} \ \& \ C \le \underline{\underline{1}})$$

$$\iff P(B_N \le -1.51 \ \& \ C_N \le -0.7)$$


V = 1
B
-1.5
-0.7
C
V=0.034

David Li's Copula

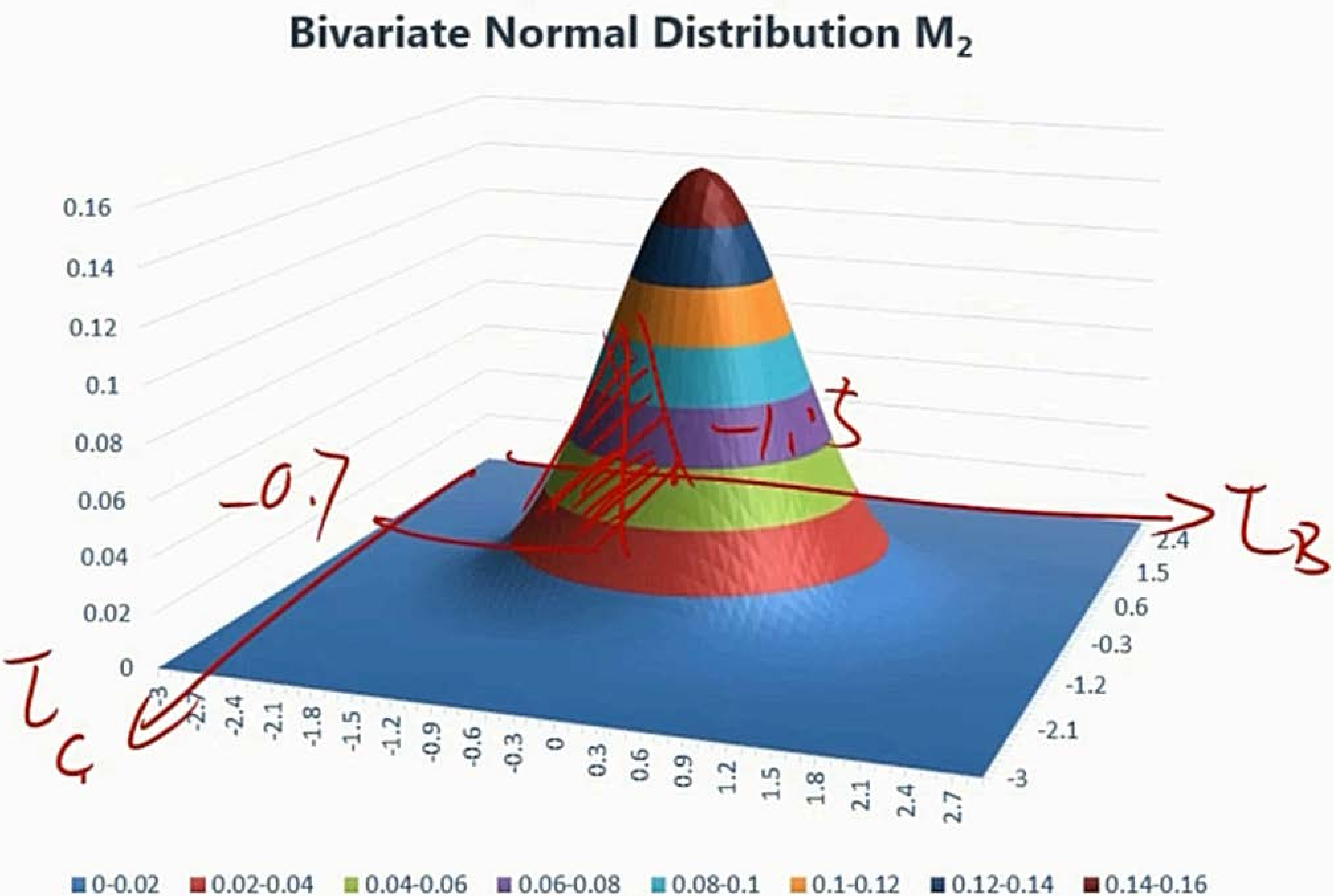
- Since we have only n = 2 companies B and Caa in our example.
- There is only one correlation coefficient ρ, not a correlation matrix ρ_M, and we have:

$$M_2\{N^{-1}[Q_B(t)], N^{-1}[Q_{Caa}(t)]; \rho\}$$

- What is the joint default probability Q of companies B and Caa in the next year assuming a one-year Gaussian default correlation of 0.4?

$$Q(\tau_B \leq 1 \cap \tau_{Caa} \leq 1)$$
$$= M(x_B \leq -1.5133 \cap x_{Caa} \leq -0.7118, \rho = 0.4)$$
$$= 3.44\%$$

David Li's Copula



- **Copula functions are designed to simplify statistical problems.** They allow the joining of multiple univariate distributions to a single multivariate distribution.

PD_B PD_C 正态 t1 t2 -- tn (B)

- **There exists a copula function C such that:**

$$C[G_1(u_1),\ldots,G_n(u_n)] = F_n[\boxed{F_1^{-1}(G_1(u_1))},\ldots,\boxed{F_n^{-1}(G_n(u_n))};\rho_F]$$

PD_B (C)

- $G_i(u_i)$: marginal uniform distributions
- F_n: the joint cumulative distribution function
- $F_i^{-1}(G_i(u_i))$: the inverse of F_i, standard normal distribution
- ρ_F: the correlation structure of F_n

Q_B

Copulas

- **Copula functions are designed to simplify statistical problems.** They allow the joining of multiple univariate distributions to a single multivariate distribution.

PD_B PD_C 正态 $t_1 t_2 \cdots t_n$ (B)

- **There exists a copula function C such that:**

$$C[G_1(u_1),\ldots,G_n(u_n)] = \boxed{F_n}[\boxed{F_1^{-1}(G_1(u_1))},\ldots,\boxed{F_n^{-1}(G_n(u_n))};\boxed{\rho_F}]$$

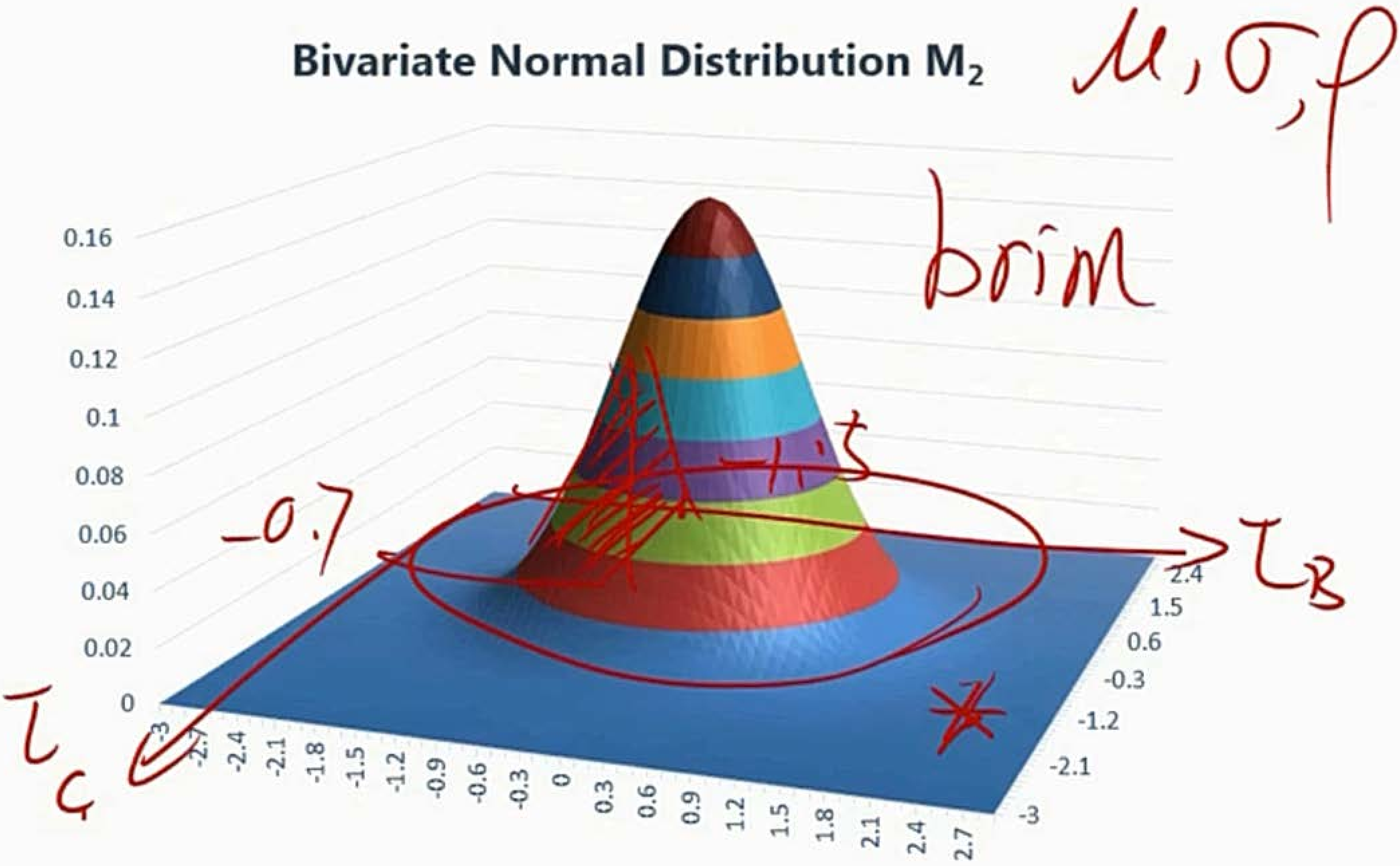
PD_B (C)

- $G_i(u_i)$: marginal uniform distributions
- F_n: the joint cumulative distribution function
- $F_i^{-1}(G_i(u_i))$: the inverse of F_i, standard normal distribution
- ρ_F: the correlation structure of F_n

Q_B

(三维正态)

Bivariate Normal Distribution M_2



Limitations of the Gaussian Copula

- **Tail Dependence: In a crisis, correlations typically increase.** Hence it would be desirable to apply a correlation model with high comovements in the lower tail of the joint distribution.

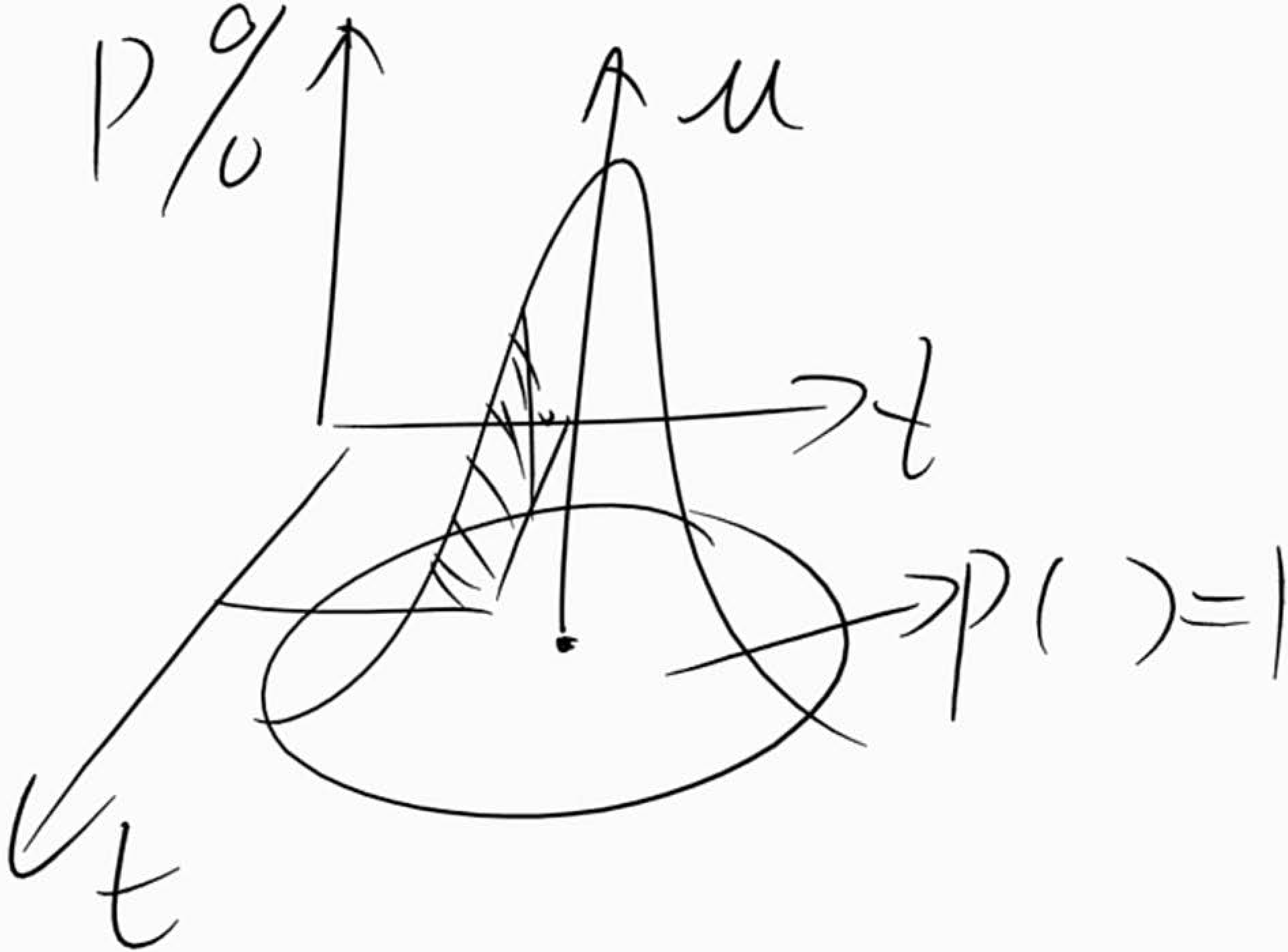
- **Copula model is static** and consequently allows only limited risk management.

ρ

$$
\begin{array}{c|ccc}
 & A & B & C \\
\hline
A & 1 & \rho_{BA} & \rho_{CA} \\
B & \rho_{AB} & 1 & \rho_{BC} \\
C & \rho_{AC} & \rho_{BC} & 1
\end{array}
$$

Principal Components Analysis

- **Principal Components Analysis (PCA)** sets up PCs with the following properties:
  - The sum of the variances of the PCs equals the sum of the variances of the individual rates. In this sense the PCs capture the volatility of this set of interest rates.
  - The PCs are uncorrelated with each other.
  - Each PC is chosen to have the maximum possible variance given all earlier PCs. In other words, the first PC explains the largest fraction of the sum of the variances of the rates; the second PC explains the next largest fraction, etc.



P%
u
t
p( )=1
t

$$f(\underline{\mu},\underline{\sigma},\boxed{\underline{\rho}}) =$$

$$f(\mu,\sigma)$$

Reading 4

Empirical Approaches to Risk Metrics and Hedges

142-203

$$\frac{\Delta P}{P} = -D \cdot P \left(\frac{\Delta Y}{Y} \right)$$

$\uparrow \quad \downarrow$

$$\Delta P \to 0$$

$$\Delta P_u = -D_u \cdot P_u \cdot \Delta Y$$

$$\Delta P_H$$

$$\Delta P_u + \Delta P_H = 0 = \Delta P_{ext}$$

$$D_u \cdot P_u \cdot \underbrace{\Delta y}_{\Delta y_1} + n\, D_H \cdot P_H \cdot \underbrace{\Delta y}_{\Delta y_2} = 0$$

The Drawbacks of DV01-neutral Hedge

- **The drawbacks of DV01-neutral hedge**

ΔY

  - This is a single-factor model (where the single factor is yield to maturity), the assumption is that movements of the entire structure can be described by a one interest rate factor. ΔY
  - By neglecting curve risk and simplistically assuming parallel shifts in the rate curve, the DV01-hedge is not necessarily a realistic hedge.

ΔY_u ΔY_H

$8\% - 3\% = 5\%$

nominal inf real

real + inf = nominal

$\%Y_R \neq \%Y_N$

- The trader plans to short \$100 million of the (nominal) 35/8 s of August 15, 2019, and, against that, to buy some amount of the TIPS 17/8 s of July 15, 2019.

real
抗通胀

Bond	Yield (%)	DV01
TIPS 1 7/8	1.237	0.081
T-Bond 3 5/8	3.275	0.067

$$F^R * \frac{0.081}{100} = 100mm * \frac{0.067}{100}$$

$$F^R = \$100mm * \frac{0.067}{0.081} = \$82.7mm$$

Traditional DV01 Hedge

Treasury: δy_N

➢ The trader plans to short \$100 million of the (nominal) 35/8 s of August 15, 2019, and, against that, to buy some amount of the TIPS 17/8 s of July 15, 2019.

$$D_N P_N = n D_R P_R$$

real

抗通胀

δy_R

Bond	Yield (%)	DV01
TIPS 1 7/8	1.237	0.081
T-Bond 3 5/8	3.275	0.067

$$F^R * \frac{0.081}{100} = 100mm * \frac{0.067}{100}$$

$$F^R = \$100mm * \frac{0.067}{0.081} = \$82.7mm$$

$$DV01 = \boxed{D \cdot P \cdot 0.0001}$$

$$D \cdot P = DV01 \times 10{,}000$$

Traditional DV01 Hedge

Treasury: Δy_N

- The trader plans to short \$100 million of the (nominal) 35/8 s of August 15, 2019, and, against that, to buy some amount of the TIPS 17/8 s of July 15, 2019.

$D_N P_N = n D_R P_R$

real

抗通胀

Δy_R

	Bond	Yield (%)	DV01
H	TIPS 1 7/8	1.237	0.081
U	T-Bond 3 5/8	3.275	0.067

$$F^R * \frac{0.081}{100} = 100mm * \frac{0.067}{100}$$

$$F^R = \$100mm * \frac{0.067}{0.081} = \$82.7mm$$

$100m \times \frac{0.067}{100} = n \times \frac{0.081}{100}$

$n = 82.7m. (H)$

Regression Hedge

$\Delta Y_N = \Delta Y_R$

➤ **A regression hedge**

- **For example**, in a trade where the trader sells T-Bond and buy TIPS (Treasury Inflation Protected Security), there is a dispersion of the change in the nominal yield for a given change in the real yield.
- with respect to improving on the DV01 hedge, We can estimate the average change in the nominal yield for a given change in the real yield and adjust the DV01 hedge accordingly.
- We can used the Least-squares regression analysis method to adjust the DV01 hedge.

Regression Hedge

- The trader has doubts about this choice because changes in yields on TIPS and nominal bonds may very well not be one-for-one.
  - For example, a five basis-point change in the yield of the TIPS does not imply a unique change in the nominal yield, not even an average change of five basis points.
- **This lack of a one-to-one yield relationship calls the DV01 hedge into question.**

Regression Hedge

➢ The trader can estimate the average change in the nominal yield for a given change in the real yield and adjust the DV01 hedge accordingly.

- For example, the nominal yield in the data changes by **1.0189** basis points per basis-point change in the real yield.

$$F^R \times \frac{0.081}{100} = 100mm \times \frac{0.067}{100} \times 1.0189$$

$$F^R = \$100mm \times \frac{0.067}{0.081} \times 1.0189 = \$84.3mm$$

$$\Delta y_N = \Delta y_R \times 1.0189$$

$$\Delta Y_N \, D_N P_N = x \cdot D_R P_R \Delta Y_R$$

$$\Delta Y_N = \Delta Y_R \times 1.0189$$

$$D_N P_N \times 1.0189 = x \cdot D_R \cdot P_R$$

$$x = \boxed{\frac{D_N P_N}{D_R P_R}} \times 1.0189$$

Regression Hedge

- Let Δy_t^N and Δy_t^R be the changes in the yields of the nominal and bonds.

$$\Delta y_t^N = \alpha + \beta \Delta y_t^R + \varepsilon_t$$

  - Δy_t^N = changes in the nominal yield
  - Δy_t^R = changes in the real yield

- Least-squares estimation of α and β finds the estimates $\hat{\alpha}$ and $\hat{\beta}$ that minimize the sum of the squares of the realized error terms over the observation period.

$$\sum_t \hat{e}_t^2 = \sum_t (\Delta y_t^N - \hat{\alpha} - \widehat{\beta \Delta y_t^R})^2$$

Regression Hedge

- Denoting the face amounts of the real and nominal bonds by F^R and F^N and their DV01s by $DV01^R$ and $DV01^N$, the regression-based hedge, characterized earlier as the DV01 hedge adjusted for the average change of nominal yields relative to real yields, can be written as follows:

$$F^R = -F^N \times \frac{DV01^N}{DV01^R} \times \hat{\beta}$$

Regression Hedge

- The whole point of the regression-based hedge is that **the risks of the two securities cannot properly be measured**.
- The market maker in question has bought or received fixed in relatively **illiquid 20-year swaps** from a customer and needs to hedge the resulting interest rate exposure.
- Immediately paying fixed or **selling 20-year swaps** would sacrifice too much if not all of spread paid by the customer, **so the market maker chooses instead to sell a combination of 10- and 30-year swaps.**
- Two variable regression model to **describe the relationship between changes in 20-year swap rates and changes in 10-and 30-year swap rates:**

$$\Delta y_t^{20} = \alpha + \beta^{10}\Delta y_t^{10} + \beta^{30}\Delta y_t^{30} + \varepsilon_t$$

Principal Components Analysis

- **How to choose the right regressors?**
- **Principal component analysis** is a powerful statistical tool that can help solve the curse of dimensionality.

Principal Components Analysis

- **Principal Components Analysis (PCA)** sets up PCs with the following properties:

SVD

  - The sum of the variances of the PCs equals the sum of the variances of the individual rates. In this sense the PCs capture the volatility of this set of interest rates.
  - The PCs are uncorrelated with each other.
  - Each PC is chosen to have the maximum possible variance given all earlier PCs. In other words, the first PC explains the largest fraction of the sum of the variances of the rates; the second PC explains the next largest fraction, etc.

Principal Components Analysis

- PCs of rates are particularly useful because of an empirical regularity:
  - The **sum of the variances of the first three PCs** is usually quite close to the sum of variances of all the rates.
  - Hence, rather than describing movements in the term structure by describing the variance of each rate and all pairs of correlation, **one can simply describe the structure and volatility of each of only three PCs**.

The Science of Term Structure Models

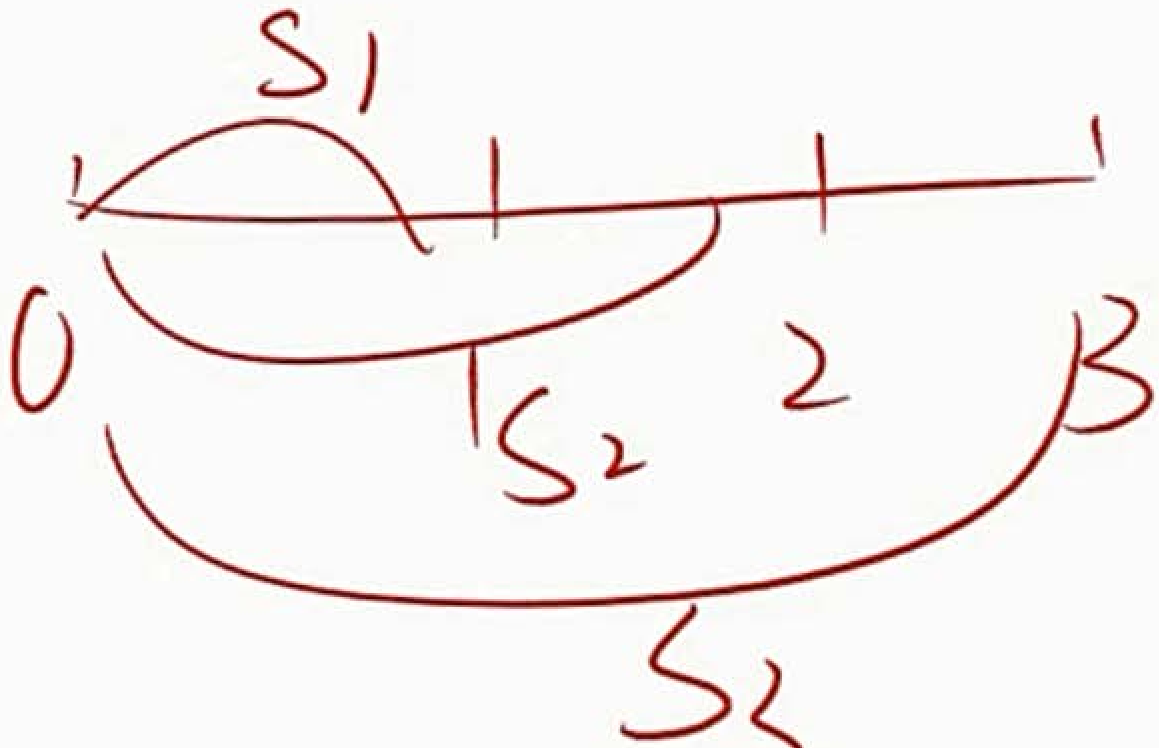
Term Structure Models of Interest Rates

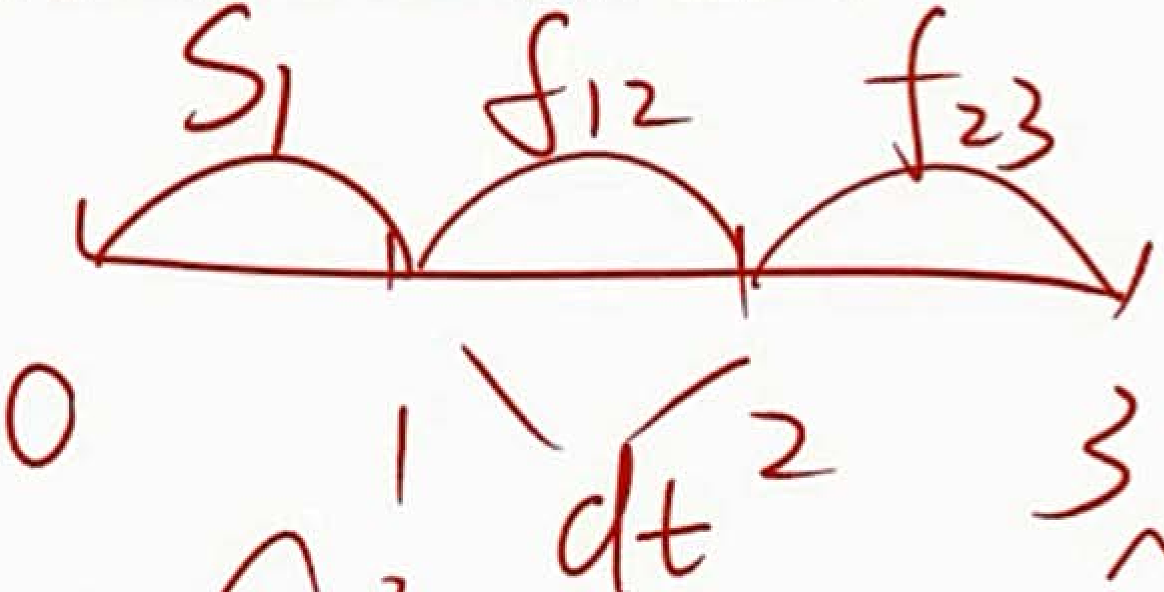
Binominal Interest Rate Tree

➢ **Interest Rate Tree (Binominal) Model**

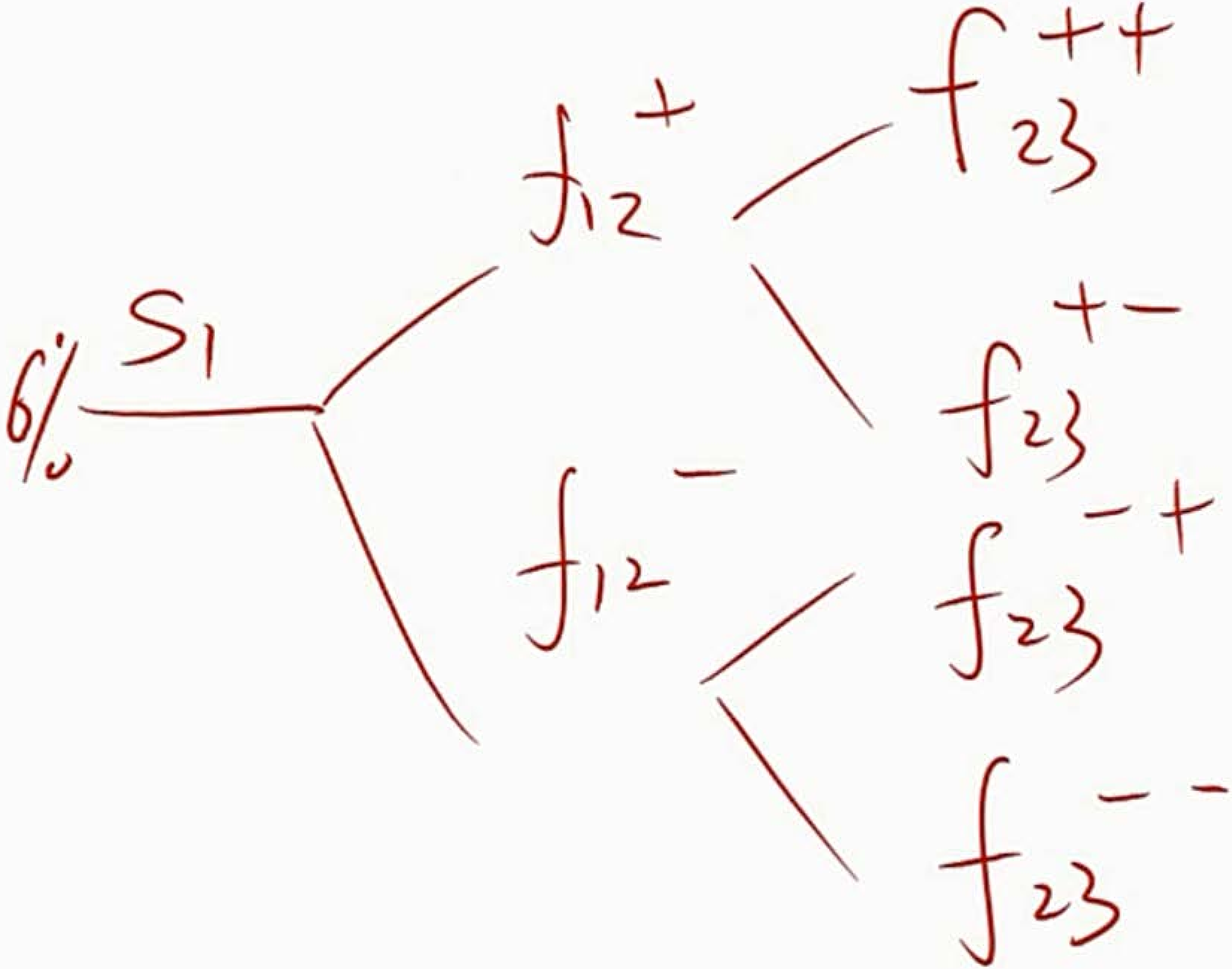
- The binomial interest rate model is used to value bonds with embedded options.
- This very strong assumption is depicted by means of a binomial tree, where "binomial" means that only two future values are possible.
- A binomial model is a model that assumes that interest rate can take only one of two possible values in the next period.
- This interest rate model makes assumptions about interest rate volatility, along with a set of paths that interest rates may follow over time. This set of possible interest rate paths is referred to as an interest rate tree.

$$P = \frac{c}{1+S_1} + \frac{c}{(1+S_2)^2} + \frac{c+Par}{(1+S_3)^3}$$

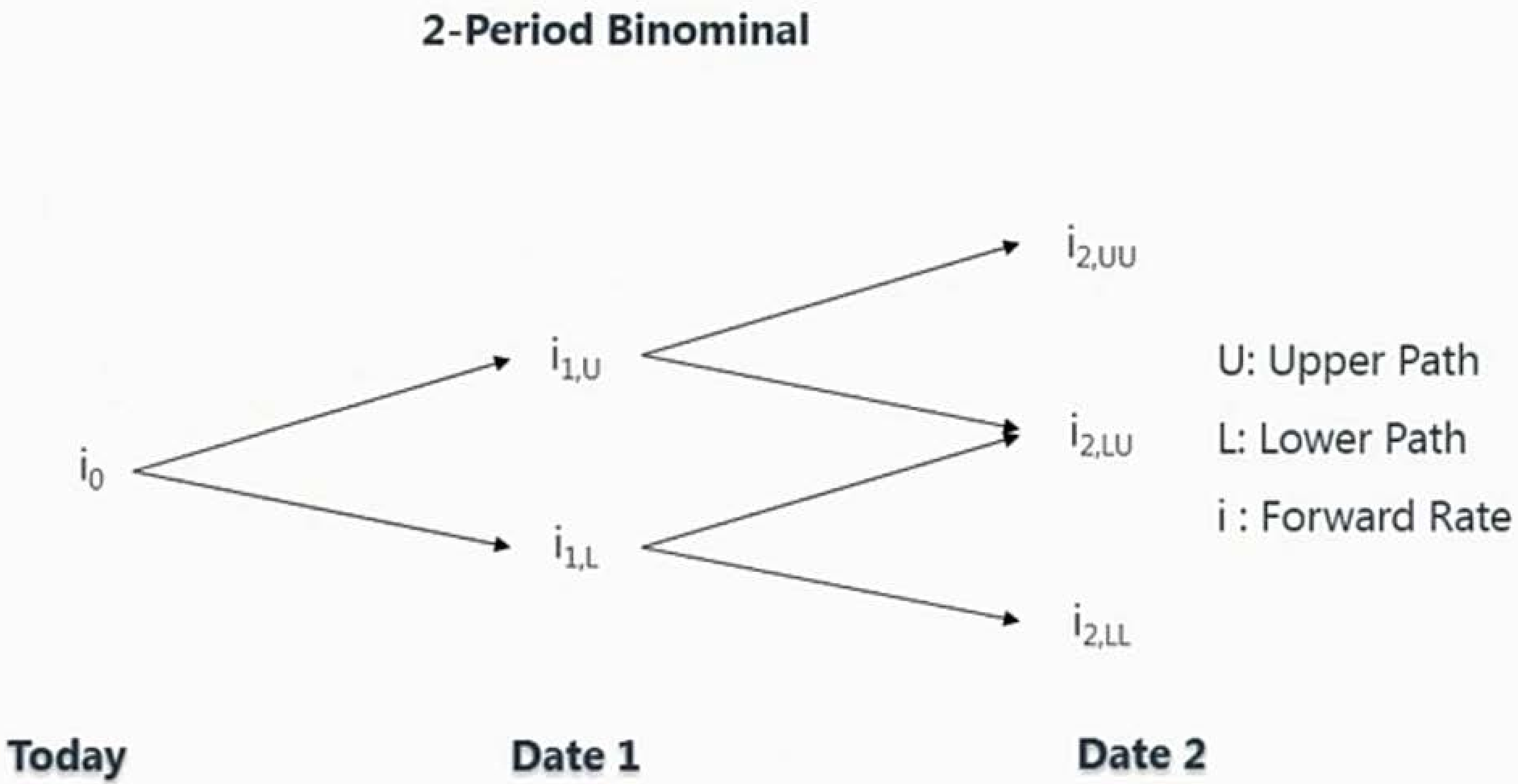




$$(1+\hat{S}_3)^3 = (1+S_1)(1+\hat{f}_{12})(1+\hat{f}_{23})$$


6%
S_1
f_{12}^{+}
f_{12}^{-}
f_{23}^{++}
f_{23}^{+-}
f_{23}^{-+}
f_{23}^{--}

Binominal Interest Rate Tree

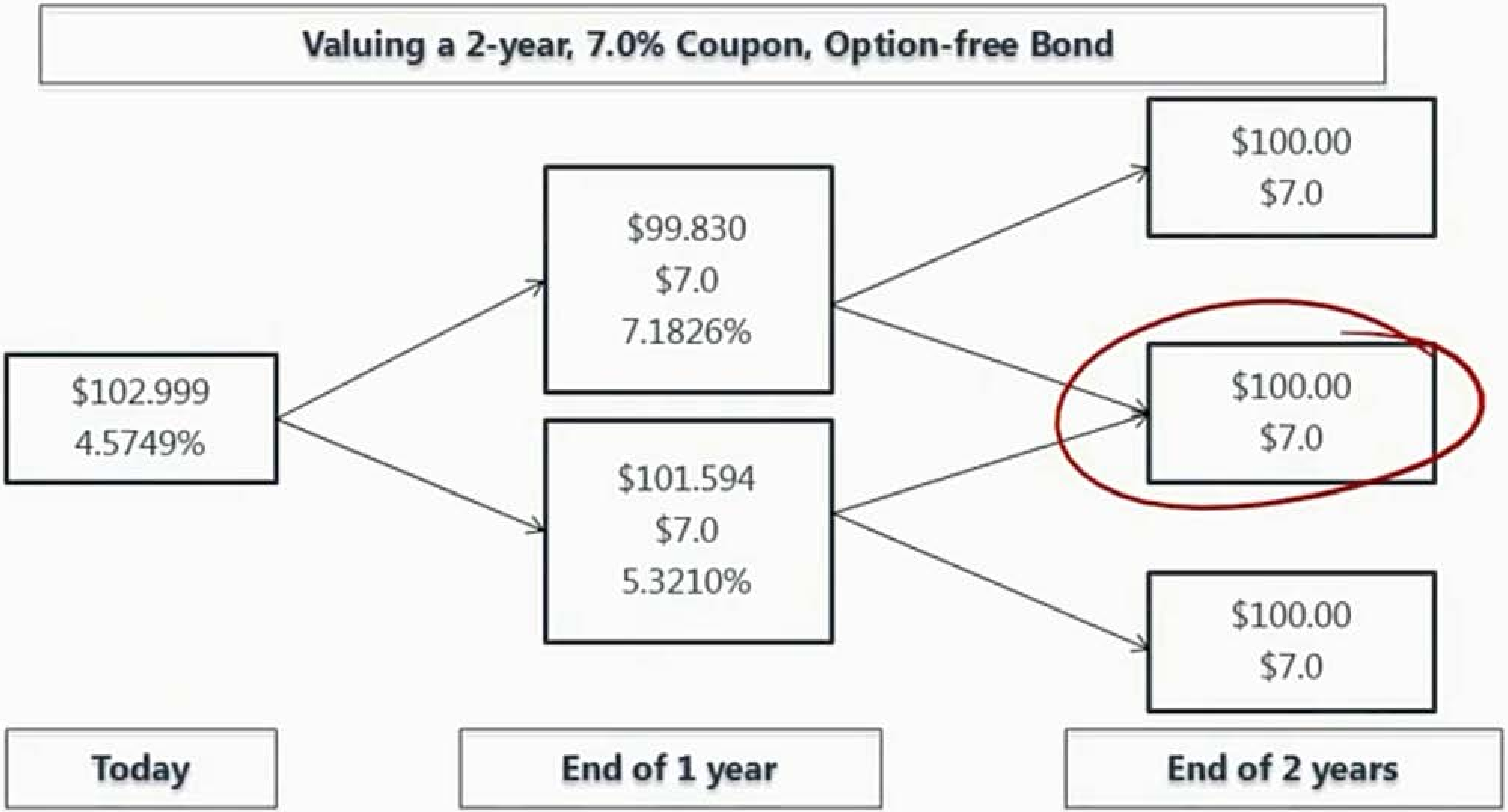


Binominal Interest Rate Tree

➢ **How to constructing the Binomial Interest Rate Tree?**

- The construction of an interest rate tree, binomial or otherwise, is a tedious process. In practice, the interest rate tree is usually generated using specialized computer software.
- There is one underlying rule governing the construction of an interest rate tree:
    - ✓ The values for on-the-run issues generated using an interest rate tree should prohibit arbitrage opportunities. This means that the value of an on-the-run issue produced by the interest rate tree must equal its market price.

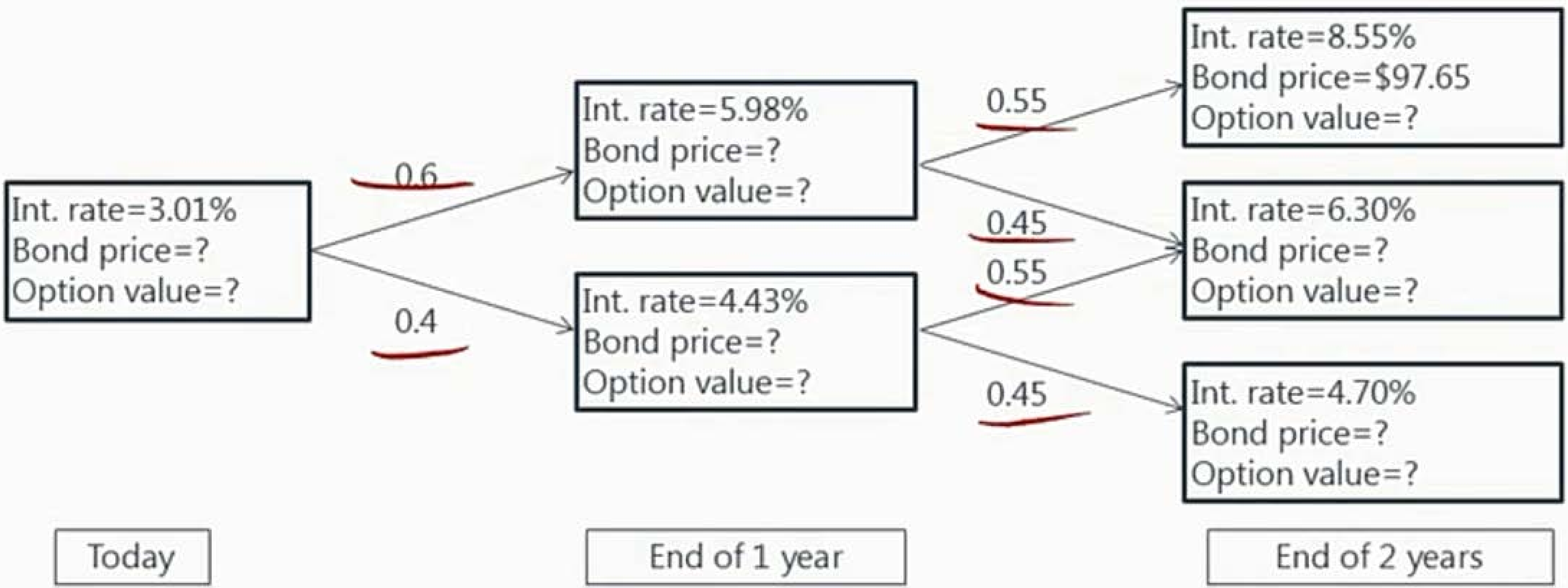
Valuing an Option Bond with the Tree



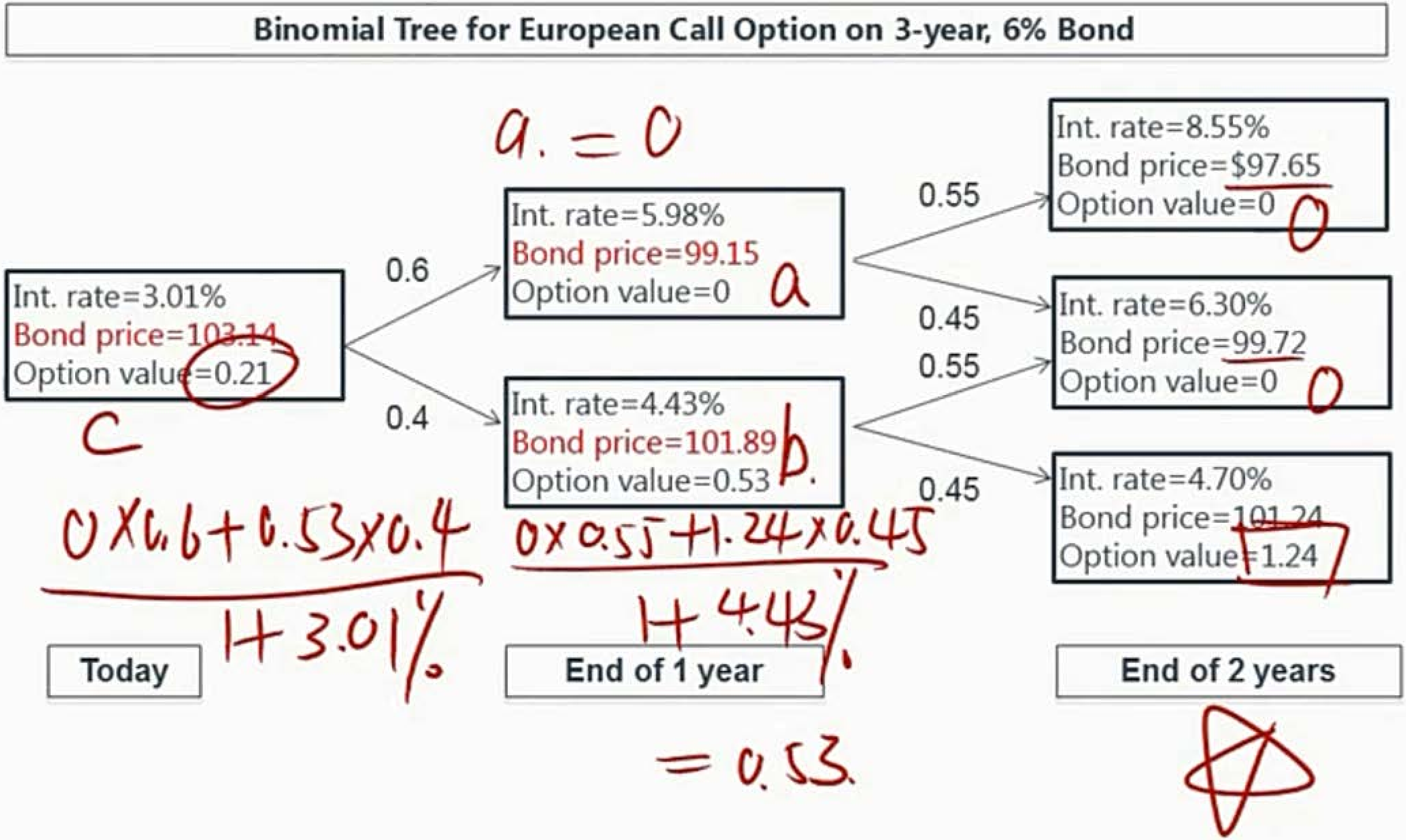
Valuing an Option Bond with the Tree

- Value a European call option with two years to expiration and a strike price of $100.00. The underlying is a 6%, annual-coupon bond with three years to maturity. Assume that the risk-neutral probability of an up move is 0.6 in year 1 and 0.55 in year 2.

Binomial Tree for European Call Option on 3-year, 6% Bond



Valuing an Option Bond with the Tree



The choice of time steps

- Reason to choose time steps smaller than six months
  - ensure that all cash flows are sufficiently close in time to some date in the tree.
  - fill the tree with enough rates to price contingent claims with sufficient accuracy.
- While smaller time steps generate more realistic interest rate distributions, it is not always desirable. $dt \to 0$
  - the greater the number of computations in pricing a security, the more attention must be paid to numerical issues like **round-off error**.
  - practitioners requiring quick results cannot make the time step too small.
- The best choice of step size ultimately depends on the problem at hand.

Fixed Income Securities & Black-Scholes-Merton

- Why the Black-Scholes-Merton model to value equity derivatives is not appropriate to value derivatives on fixed-income securities?
  - The price of a bond must converge to its face value at maturity while the random process describing the stock price need not be constrained in any similar way. $S_T = Par.$
  - Because of the maturity constraint, the volatility of a bond's price must eventually get smaller as the bond approaches maturity. The simpler assumption that the volatility of a stock is constant is not so appropriate for bonds. r
  - Since stock volatility is very large relative to short-term rate volatility, it may be relatively harmless to assume that the short-term rate is constant. By contrast, it can be difficult to defend the assumption that a bond price follows some random process while the short-term interest rate is constant.

Constant Maturity Treasury Swap

- **Calculate the value of a constant maturity Treasury swap, given an interest rate tree and the risk-neutral probability.**
  - The following example calculates the price of a constant maturity Treasury swap.
  - A CMT swap is an agreement to swap a floating rate for a Treasury rate such as the 10-year rate.

Constant Maturity Treasury Swap

收 float 付 fixed

- **Example: CMT swap**
  - The rate tree can be used to price a constant-maturity Treasury (CMT) swap. In the example, the strike (fixed rate) is 5.0% such that the swap pays:

$$\text{payoff} = \left(\frac{\$1{,}000{,}000}{2}\right) \times \left(y_{CMT} - \overset{K}{5\%}\right) = ?$$

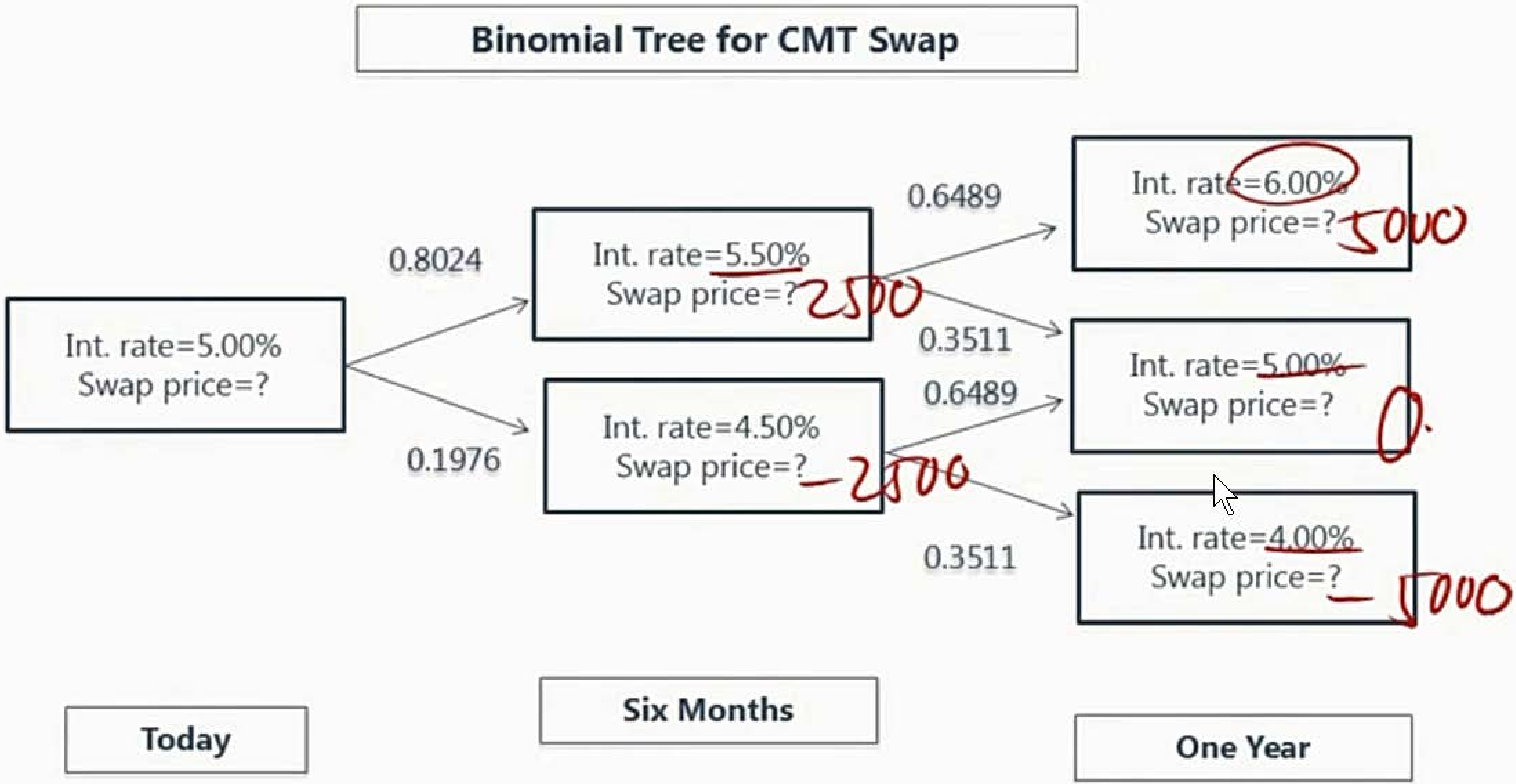
dt=0.5

NP=1m

every six months until maturity

  - y_{CTM} a **semiannually compounded yield**, of a predetermined maturity, on the payment date. The text prices a one-year CMT swap on the six-month yield. In practice, CMT swaps trade most commonly on the yields of the most liquid maturities, i.e., on 2-, 5- and 10-year yields.

Constant Maturity Treasury Swap

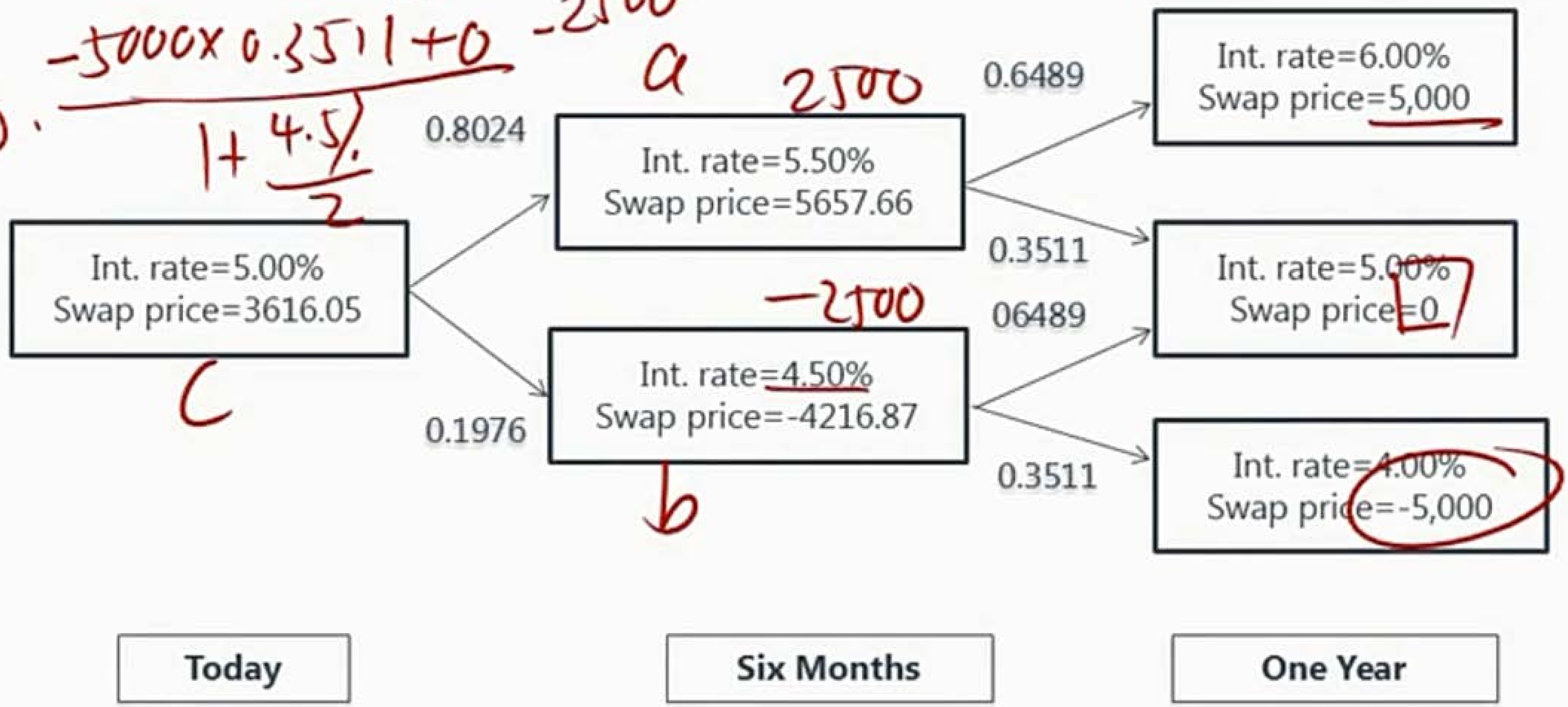


Constant Maturity Treasury Swap

$$V_a: \frac{5000 \times 0.6489 + 0}{1 + \frac{5.5\%}{2}} + 2500$$

$$b: \frac{-5000 \times 0.3511 + 0}{1 + \frac{4.5\%}{2}} - 2500$$

Binomial Tree for CMT Swap

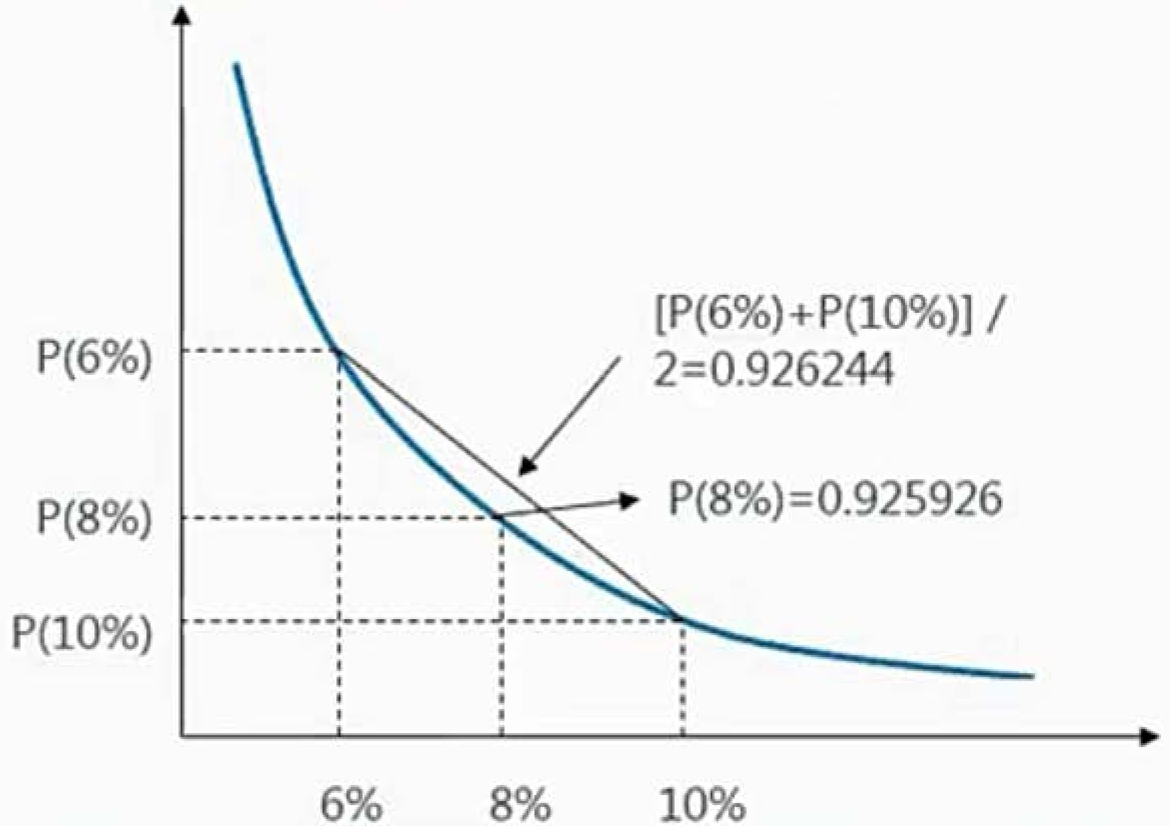


Interest Rate Expectations

- **Under pure expectations theory**, the term structure is a function of expected future spot rates and the shape of the spot rate curve is determined solely by expected future spot rates.
- Hence, <u>forecasts can be very useful in describing the shape and level of the term structure over short-term horizons **but** probably only the level of rates at very long horizons.</u>

Convexity effect

- To isolate the implications of volatility on the shape of the term structure, it assumes that investors are risk-neutral and investors believe that their forecasts will be realized.



- The difference between the two YTMs is called **convexity effect**.

Jensen's Inequality

➢ **Convexity Effect**

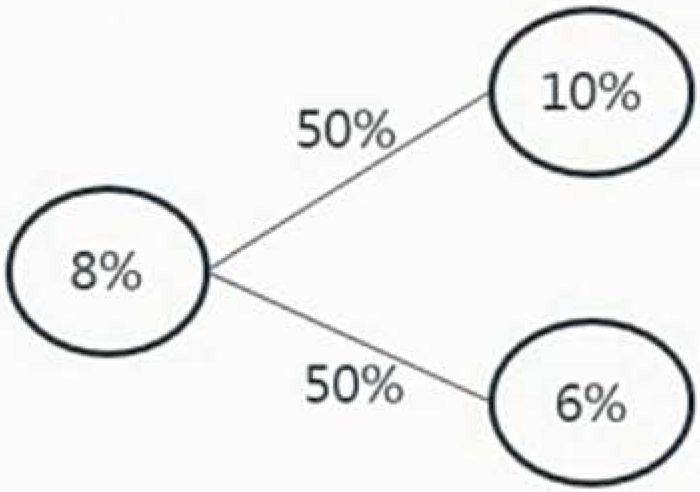
- The convexity effect arises from a special case of Jensen's Inequality:

$$E\left[\frac{1}{(1+\mathrm{r})}\right] > \frac{1}{E(1+r)}$$

- All else held equal, the value of convexity increases with **maturity and volatility**.

Risk Premium

- **Risk premium** is the price for risk averse investors to bear uncertainty.
- **Calculate** the price for the 2-year zero-coupon bond with a 20 basis point risk premium.



$$P = \frac{\left[\frac{\$1}{1.102} + \frac{\$1}{1.062}\right] \times 0.5}{1.08}$$

$$= \frac{[\$0.90744 + \$0.94162] \times 0.5}{1.08}$$

$$= \$0.85605$$

Bond: $T = 2$. Par $= \$1$

(Zero)

$V_0 =$

	A	B	C	D
50		10%	$1	
51		0.909091		
52	8%			
53	=(0.5*B51+0.5*B54)/(1+A52)			
54		0.943396		
55				
56				
57				

	A	B	C
50		10%	$1
51		0.909091	
52	8%		
53	0.8576	6%	$1
54		0.943396	
55			
56			

	A	B	C	D	E
50		10%	$1		
51		0.909091			
52	8%				
53	0.8576	6%	$1	Method 1	
54		0.943396			
55					
56					
57		=1/(1+8%)		Method 2	
58					

A57 =B57/(1+8%)

	A	B	C	D	E
50		10%	$1		
51		0.909091			
52	8%				
53	0.8576	6%	$1	Method 1	
54		0.943396			
55					
56					
57	=B57/(1+8%)			Method 2	
58					
59					
60					

	A	B	C	D	E
50		10%	$1		
51		0. 909091			
52	8%				
53	0. 8576	6%	$1	Method 1	
54		0. 943396			
55					
56					
57	0. 8573	0. 925926		Method 2	
58					

	A	B	C	D
49				
50		10%	$1	
51		0.909091		
52	8%			
53	0.8576	6%	$1	LHS
54		0.943396		
55				
56				
57	0.8573	0.925926	$1	RHS
58				

Jensen's Inequality

➢ **Convexity Effect**

- The convexity effect arises from a special case of Jensen's Inequality:

$$E\left[\frac{1}{(1+\mathrm{r})}\right] > \frac{1}{E(1+r)}$$

- All else held equal, the value of convexity increases with **maturity and volatility**.

➢ Convexity Effect

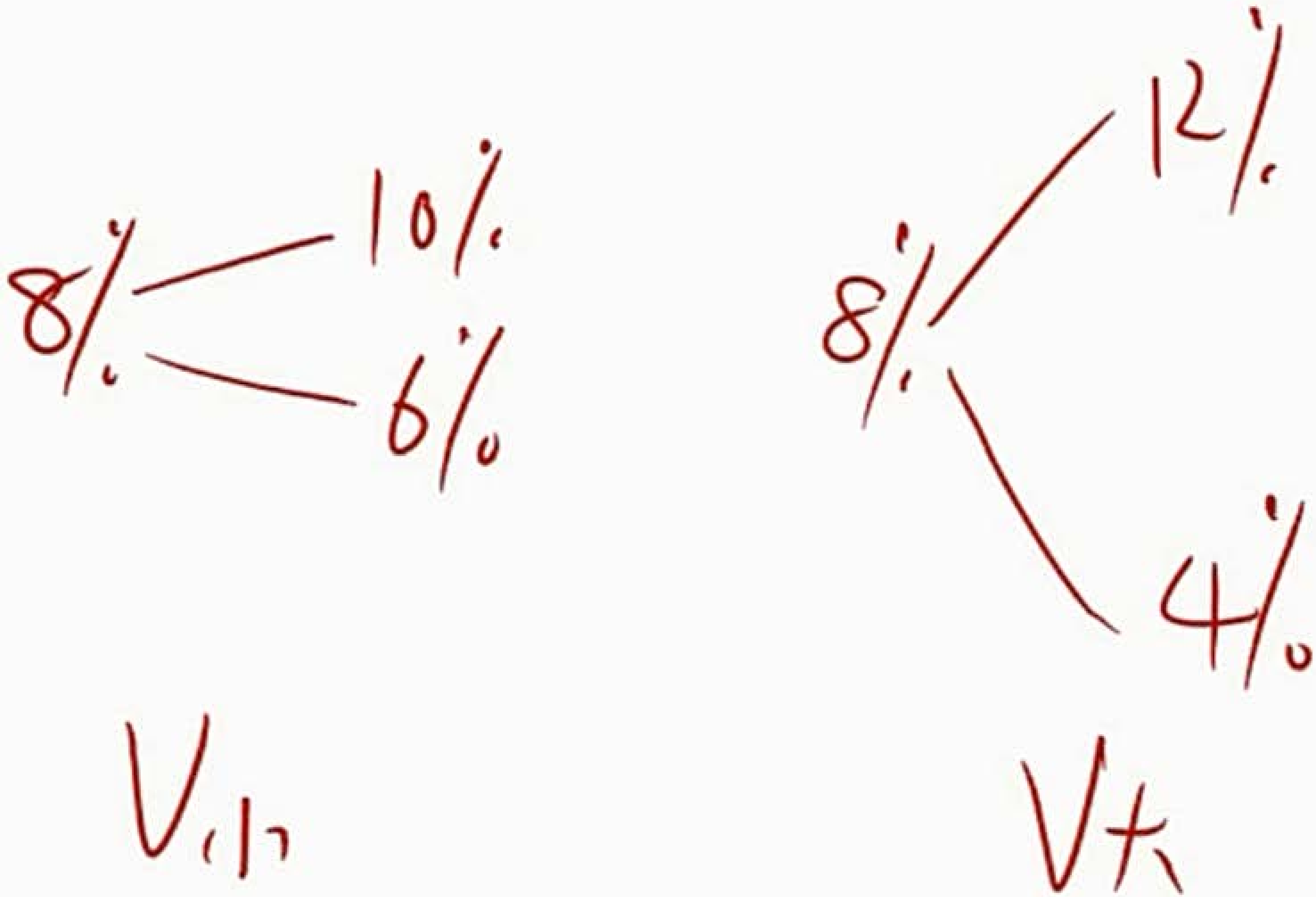
- The convexity effect arises from a special case of Jensen's Inequality:

$$E\left[\frac{1}{(1+r)}\right] > \frac{\$1}{E(1+r)}$$

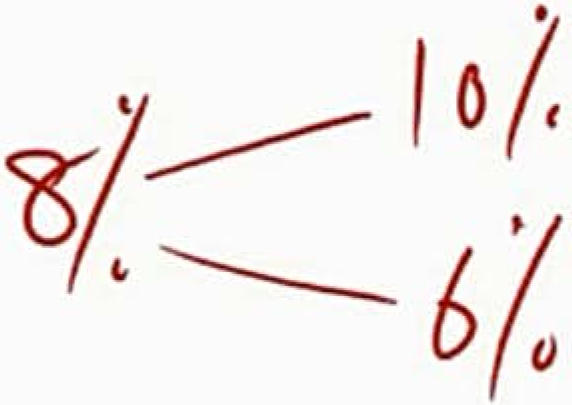
ave

- All else held equal, the value of convexity increases with **maturity and volatility**.

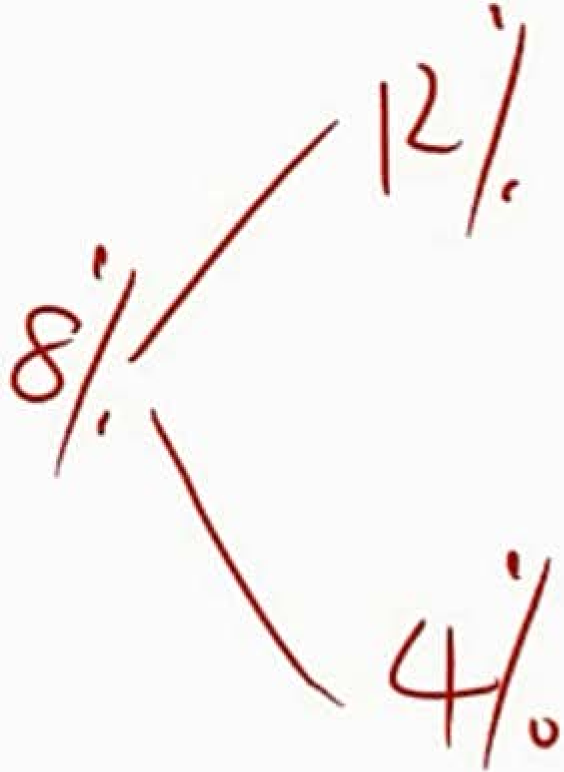
$$0.5 \times \frac{1}{1+10\%} + 0.5\,\frac{1}{1+6\%} > \frac{\$1}{1+\frac{10\% + 6\%}{2}}$$



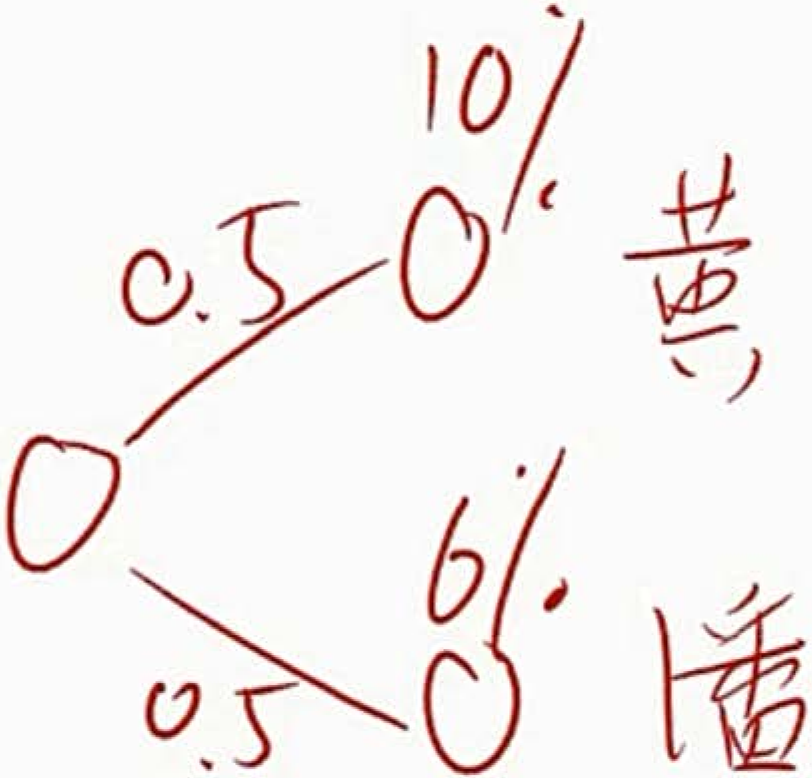
8%
10%
6%
8%
12%
4%



V(1)



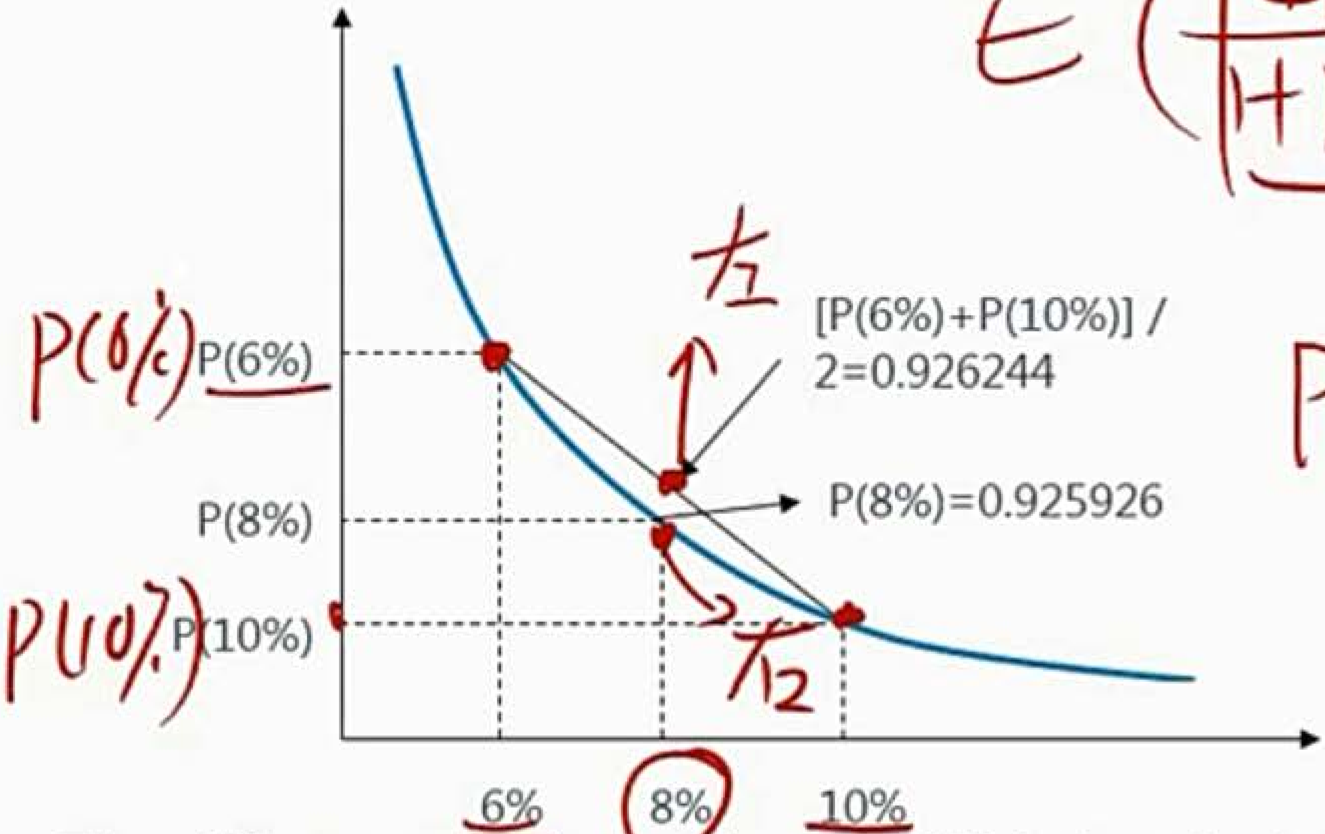
Vt



$V_0 \to$ 大(左) $> V_1$,(右)

Convexity effect

- To isolate the implications of volatility on the shape of the term structure, it assumes that investors are risk-neutral and investors believe that their forecasts will be realized.



$$E\left(\boxed{\frac{\$1}{1+10\%}}, \boxed{\frac{\$1}{1+6\%}}\right)$$

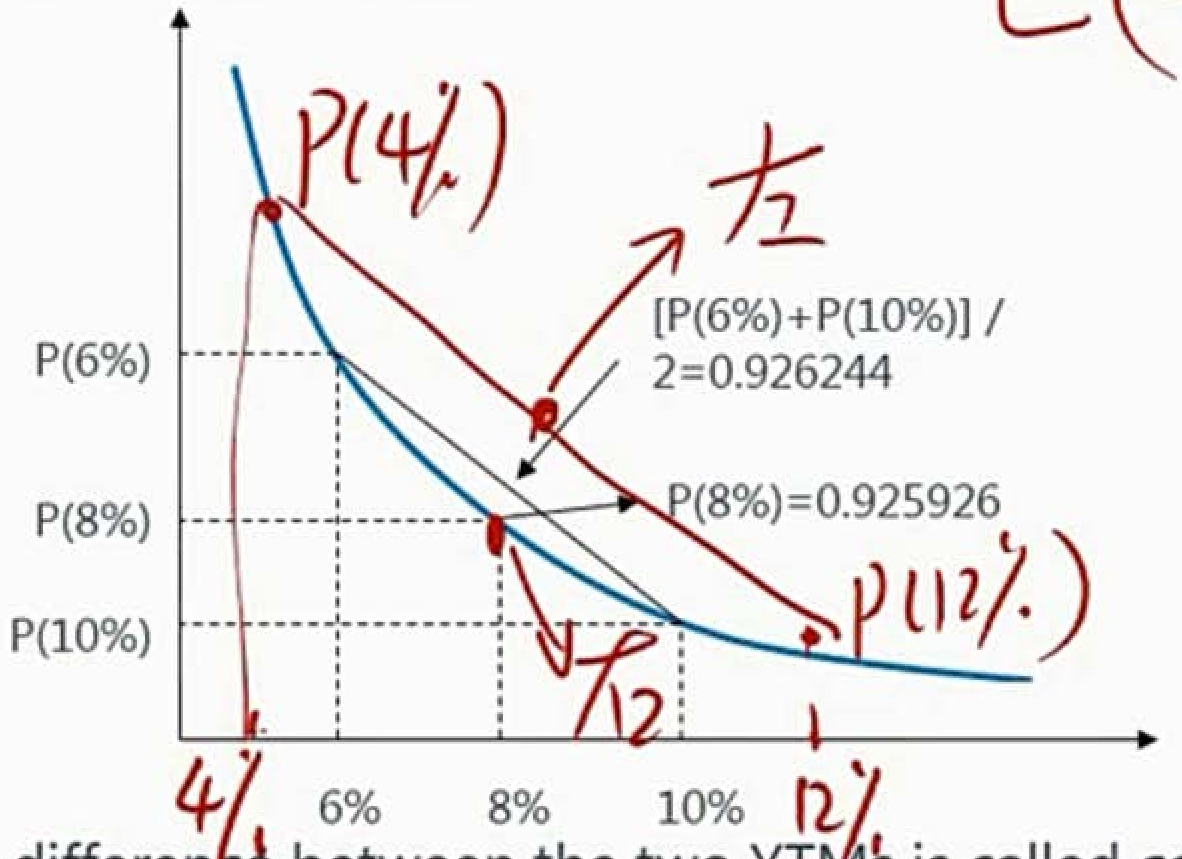
$P(10\%)$ $\quad$ $P(6\%)$

$$\boxed{\frac{\$1}{1+8\%}} \quad P(8\%)$$

- The difference between the two YTMs is called **convexity effect**.

- To isolate the implications of volatility on the shape of the term structure, it assumes that investors are risk-neutral and investors believe that their forecasts will be realized.

E(4%, 12%)



- The difference between the two YTMs is called **convexity effect**.

The Art of Term Structure Models: Drift & Volatility and Distribution

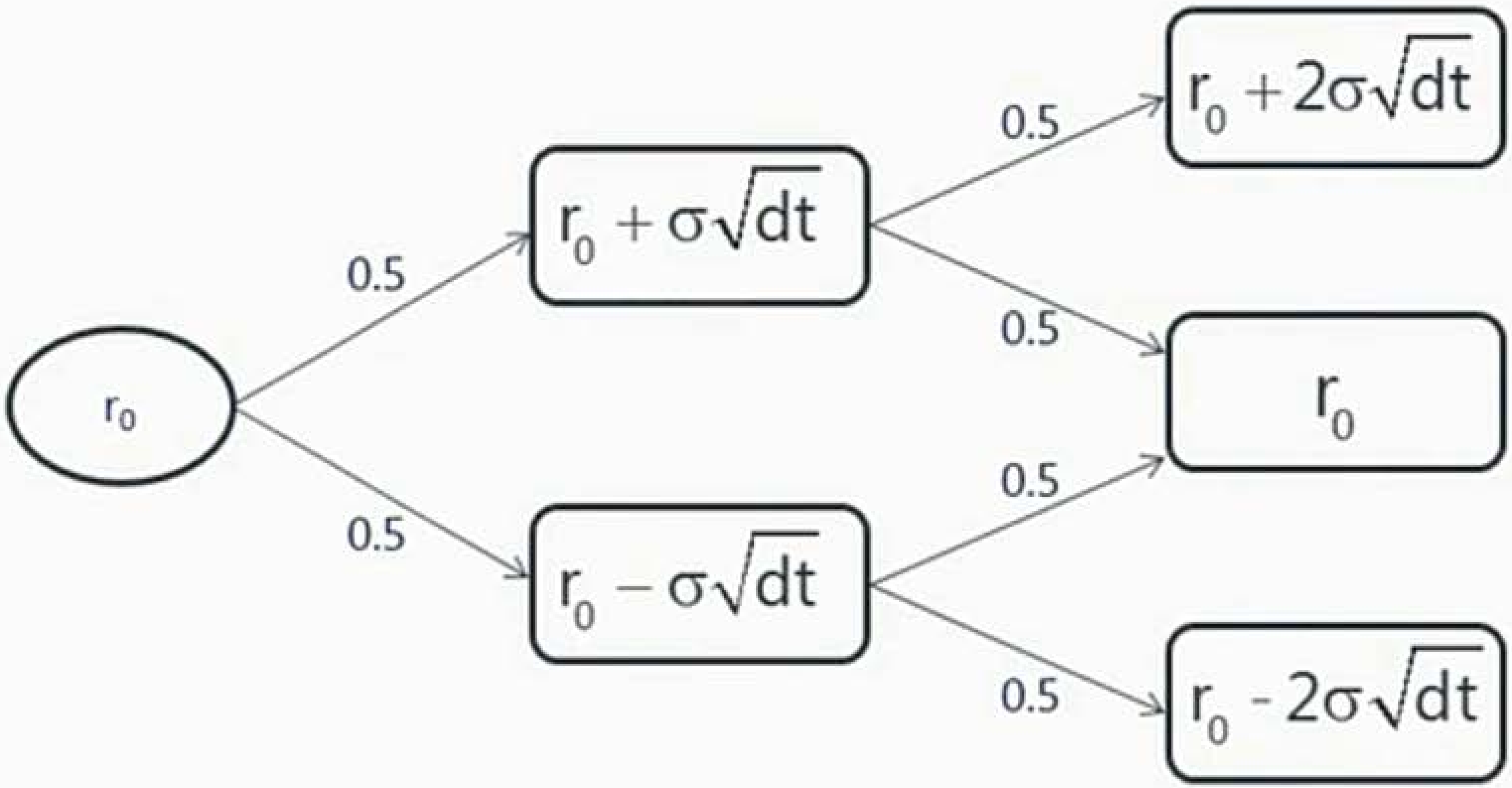
Term Structure Models of Interest Rates

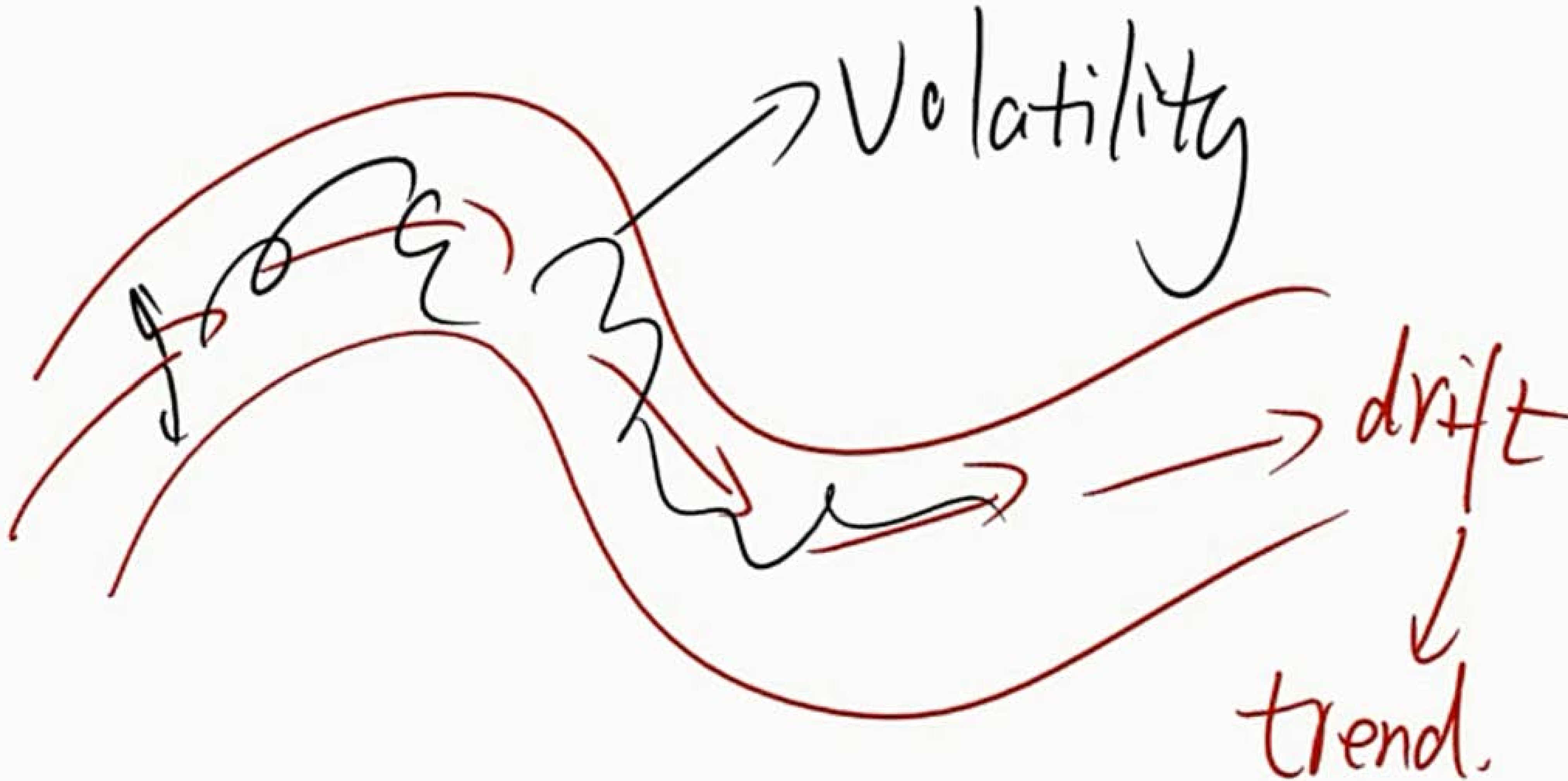
175-203

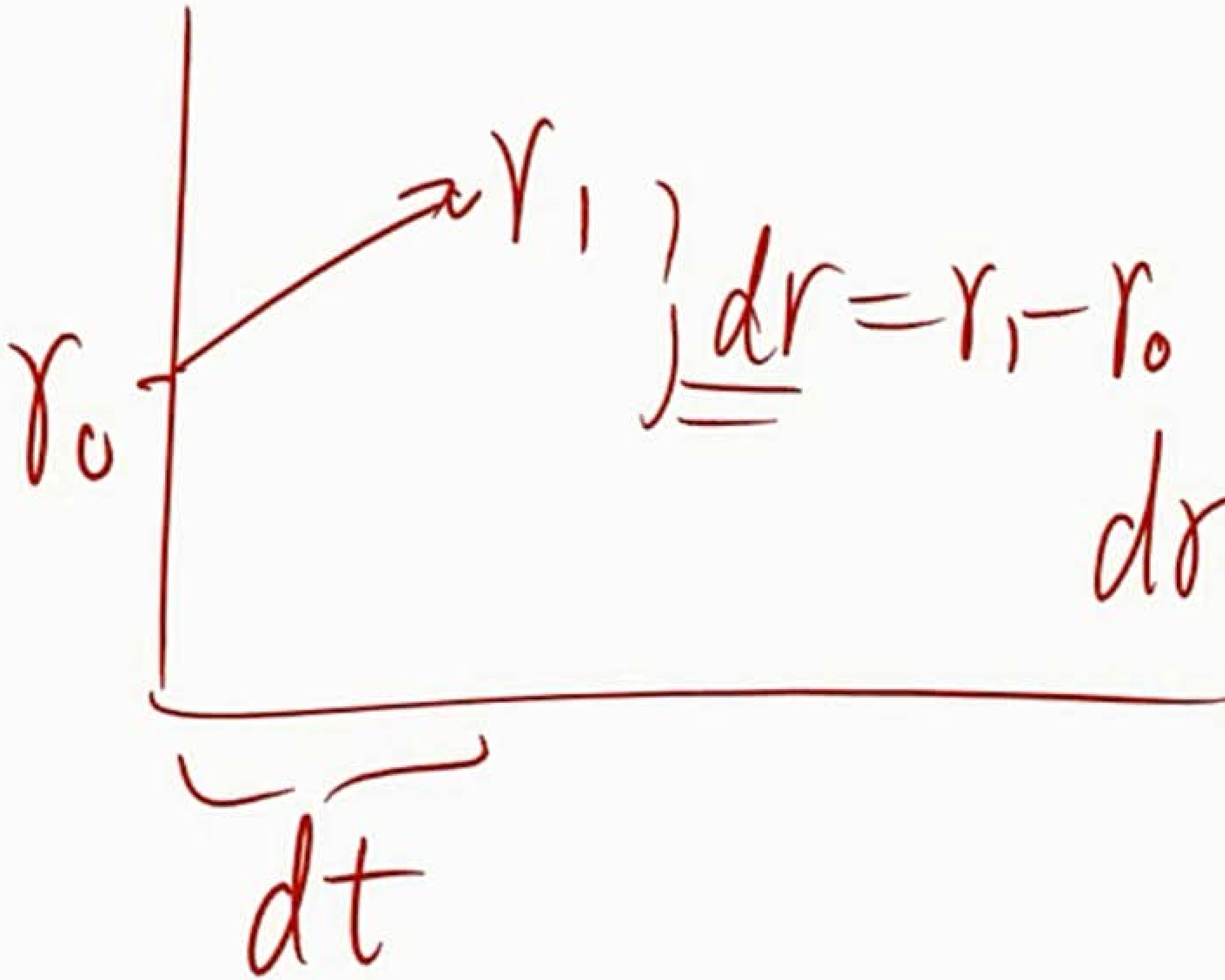
Model 1

- Model 1 is assumed to have an interest rate term structure with no drift.
- As the expected value of (dw) is zero, the expected change in the rate (a.k.a., the drift) is zero.

$$dr = \sigma dw \qquad dw = \varepsilon\sqrt{dt}$$

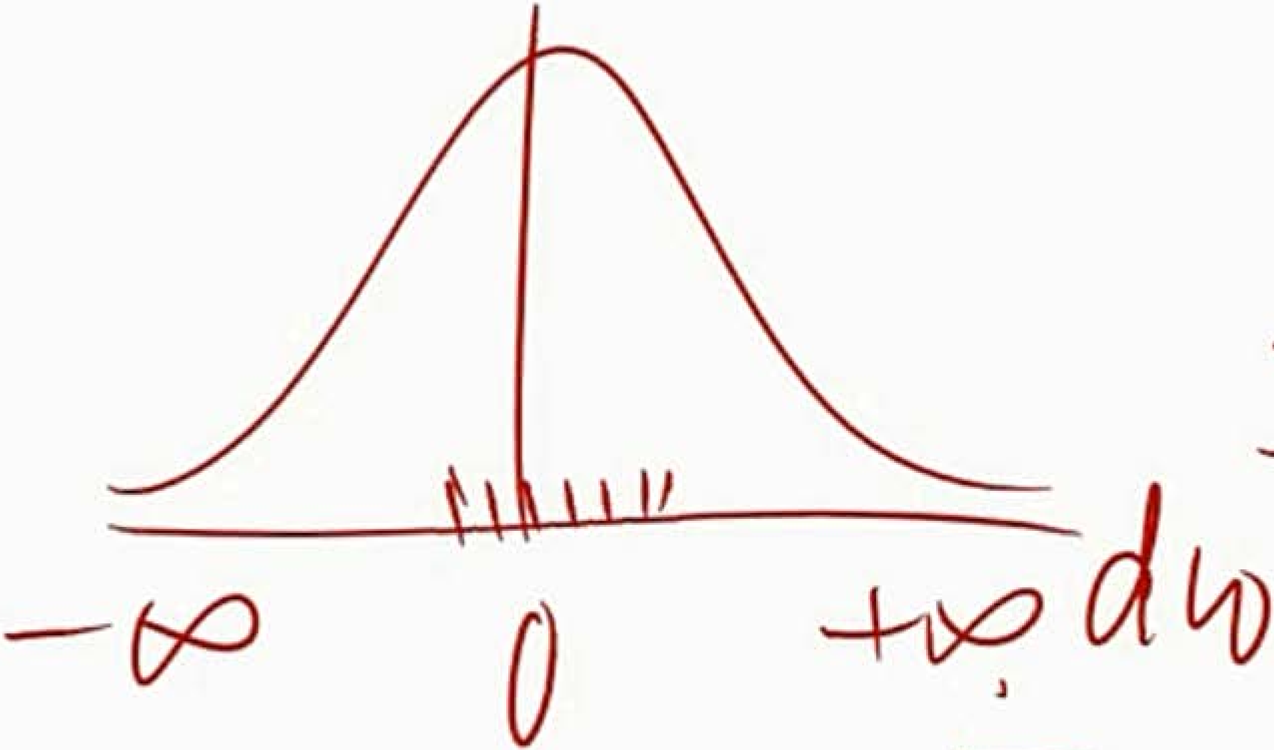


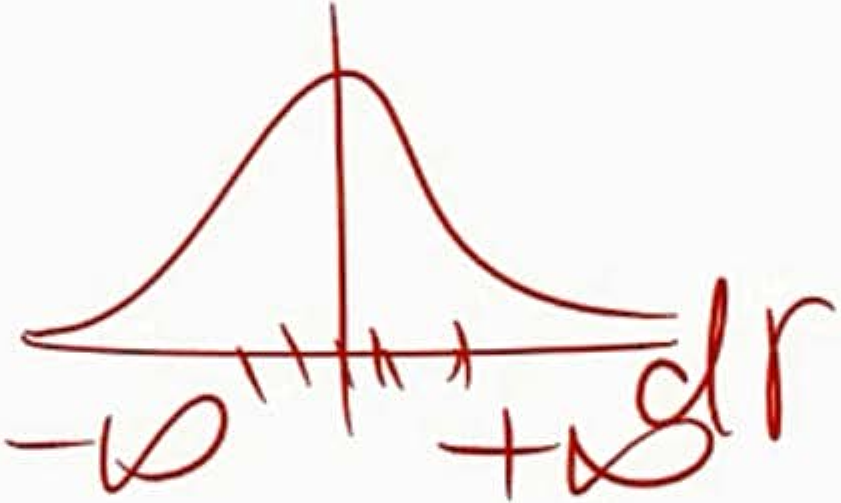

Volatility
drift
trend.



$$dr = r_1 - r_0$$

$$dr = \underbrace{\sigma}_{30\%} \times \underbrace{\boxed{dw}}_{\text{随机项}}^{\overset{N(0,dt)}{|}}$$





$$dr = \boxed{\sigma}\, dw$$

Model 1

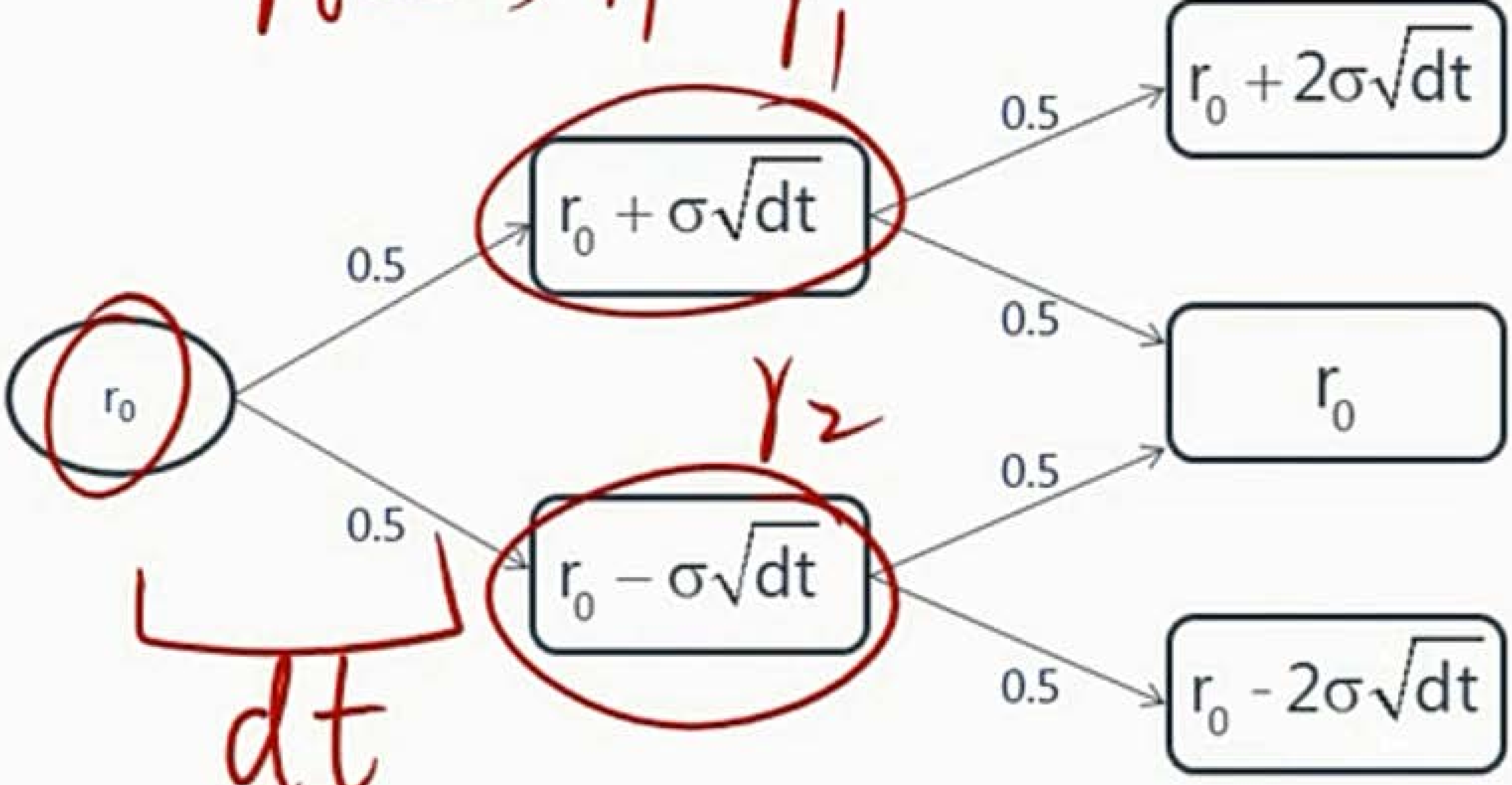
- Model 1 is assumed to have an interest rate term structure with no drift.
- As the expected value of (dw) is zero, the expected change in the rate (a.k.a., the drift) is zero. $dw \sim N(0, dt)$

$r_0 + \sigma\sqrt{dt}$

$$\boxed{dr = \sigma dw} \qquad dw = \varepsilon\sqrt{dt}$$

$r_0 \xrightarrow{r_0+dr} r_1$

$r_0 - \sigma\sqrt{dt}$



$$
\begin{tikzcd}[column sep=large]
 & & r_0 + 2\sigma\sqrt{dt} \\
 & r_0 + \sigma\sqrt{dt} \arrow[ur, no head] \arrow[dr, no head] & \\
r_0 \arrow[ur, no head] \arrow[dr, no head] & & \boxed{r_0} \\
 & r_0 - \sigma\sqrt{dt} \arrow[ur, no head] \arrow[dr, no head] & \\
 & & r_0 - 2\sigma\sqrt{dt}
\end{tikzcd}
$$

$$dr = \sigma \underline{dw}$$

$$N(0, dt)$$

$$dw = \boxed{\sqrt{dt} \cdot \varepsilon}$$

$$\varepsilon \rightarrow N(0,1)$$

$$E(dw) = \sqrt{dt}\, E(\varepsilon) = 0$$

$$V(dw) = (\sqrt{dt})^2 V(\varepsilon) = dt \cdot \times 1 = dt$$

$$dr = \sigma \, dw$$

$$\delta r = \sigma \cdot \sqrt{dt} \cdot \varepsilon$$

$N(0,1)$

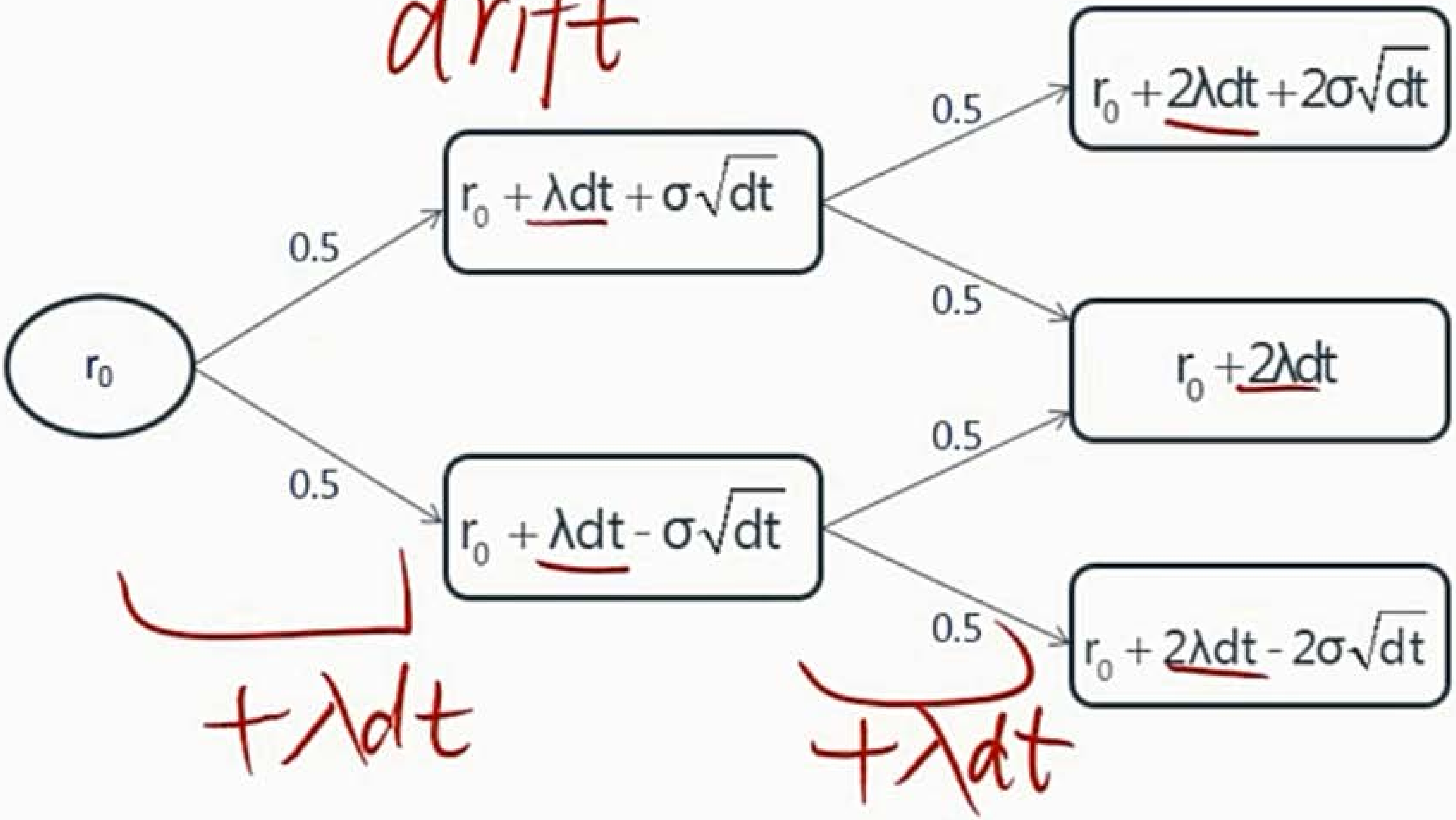
Model 1

- A problem with Gaussian models is that the short-term rate can become **negative. Solutions are as follow:**
  - **A non-normal distribution can be assumed**: For example, if we assume interest rates are lognormally distributed. However, building a model around a probability distribution that rules out negative rates or makes them less likely may result in volatilities that are unacceptable.
  - **Use shadow rates** (force the "adjusted" tree rates to be non-negative): Another popular method of ruling out negative rates is to construct rate trees with whatever distribution is desired, as done in this section, and then simply set all negative rates to zero. In this methodology, rates in the original tree are called the shadow rates of interest while the rates in the adjusted tree could be called the observed rates of interest.

Model 2

- **Constant drift model: Model 2**

$$dr = \underbrace{\lambda dt}_{drift} + \boxed{\sigma dw}$$



Ho-lee Model

$\lambda(t)$

- **Ho-lee Model:** Short-term rate process with time-dependent drift.
- It is clear that if $\lambda_1 = \lambda_2$ then Ho-Lee model reduces to Model 2

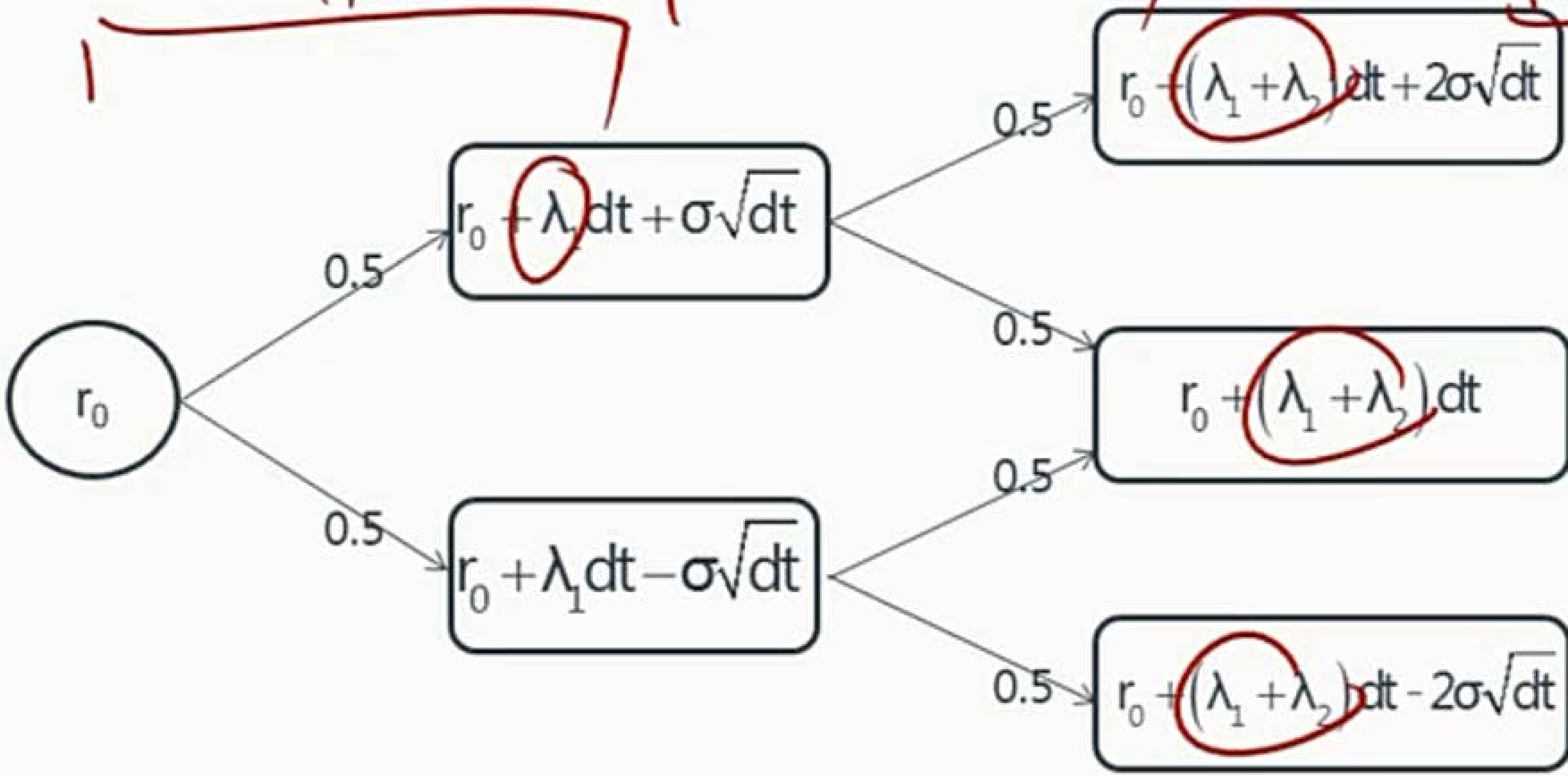
r

$\lambda(t)$

t

λ_1

λ_2



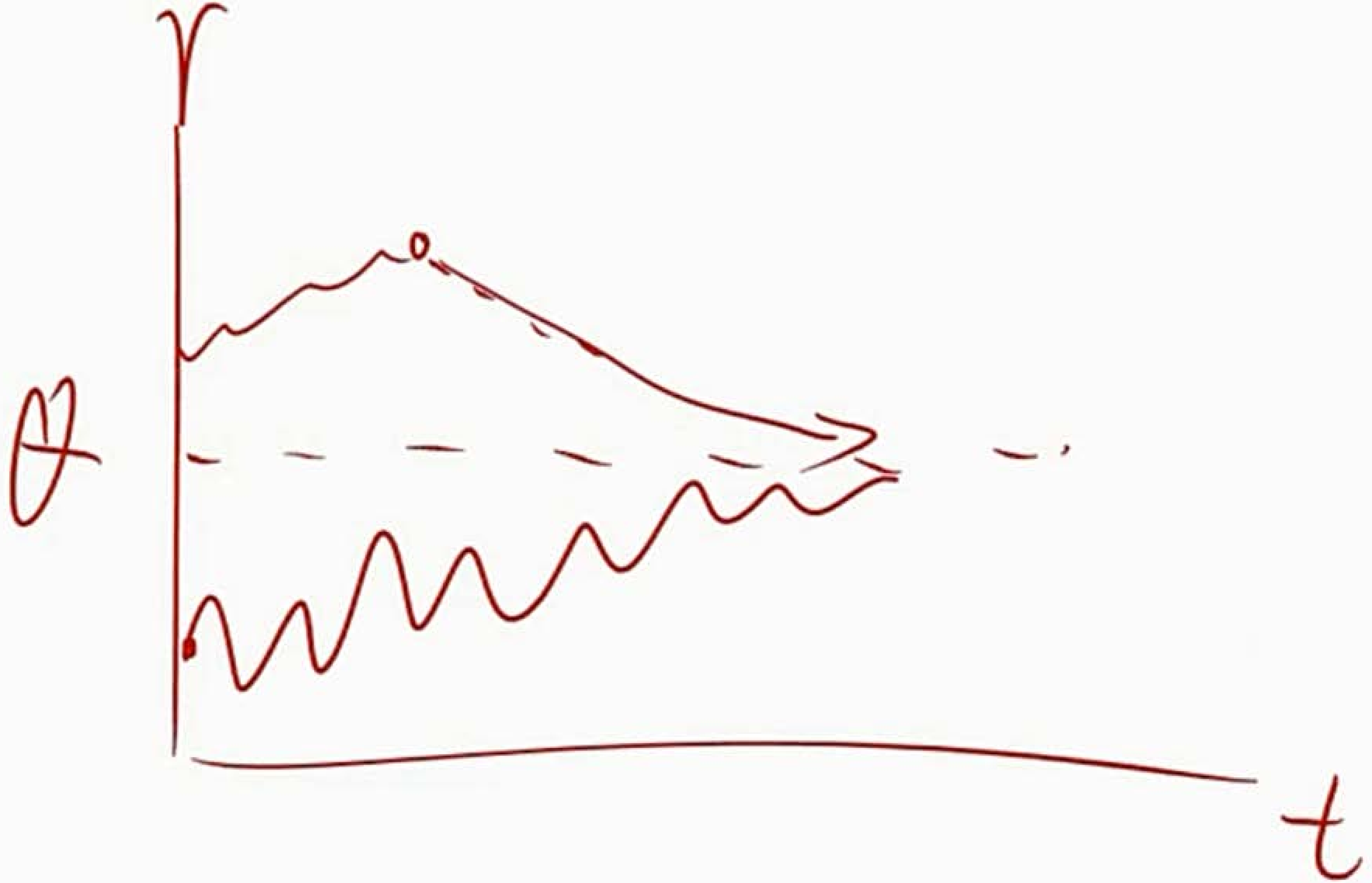
Vasicek Model

- The **Vasicek Model** introduces mean reversion into the rate model, which is a common assumption for the level of interest rates. The Vasicek Model is given by:

$$dr = \kappa(\theta - r)dt + \sigma dw$$

  - k = a parameter that measures the speed of reversion adjustment
  - θ = long-run value of the short-term rate assuming risk neutrality
  - r = current interest rate level

- The greater the difference between r and θ, the greater the expected change in the short-term rate toward θ.



r
θ
t

$\theta = 5\% \quad r_0 = 3\% \quad k = 0.5$

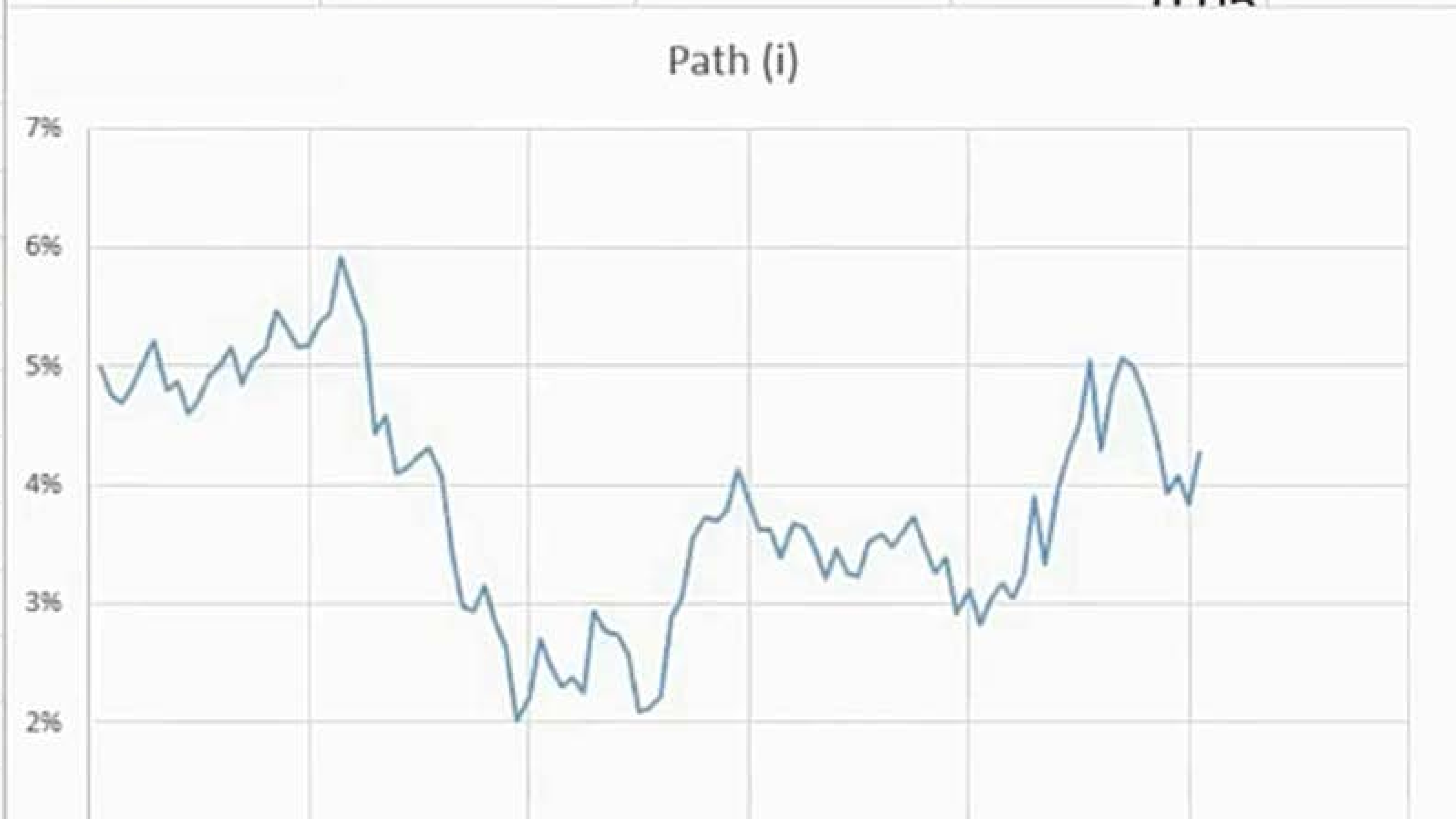
$$dr = 0.5(5\% - 3\%) = 1\%$$

$$r_1 = r_0 + dr = 3\% + 1\% = 4\%$$

E2 =E2+kappa*(theta-E2)*dt+sigma*sqrt(dt)*norm.s.inv(rand())

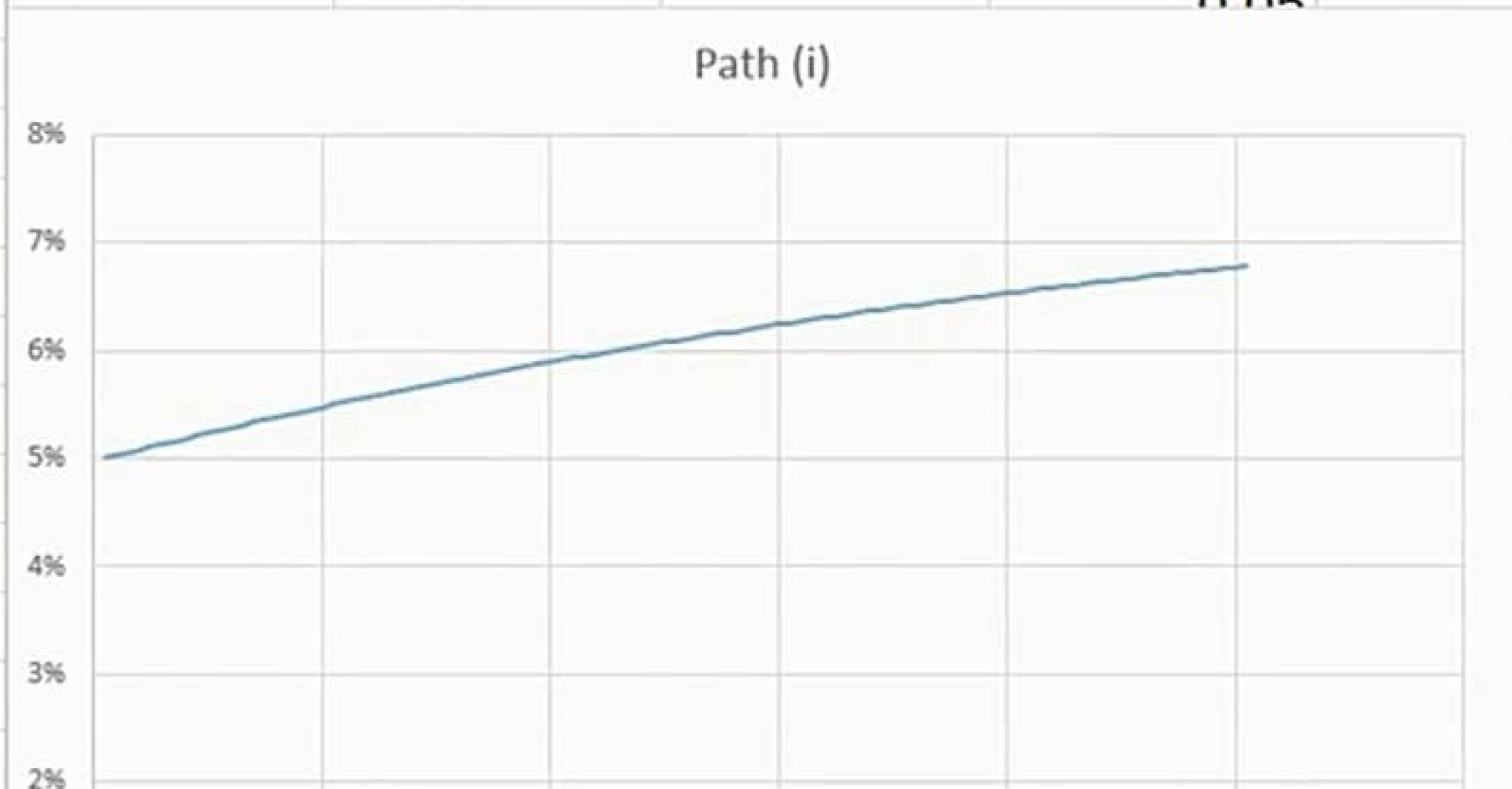
	A	B	C	D	E	F
1	r(0)	5%		t	Path (i)	
2	theta	8%		0	5%	
3	k	0.7		0.01	=E2+kappa*(theta-E2)*dt+sigma*sqrt(dt)*norm.s.inv(rand())	
4	T	1		0.02		
5	dt	0.01		0.03		
6	sigma	3%		0.04		
7				0.05		
8				0.06		
9				0.07		
10				0.08		
11				0.09		
12				0.1		
13				0.11		
14				0.12		
15				0.13		
16				0.14		

	A	B	C	D	E	F
1	r(0)	5%		t	Path (i)	
2	theta	8%		0	5%	5%
3	k	0.7		0.01	4.759%	5.086%
4	T	1		0.02	4.696%	5.308%
5	dt	0.01		0.03	4.832%	5.263%
6	sigma	3%		0.04	5.031%	5.204%
7				0.05	5.220%	5.010%
8					4.803%	5.046%
9					4.865%	5.368%
10					4.603%	5.618%
11					4.731%	5.529%
12					4.929%	5.260%
13					5.010%	4.767%
14					5.162%	5.100%
15					4.851%	4.011%
16					5.059%	4.408%
17					5.138%	5.034%
18					5.461%	5.410%
19					5.315%	5.690%

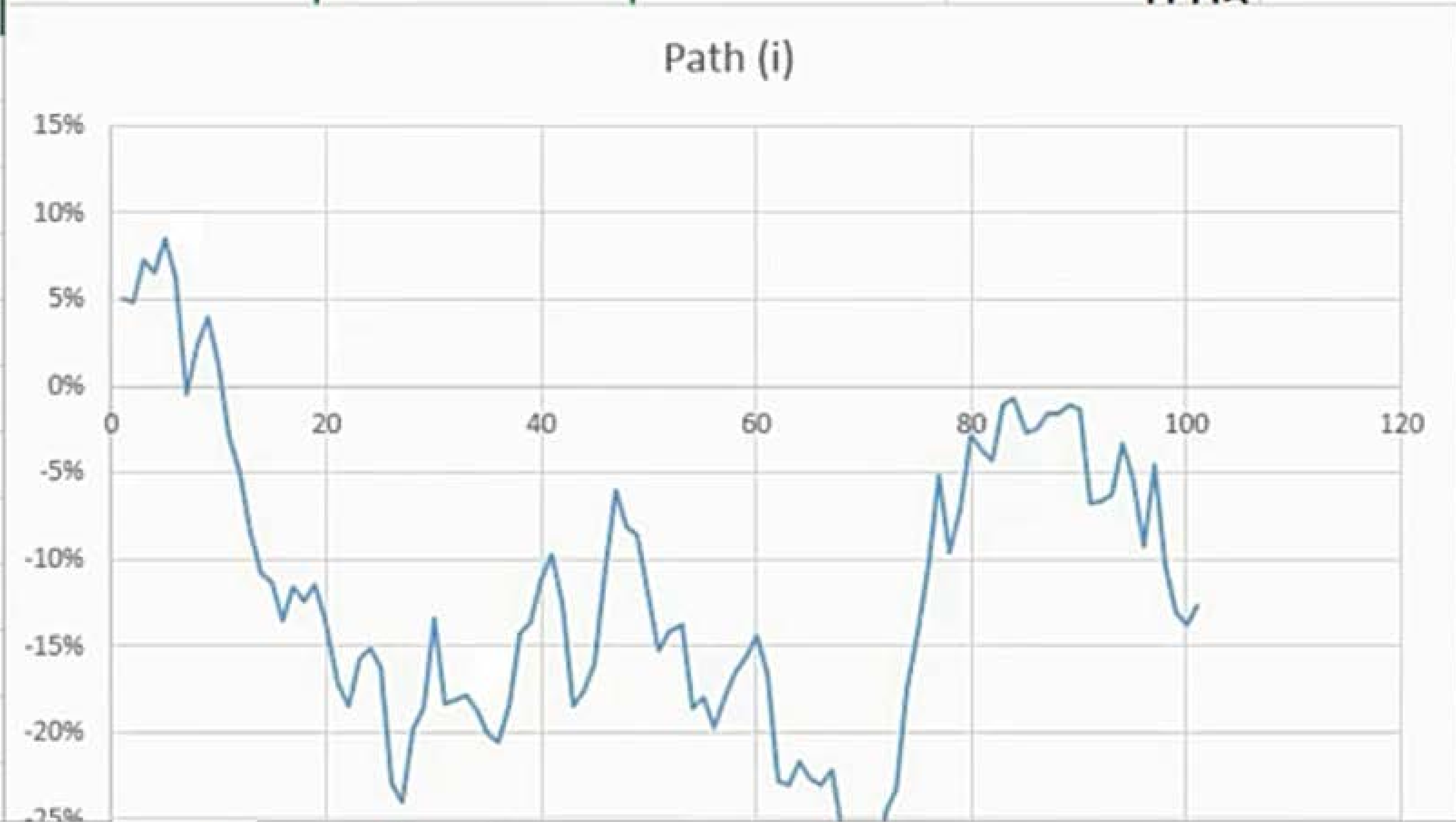


	A	B	C	D	E	F
1	r(0)	5%		t	Path (i)	
2	theta	8%		0	5%	5%
3	k	0.7		0.01	5.021%	5.021%
4	T	1		0.02	5.042%	5.042%
5	dt	0.01		0.03	5.063%	5.063%
6	sigma	0%		0.04	5.083%	5.083%
7				0.05	5.104%	5.104%
8					5.124%	5.124%
9					5.144%	5.144%
10					5.164%	5.164%
11					5.184%	5.184%
12					5.204%	5.204%
13					5.223%	5.223%
14					5.243%	5.243%
15					5.262%	5.262%
16					5.281%	5.281%
17					5.300%	5.300%
18					5.319%	5.319%
19					5.338%	5.338%

	A	B	C	D	E
1	r(0)	5%		t	Path (i)
2	theta	8%		0	5%
3	k	0.9		0.01	5.027%
4	T	1		0.02	5.054%
5	dt	0.01		0.03	5.080%
6	sigma	0%		0.04	5.107%
7				0.05	5.133%
8					5.158%
9					5.184%
10					5.209%
11					5.234%
12					5.259%
13					5.284%
14					5.308%
15					5.333%
16					5.357%
17					5.380%
18					5.404%



	A	B	C	D	E
1	r(0)	5%		t	Path (i)
2	theta	8%		0	5%
3	k	0.7		0.01	4.796%
4	T	1		0.02	7.274%
5	dt	0.01		0.03	6.477%
6	sigma	30%		0.04	8.557%
7				0.05	6.219%
8					-0.489%
9					2.345%
10					3.910%
11					1.252%
12					-2.898%
13					-5.084%
14					-8.555%
15					-10.808%
16					-11.326%
17					-13.526%
18					-11.558%
19					-12.390%



Model 3

- The traditional model with time-dependent volatility is written as

$$dr = \lambda(t)dt + \boxed{\sigma(t)dw}$$

- If the volatility start with a constant, σ, and then exponentially declines to zero, the model as simply called **Model 3**.

$$dr = \lambda(t)dt + \sigma e^{-\alpha t}dw$$

- Cox-Ingersoll-Ross (CIR) Model

$$dr = k(\theta - r)dt + \sigma\sqrt{r}dw$$

Model 3

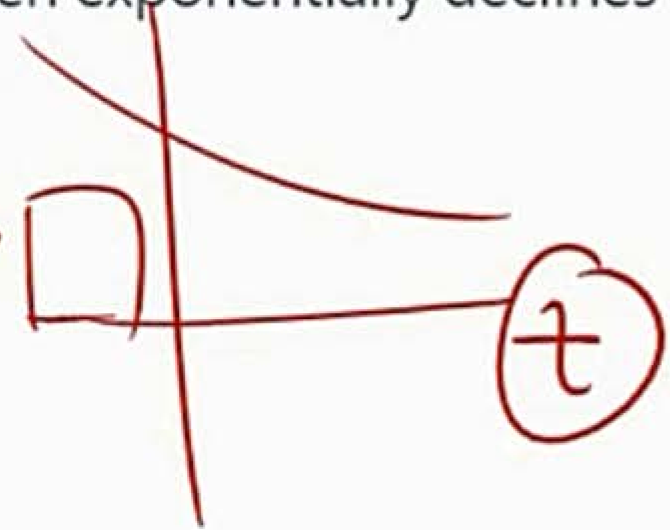
➢ The traditional model with time-dependent volatility is written as

$$dr = \lambda(t)dt + \boxed{\sigma(t)dw}$$

- If the volatility start with a constant, σ, and then exponentially declines to zero, the model as simply called **Model 3**.

$$dr = \lambda(t)dt + \boxed{\sigma \boxed{e^{-\alpha t}} dw}$$

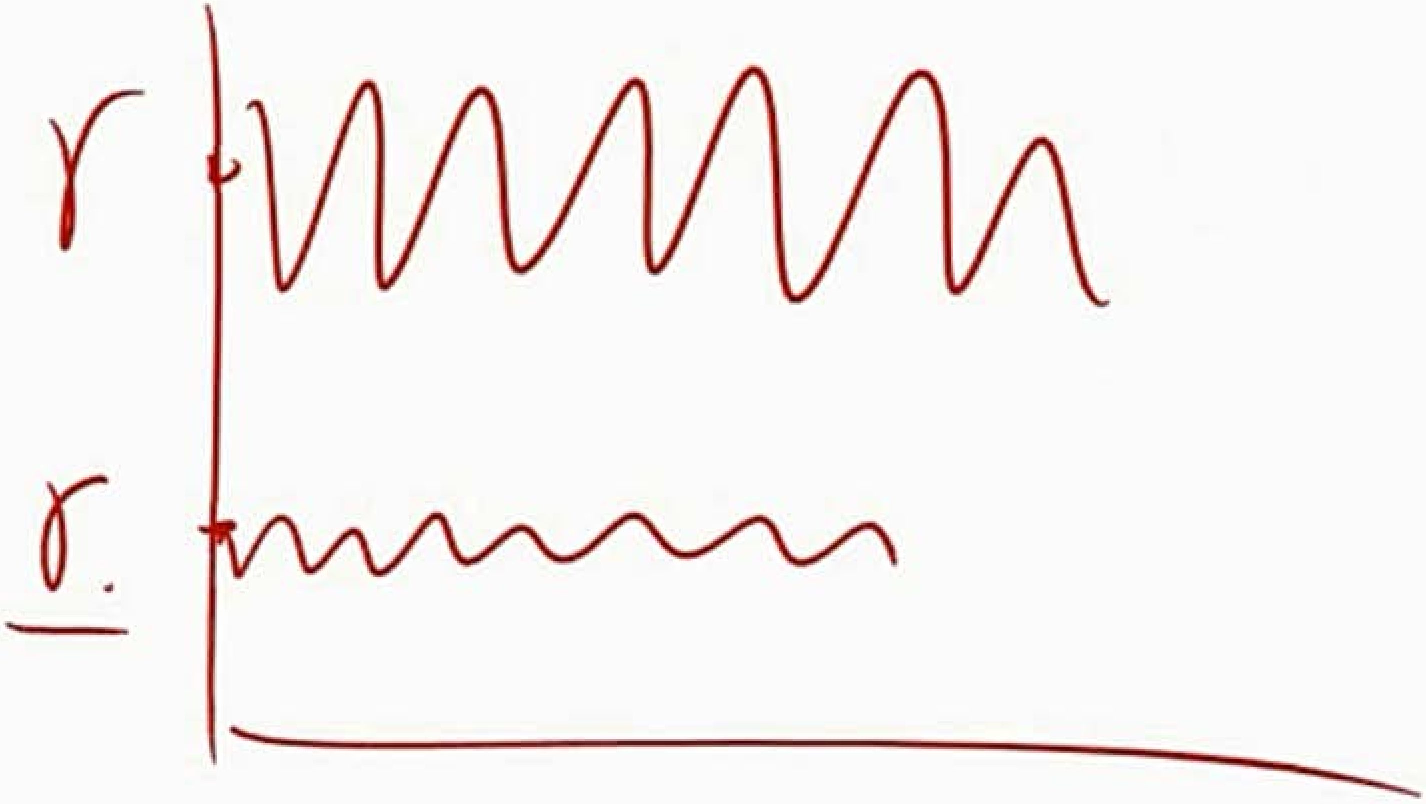
$\sigma \times \square$



- Cox-Ingersoll-Ross (CIR) Model

$$dr = k(\theta - r)dt + \boxed{\sigma\sqrt{r}dw}$$

$\boxed{\sqrt{r}}$ dw



Cox-Ingersoll-Ross (CIR) Model

- If the basis-point volatility of the short term is an increasing function of the short term rate and interest rate appears to have a feature of mean reverting, the model is called **Cox-Ingersoll-Ross (CIR) Model.**

$$\mathrm{d}r = k(\theta - r)dt + \sigma\sqrt{r}dw$$

Lognormal Models

- The simplest Lognormal model **(Model 4)**

$$\mathrm{d}r = ardt + \sigma rdw$$

- **The Salomon Brothers Model** (Lognormal model with deterministic drift)

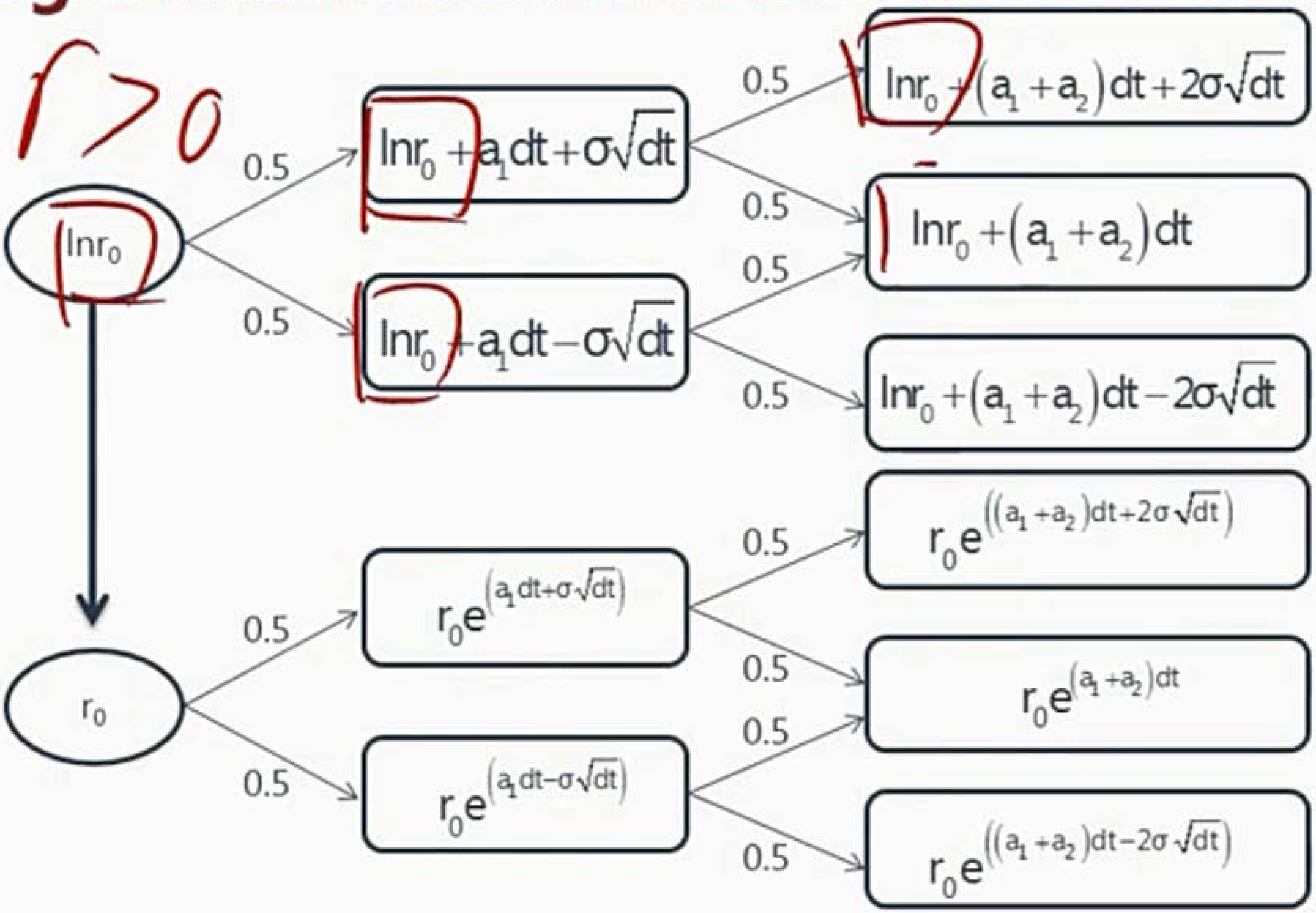
$$\mathrm{d}[\ln(r)] = a(t)dt + \sigma dw$$

- **The Black-Karasinski Model** (Lognormal model with mean reversion)

$$\mathrm{d}[\ln(r)] = k(t)[\ln\theta(t) - \ln(r)]dt + \sigma(t)dw$$

Lognormal Model with Deterministic Drift

$r > 0$



- It can be noted that the future movements of the short rate in a lognormal model **are multiplicative instead of being additive as in normal models.**

$$\begin{tikzcd}
 & x_1 \\
\boxed{x} \arrow[ur, dash] \arrow[dr, dash] & \\
 & x_2
\end{tikzcd}$$

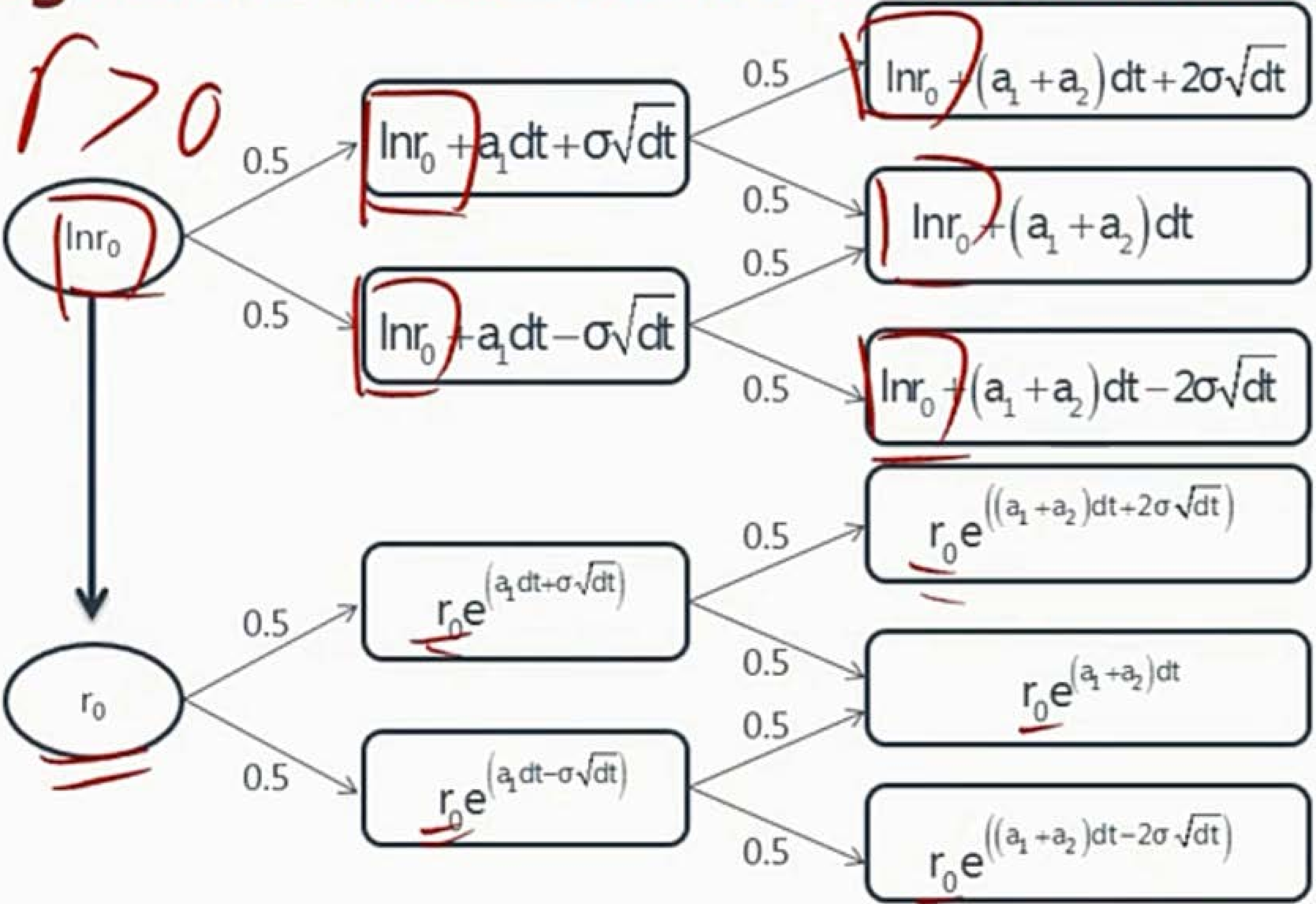
$$\begin{tikzcd}
 & e^{x_1} \\
e^{x} \arrow[ur, dash] \arrow[dr, dash] & \\
 & e^{x_2}
\end{tikzcd}$$

$$e^{\ln x} = x$$

$$e^{\ln r} = r$$

Lognormal Model with Deterministic Drift

$r > 0$



$$e^{\ln r_0 + a\,dt + \sigma\sqrt{dt}}$$

$$= \boxed{e^{\ln r_0}} \times e^{a\,dt + \sigma\sqrt{dt}}$$

$$= \underline{r_0}\, \underline{e}^{a\,dt + \sigma\sqrt{dt}}$$

The Art of Term Structure Models

- An analyst constructed an interest rate tree with monthly time steps, where t = 1/12. The current short-term rate is 3.0%. His term structure model assumes an annual basis point volatility of 200bps with an annual drift of 50bps. He employs Model 2 which assumes normally distributed rates and incorporating drift. Here is his rate tree:

Model 2: Normally Distributed Rates but Incorporating Annual Drift

What is the un-displayed missing value?

A. 1.93%

B. 2.17%

C. 2.38%

D. 3.01%

λ_1 λ_2

$r_0 + (\lambda_1 + \lambda_2)dt + \sigma dw$

- **Answer: A**

Reading 6

Volatility Smiles

186-203

What is Volatility Smile?

- **What is volatility smile?**
  - Volatility smile is a plot of the implied volatility of an option as a function of its strike price.
    - ✓ This chapter describes the volatility smiles that traders use in equity and **foreign currency** markets.

BSM

$$C = S_0 N(d_1) - ke^{-rt} N(d_2)$$

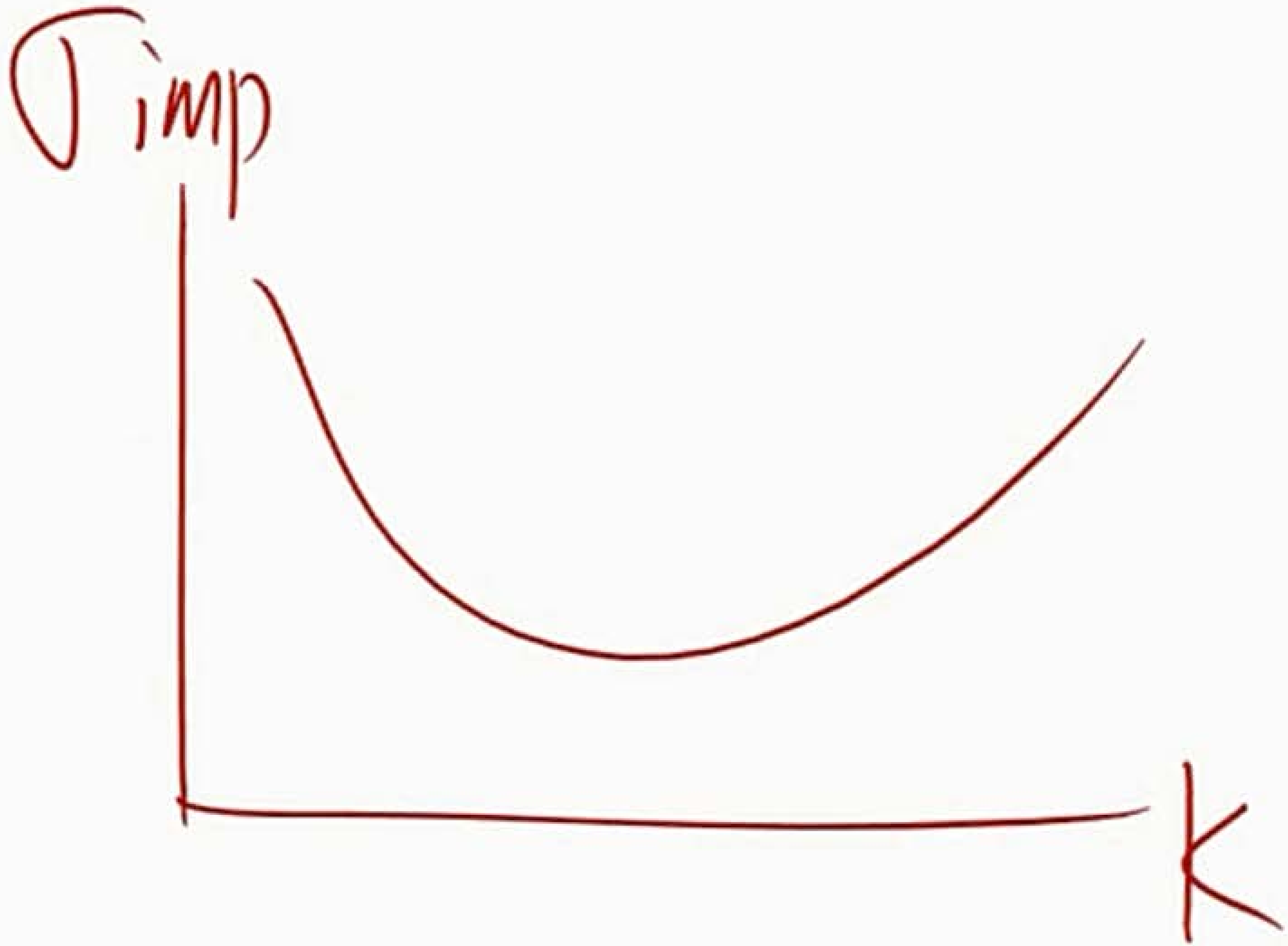
S_0 ✓

K ✓ $\xrightarrow{BSM}$ C_{BS}

r ✓ P_{BS}

T ✓

$\sigma = ?$


σ_{imp}
K

Volatility Smiles is the Same for Calls and Puts

- **Based on the put-call parity**

$$p + S_0 e^{-qT} = c + Ke^{-rT}$$

$$\begin{cases} p_{BS} + S_0 e^{-qT} = c_{BS} + Ke^{-rT} - - - - (1) \\ p_{mkt} + S_0 e^{-qT} = c_{mkt} + Ke^{-rT} - - - - (2) \end{cases}$$

- (1) - (2): we can get:

$$p_{BS} - p_{mkt} = c_{BS} - c_{mkt}$$

Volatility Smiles is the Same for Calls and Puts

- This shows that the dollar pricing error when the Black-Scholes model is used to price a European put option **should be exactly the same as** the dollar pricing error when it is used to pricing a European call option with the same strike price and time to maturity.

- **The implied volatility of a European call option is always the same as the implied volatility of a European put option** when the two have the same strike price and maturity date.

GOLDEN FUTURE

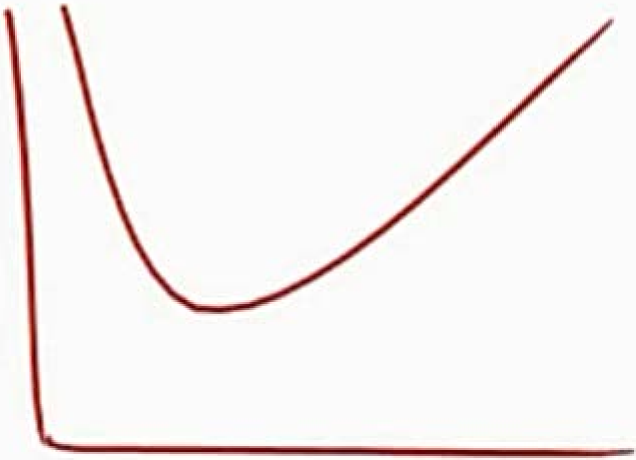
$$P_{BS} + S = C_{BS} + ke^{-rt} \quad ①$$

$$P_{mkt} + S = C_{mkt} + ke^{-rt} \quad ②$$

$$① - ② \quad P_{BS} - P_{mkt} = C_{BS} - C_{mkt}$$

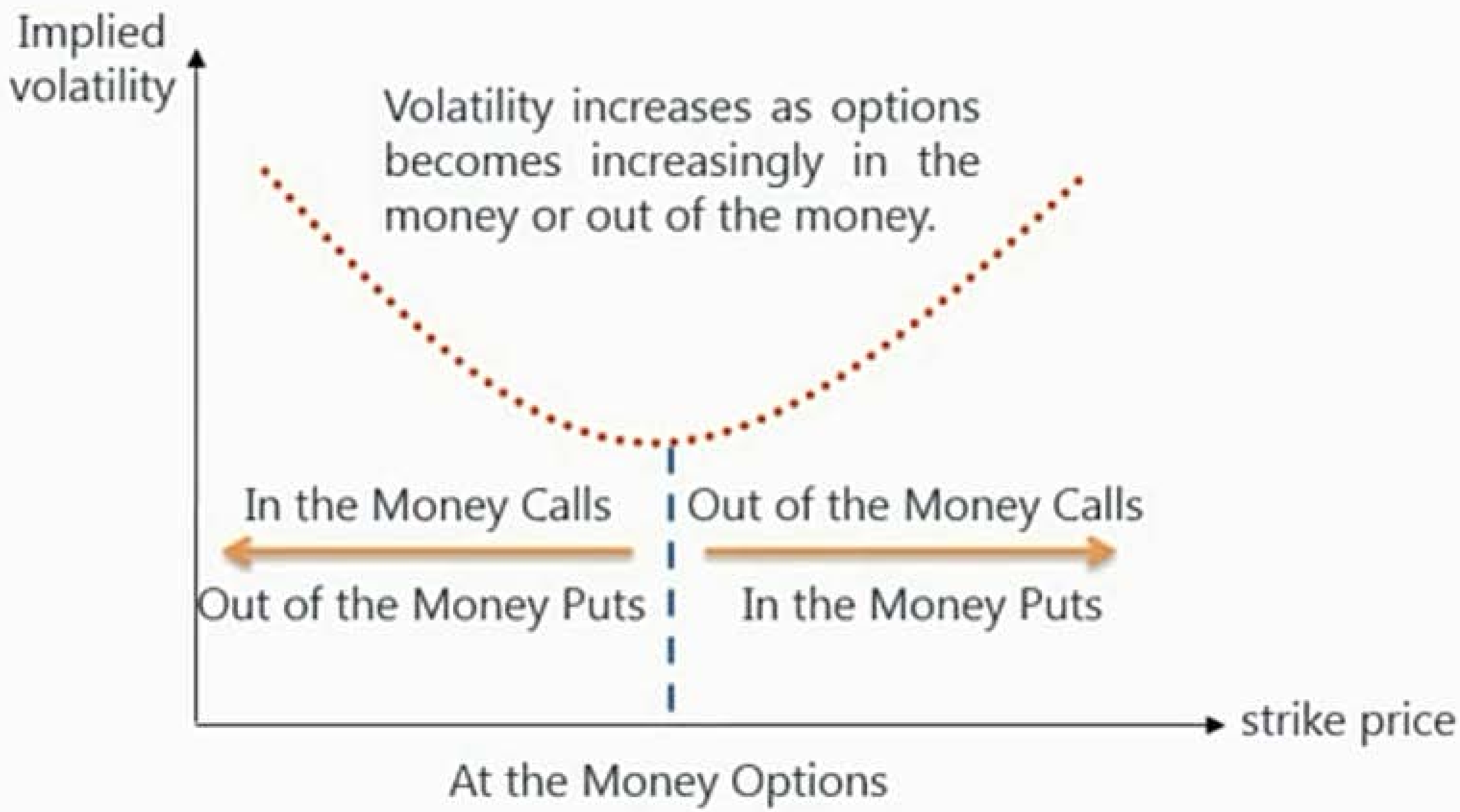
$$\underline{\sigma_{BS,p}} - \sigma_{mkt,p} = \sigma_{BS,c} - \sigma_{mkt,c}$$

$$\sigma_{\underset{\substack{\uparrow \\ imp}}{mkt},\,p} = \sigma_{\underset{\substack{\uparrow \\ imp}}{mkt},\,c}$$



Volatility Smile for Foreign Currency Options

- The implied volatility is relatively low for **at-the-money** options. It becomes progressively higher as an option moves either **into the money or out of the money.**



call

3/1 4/1 5/1

27 28 $\boxed{29}$

put

3/1 4/1 5/1

0.7 0.8 $\boxed{0.9}$

K=5

K=10

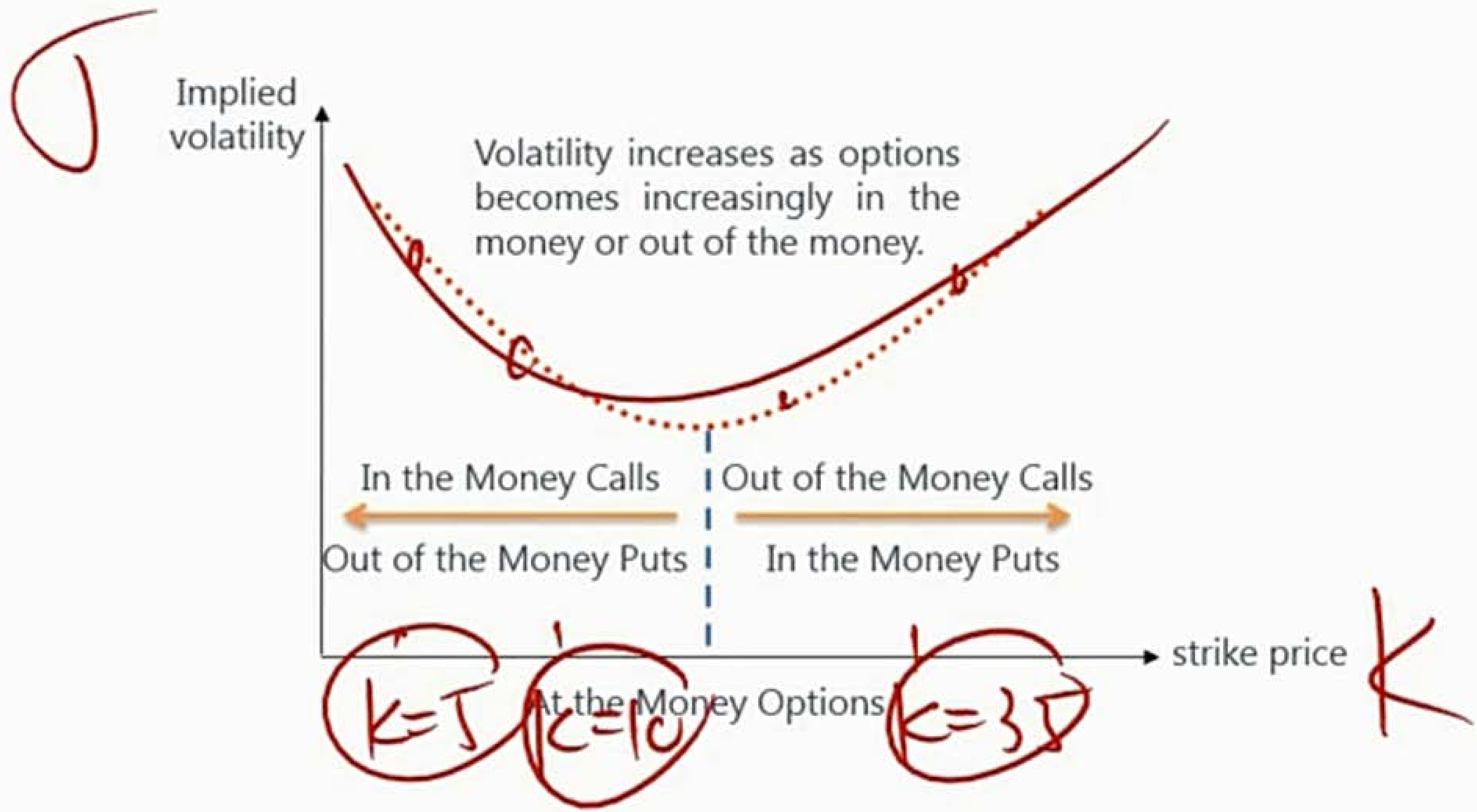
K=15

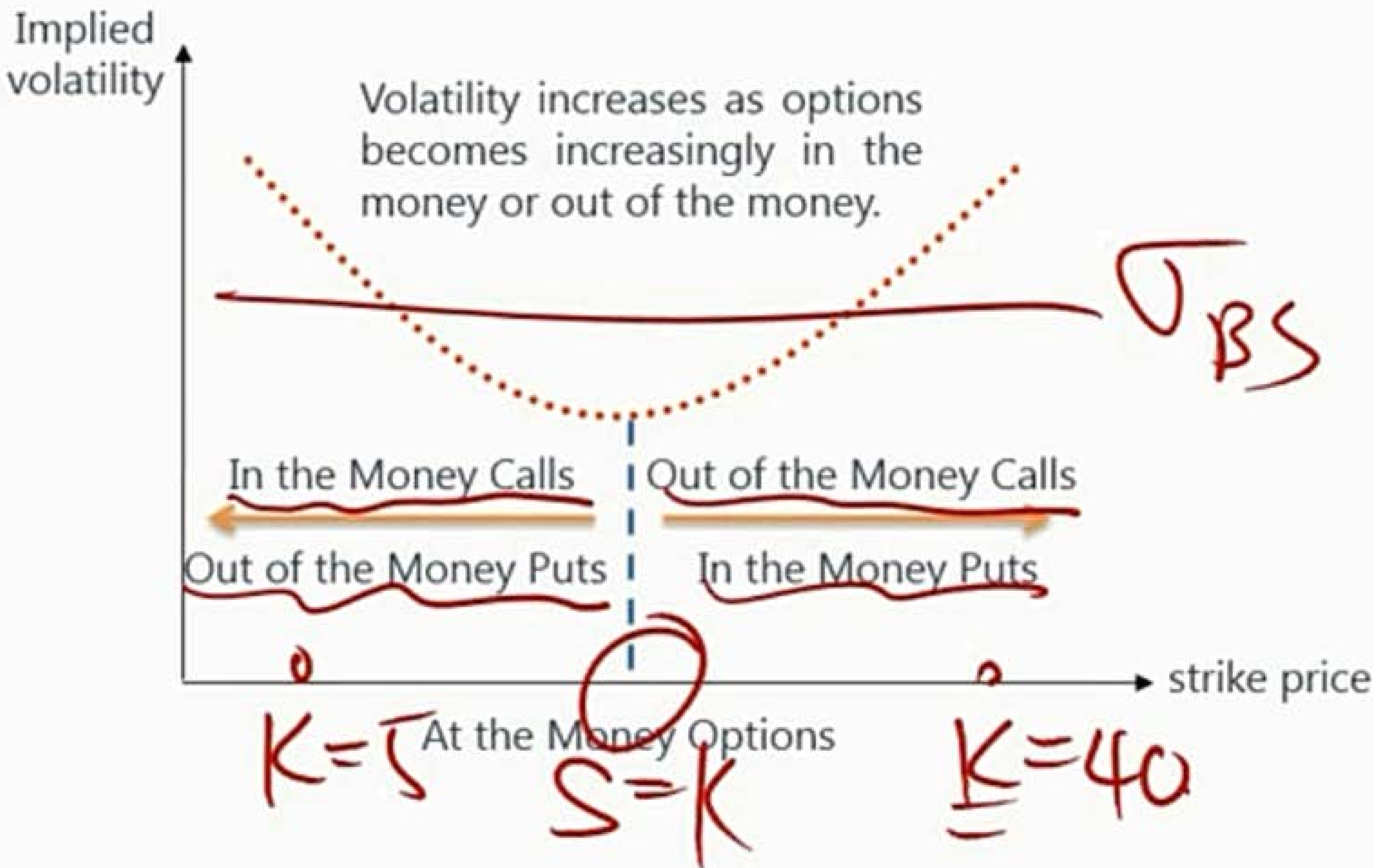
K=20

⋮

Volatility Smile for Foreign Currency Options

- The implied volatility is relatively low for **at-the-money** options. It becomes progressively higher as an option moves either **into the money or out of the money.**

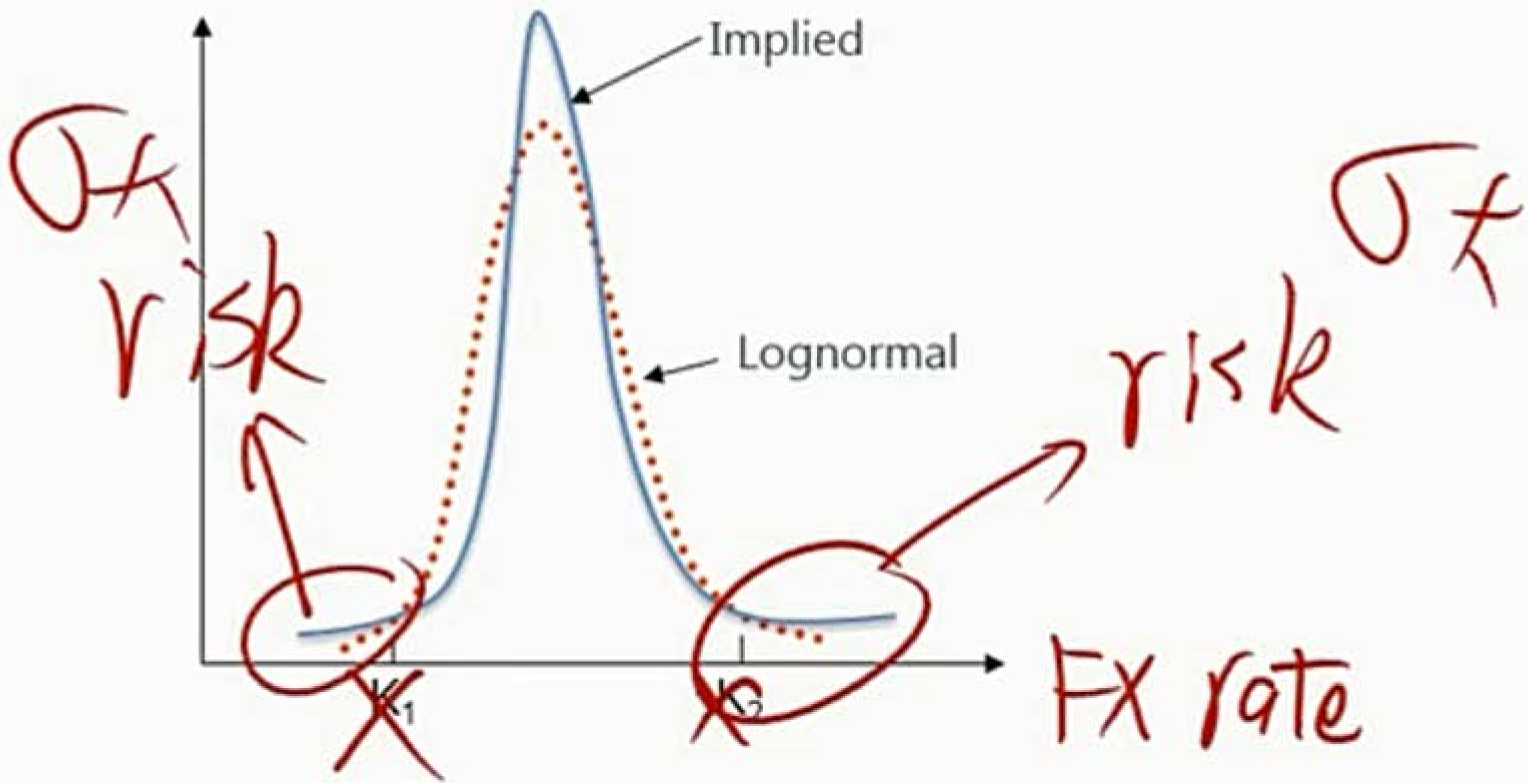




Implied volatility
Volatility increases as options becomes increasingly in the money or out of the money.
σ_{BS}
In the Money Calls
Out of the Money Calls
Out of the Money Puts
In the Money Puts
strike price
At the Money Options
K=5
S=K
K=40

Reasons for Smile in Foreign Currency Options

- The comparison of implied distribution and lognormal distribution for foreign currency options

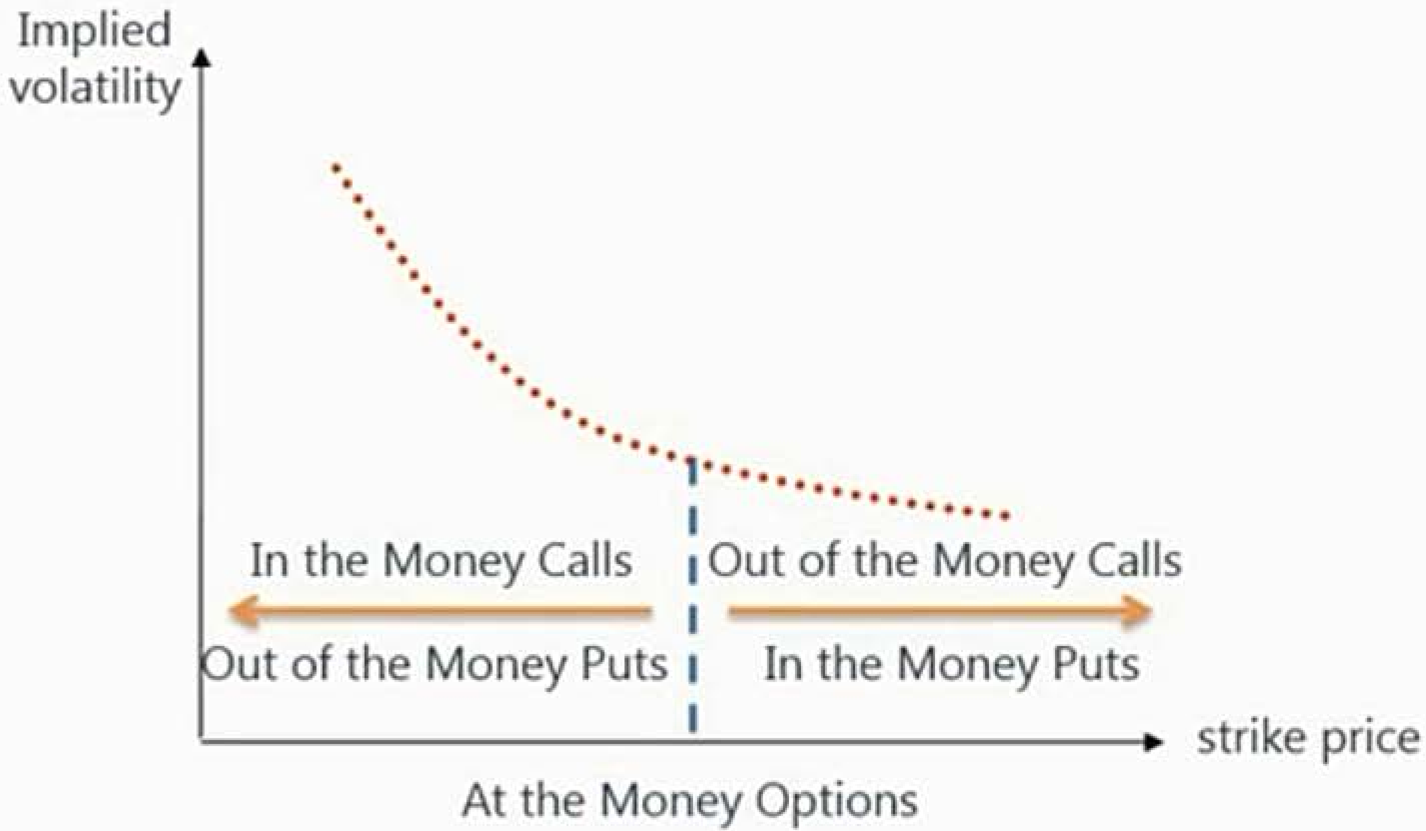


Reasons for Smile in Foreign Currency Options

- Why are exchange rate **not lognormally distributed?** Two of the contidions for an asset price to have a lognormal distribution are:
  - **The volatility of the asset is constant.**
  - **The price of the asset changes smoothly with no jumps.**
- **In practice, neither of these conditions is satisfied** for an exchange rate. The volatility of an exchange rate is far from constant, and exchange rates frequently exhibit jumps (sometimes the jumps are in response to the actions of **central banks**).

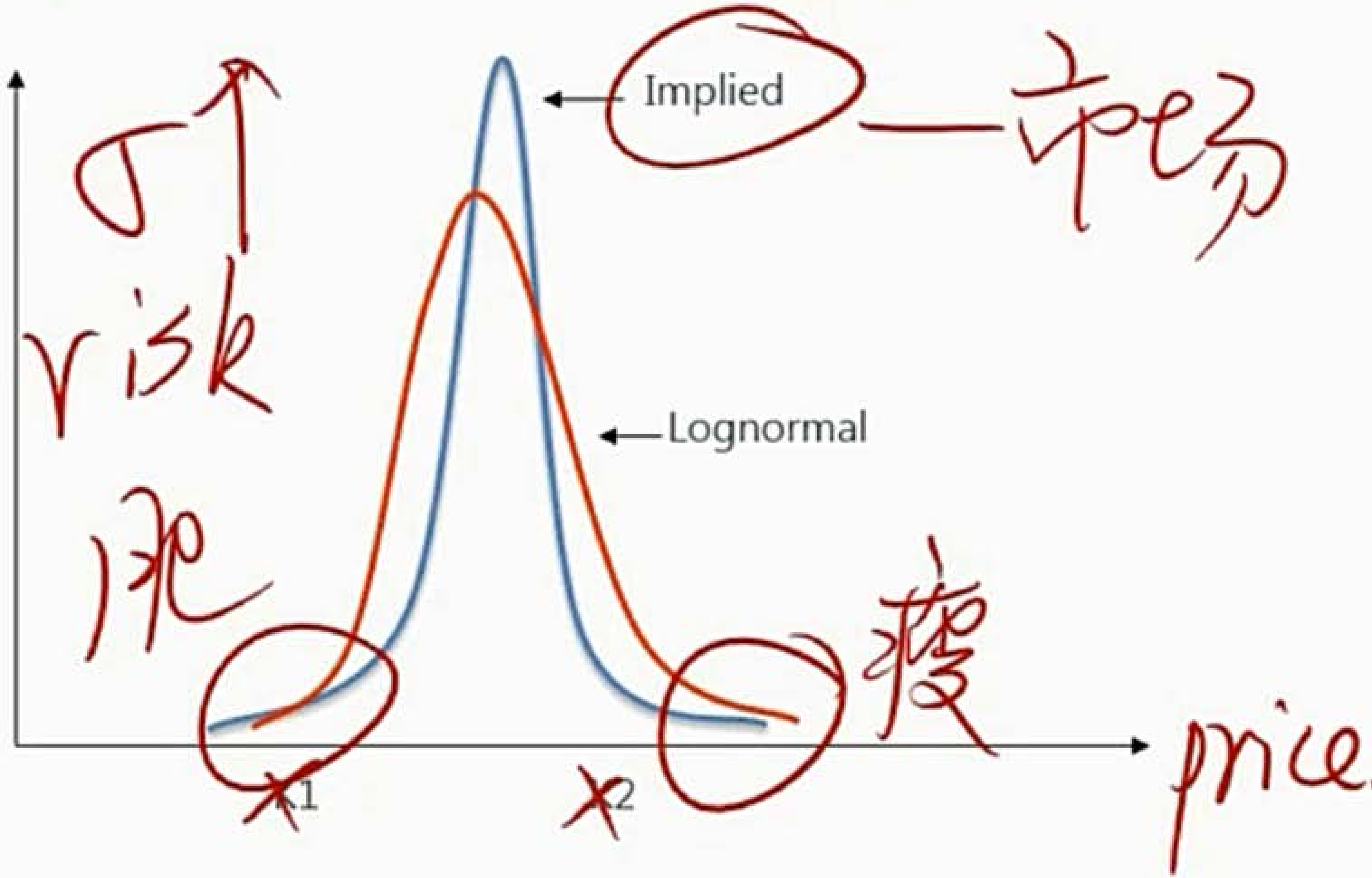
Volatility Smiles (skew) for Equity Options

- The volatility used to price a low-strike-price option (i.e., a deep out of the money put or a deep in the money call) is significantly higher than that used to price a high-strike-price option (i.e., a deep in the money put or a deep out of the money call).



Reasons for the Smile in Equity Options

➢ **Implied and lognormal distribution for equity options**



- It can be seen that the implied distribution has a heavier left tail and a less heavy right tail than the lognormal distribution.

Reasons for the Smile in Equity Options

- **Leverage (equity price→volatility)**
  - As a company's equity declines in value, the company's leverage increases. This means that the equity becomes more risky and its volatility increases.
- **Volatility Feedback Effect (volatility→equity price)**
  - As volatility increases (decreases) because of external factors, investors require a higher (lower) return and as a result the stock price declines (increases).
- **Crashophobia (expected equity price→implied volatility)**
  - 1987 stock market crash: higher premiums for put prices when the strike prices lower.

OTM put

$K = \underline{\underline{\$1}}$ $S = \$10$

$S < K$ 行权 $P_{OTM\,put} = \underline{\underline{0.0}}$

1987 +0

$10 ⟶ $0.8

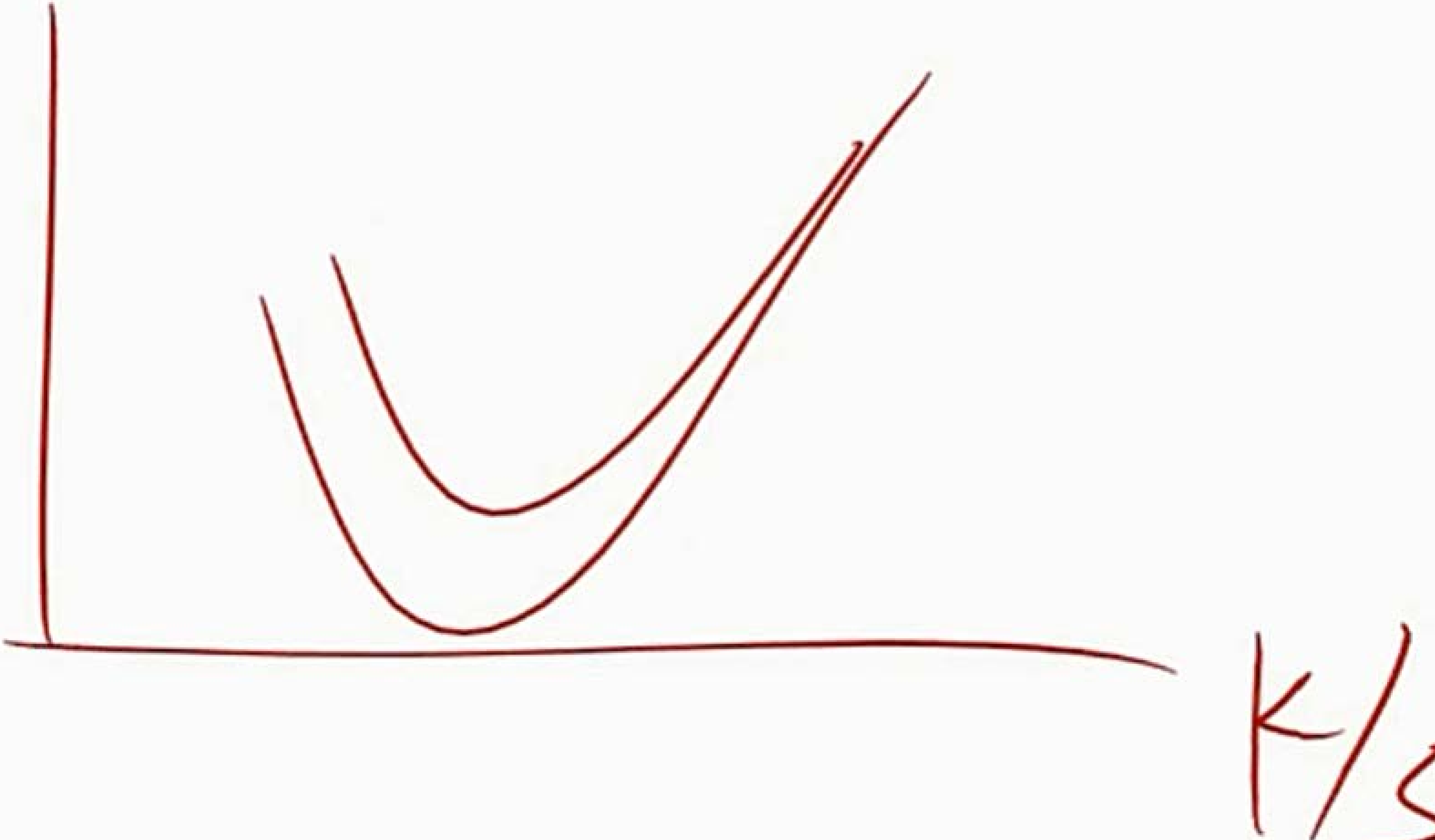
put ⟶ ITM

- **Leverage (equity price→volatility)**
  - As a company's equity declines in value, the company's leverage increases. This means that the equity becomes more risky and its volatility increases.

$\sigma = 30\%$

- **Volatility Feedback Effect (volatility→equity price)**
  - As volatility increases (decreases) because of external factors, investors require a higher (lower) return and as a result the stock price declines (increases).
- **Crashophobia (expected equity price→implied volatility)**
  - 1987 stock market crash: higher premiums for put prices when the strike prices lower.

OTM



a.

Reasons for Smiles in Foreign Currency Options

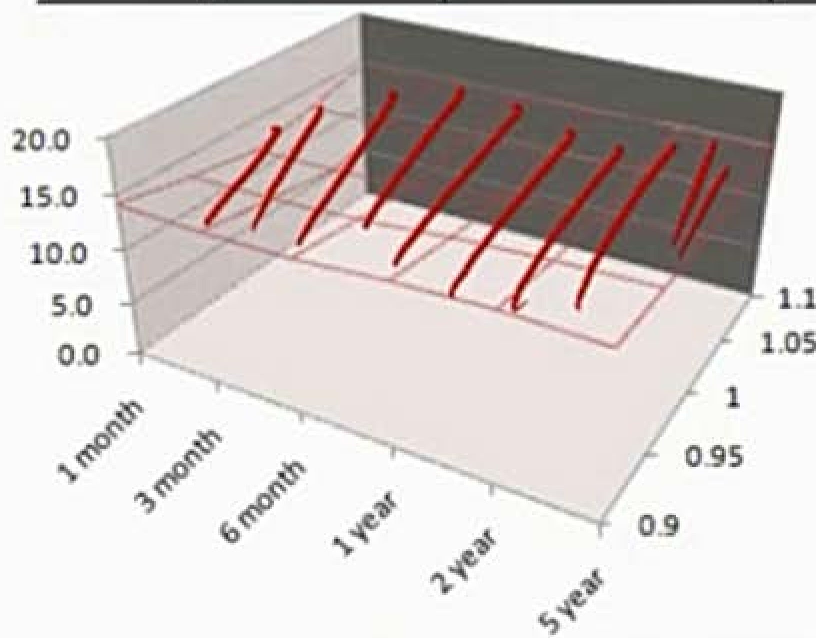
- **Alternative ways of characterizing the volatility smile**
  - The volatility smile is often calculated as the relationship between the **implied volatility** and **K/S_0** rather than as the relationship between the implied volatility and K.
    - ✓ A refinement of this is to calculate the volatility smile as the relationship between the **implied volatility** and **K/F_0**, where F_0 is the forward pirce of the asset for a contract maturing at the same time as the options that are considered.
  - Another approach to defining the volatility smile is as the relationship between **the implied volatility** and the **delta** of the option.

Volatility Term Structure and Volatility Surface

- Traders allow the implied volatility to **depend on time** to maturity as well as strike price.
- **Volatility surfaces** combine volatility smiles with the time to maturity and K/S_0.
  - Implied volatility tends to be an **increasing function** of maturity when short-dated volatilities are historically low.
  - Volatility tends to be a **decreasing function** of maturity when short-dated volatilities are historically high.

Volatility Term Structure and Volatility Surface

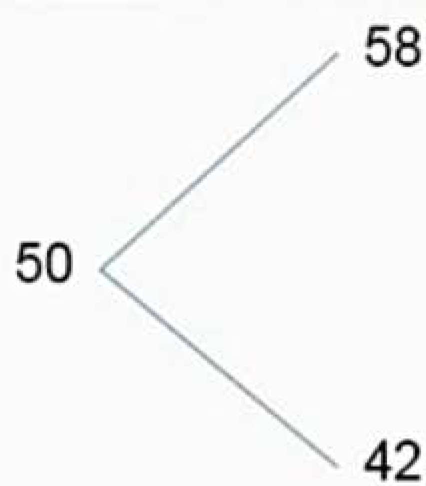
Volatility surface					
	K / S_0				
	0.90	0.95	1.00	1.05	1.10
1 month	14.2	13.0	12.0	13.1	14.5
3 month	14.0	13.0	12.0	13.1	14.2
6 month	14.1	13.3	12.5	13.4	14.3
1 year	14.7	14.0	13.5	14.0	14.8
2 year	15.0	14.4	14.0	14.5	15.1
5 year	14.8	14.6	14.4	14.7	15.0



The Impact of Large Asset Price Jumps on Volatility Smiles

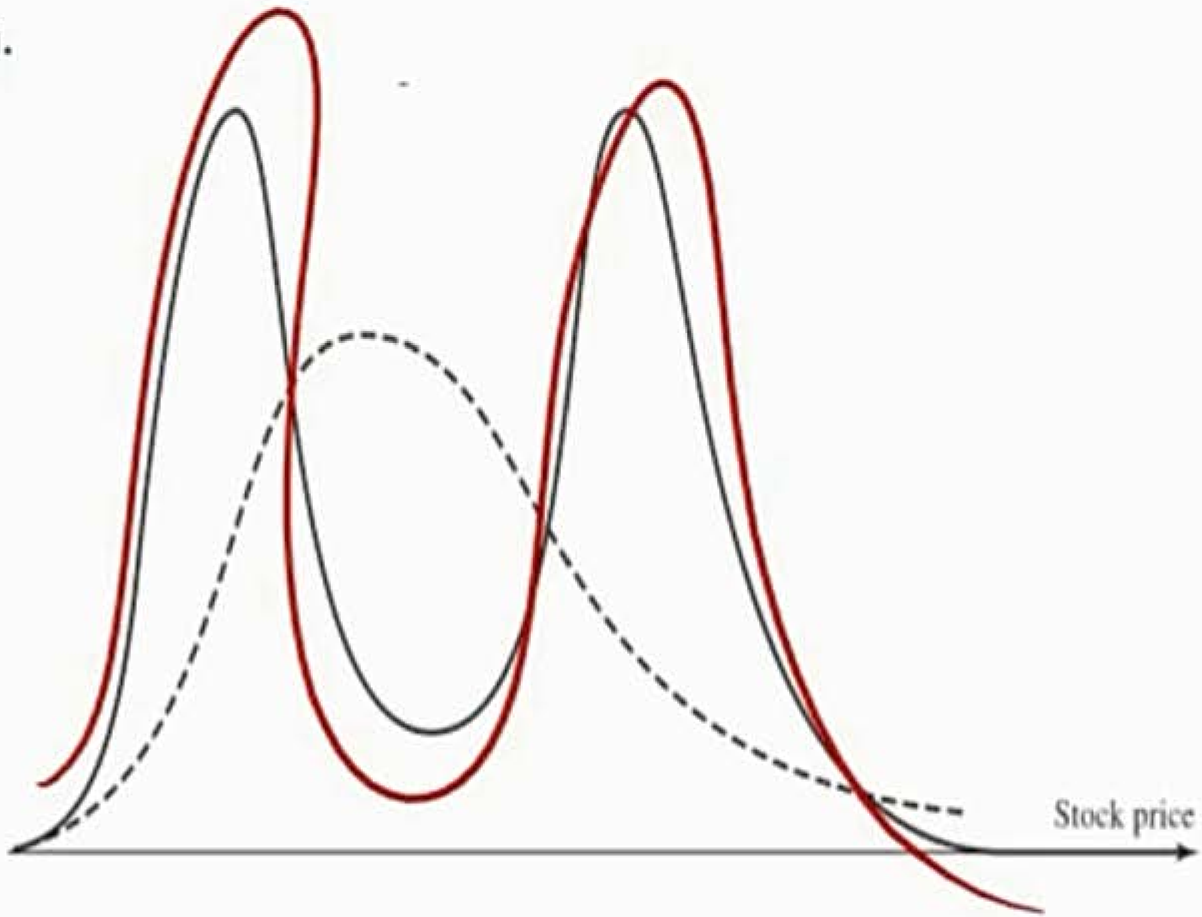
- Suppose that a stock price is currently \$50 and an important news announcement due in a few days is expected either to increase the stock price by \$8 or to reduce it by \$8.

Change in stock price in 1 month



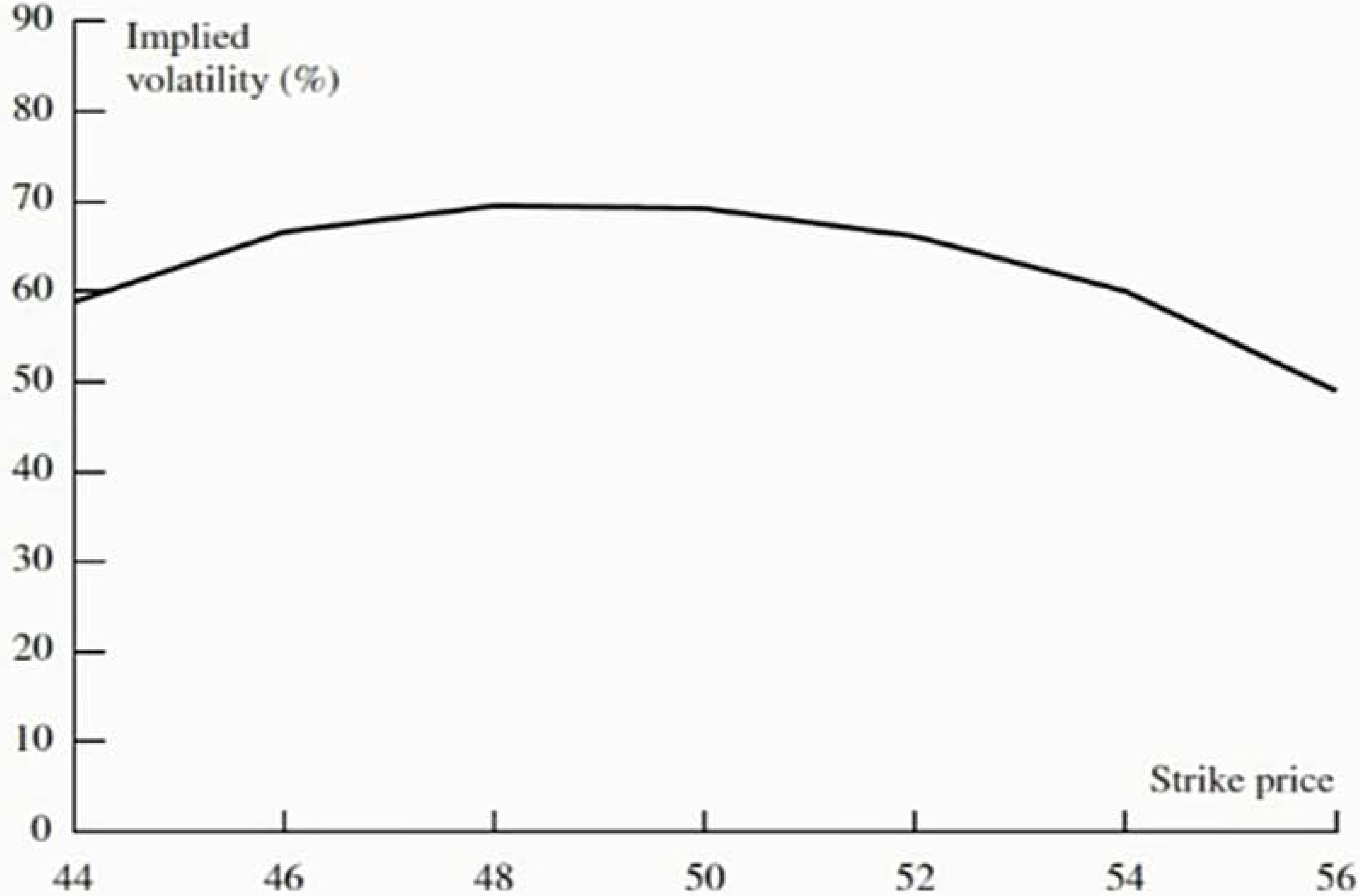
The Impact of Large Asset Price Jumps on Volatility Smiles

- The probability distribution of the stock **price consist of a mixture of two lognormal distribution**
  - The solid line is the true distribution; the dashed line is the lognormal distribution.



The Impact of Large Asset Price Jumps on Volatility Smiles

- **Volatility Frown**



Volatility Smiles

- Which of the following statements is incorrect regarding volatility smiles?
  A. Currency options exhibit volatility smiles because the at-the-money options have higher implied volatility than away-from-the-money options.
  B. Volatility frowns result when jumps occur in asset prices.
  C. Equity options exhibit a volatility smirk because low strike price options have greater implied volatility.
  D. Relative to currency traders, it appears that equity traders' expectations of extreme price movements are more asymmetric.

- **Answer: A**