In the mathematical finance literature, most articles written in the last two and half decades start with the words "Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, ...". Whether it is explicitly mentioned or not in the articles or texts, the probability space is always there somewhere in the background. The triple $(\Omega, \mathcal{F}, \mathbb{P})$, is called a *probability space*, and its inclusion in quantitative finance literature refects the increasing infuence of probability theory and probability theorists over the subject. It forms the foundation of the *probabilistic universe*.

This probability space comprises of three components

1. the sample space Ω ✓

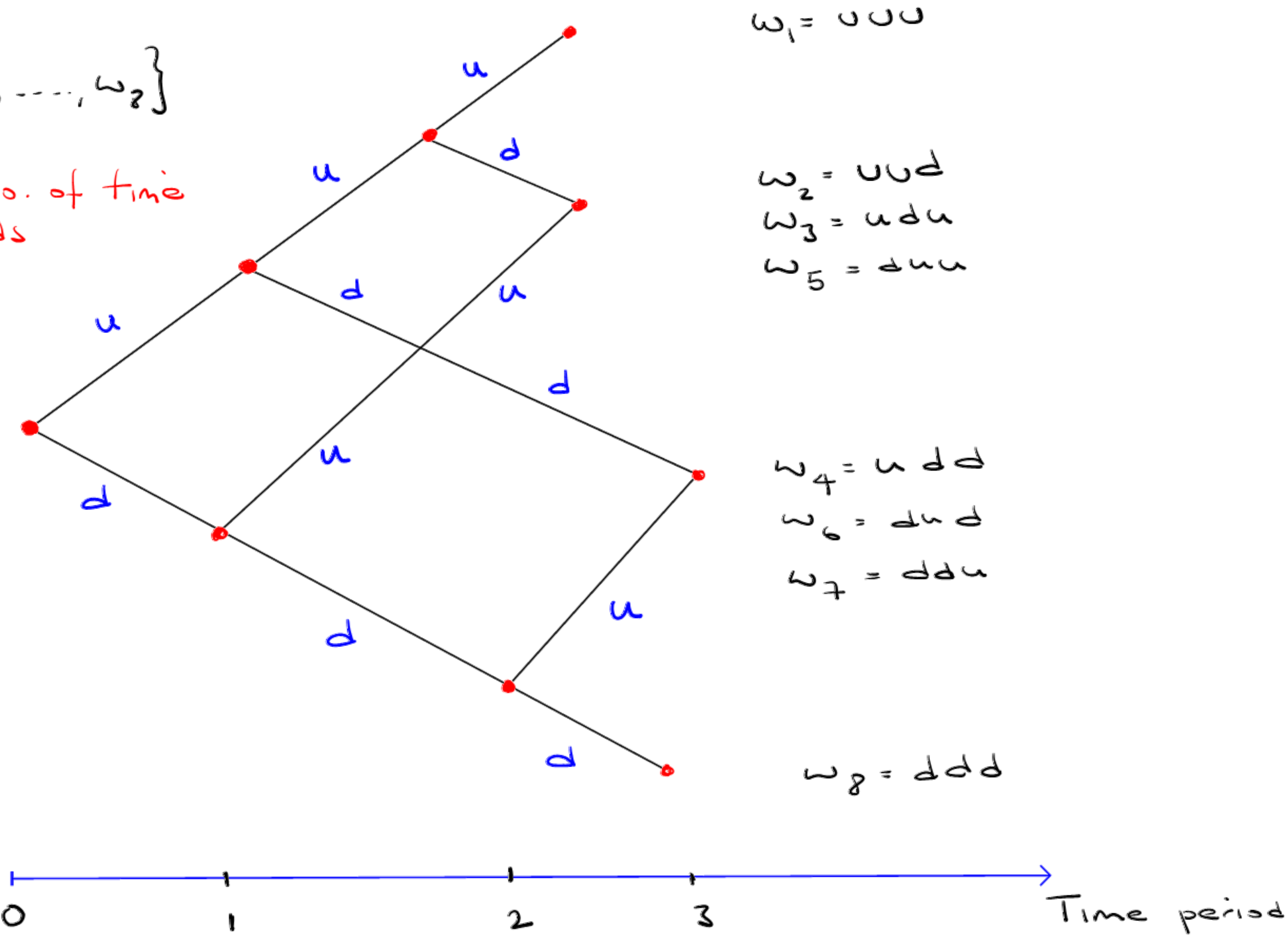
2. the filtration $\mathcal{F}$

3. the probability measure $\mathbb{P}$.

$\Omega_3 = \{\omega_1, \omega_2, \ldots, \omega_8\}$

Ω_n n – no. of time periods

$|\Omega_n| = 2^n$



Example (3-period binomial model). Take $n = 3$; so $\Omega = \Omega_3$, given by the finite set

$$\Omega = \{UUU, UUD, UDU, UDD, DUU, DUD, DDU, DDD\},$$

Or the set of all possible outcomes of three coin tosses.

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

Now define the following two subsets of Ω :

$$A_U = \{UUU, UUD, UDU, UDD\}, \; A_D = \{DUU, DUD, DDU, DDD\}.$$

We see A_U is the subset of outcomes where a Head appears on the first throw, A_D is the subset of outcomes where a Tail lands on the first throw.

Define also

$$\begin{aligned} A_{UU} &= \overbrace{\{UUU, UUD\}}^{\leftarrow\, 2\, \rightarrow}, \; A_{UD} = \{UDU, UDD\}, \\ A_{DU} &= \{DUU, DUD\}, \; A_{DD} = \{DDU, DDD\} \end{aligned}$$

corresponding to the events that the first two coin tosses result in HH, HT, TH, TT respectively. At the end of each period, new information becomes available to help us predict the actual stock trajectory.

At time 0, before the start of trading, we only have the trivial filtration

$$\mathcal{F}_0 = \{\Omega, \emptyset\}$$

since we have no information regarding the trajectory of the stock.

$\Omega = 2^3$

So e.g. three trading periods e.g. $\overbrace{\text{9am - 11am}}^{1}$; $\overbrace{\text{11am - 1pm}}^{2}$; $\overbrace{\text{1pm - closing bell}}^{3}$. After the first trading period we know whether the initial move was an up move or down move. Hence

$\emptyset$

elegant record of how information builds & is stored at each time step.

$$\mathcal{F}_1 = \{\Omega, \emptyset, A_U, A_D\}$$

ϕ

$$\mathcal{F}_2 = \{\Omega, \emptyset, A_U, A_D, A_{UU}, A_{UD}, A_{DU}, A_{DD}, \text{ unions and complements}\}$$

The filtration is an indication of how information about a probabilistic experiment builds up over time as more results become available. It can be thought of as an increasing family of events. The pair $(\Omega, \mathcal{F})$ is called a *measurable space*.

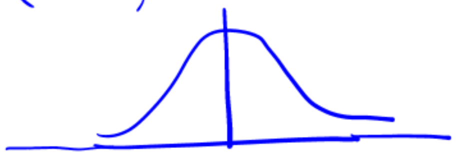
$A_u \leftarrow U$ D

C.D.F ? Assigns probabilities to outcomes.

We place a *probability measure* $\mathbb{P}$ on $\{\Omega, \mathcal{F}\}$. $\mathbb{P}$ is a special type of "function", called a measure which assigns probabilities to subsets (i.e. the outcomes); the theory also comes from Measure Theory. Whereas cumulative density functions (CDF) are defined on intervals such as $\mathbb{R}$, probability measures are defined on general sets, giving greater power, generlaisation and flexibility. A probability measure $\mathbb{P}$ is a function mapping $\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$ with the properties

$\Phi \sqrt{dt}$ $\qquad \phi \sim N(0,1)$

$X : \omega \in \Omega \longrightarrow \mathbb{R}$ $\quad (-\infty, \infty)$



(i) $\mathbb{P}(\Omega) = 1$,

$$F(x) = \int_{-\infty}^{x} p(x)\, dx \qquad x \quad x+dx$$

S.D.E

(ii) if $A_1, A_2, \ldots.$ is a sequence of disjoint sets in $\mathcal{F}$, then $\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$.

Measure theory $\qquad$ probability fine

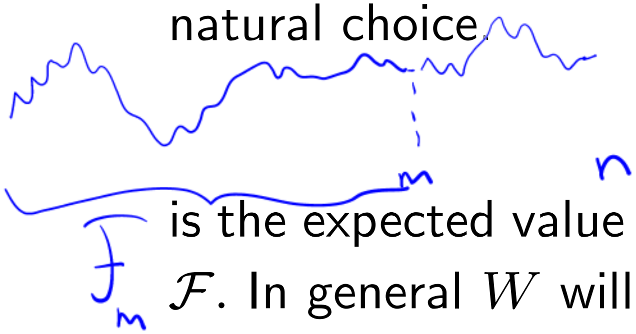
Set theory $\qquad$ small print

B.M.

$$\longrightarrow (\overbrace{\underbrace{\Omega, \mathcal{F}}_{\text{measure space}}, \mathbb{P}}^{\text{prob space.}}) \qquad (\Omega, \mathcal{F}) \qquad \int_{-\infty}^{\infty} \mathbb{P} \longrightarrow \text{prob. space}$$

Conditional Expectations

What makes a conditional expectation different (from an unconditional one) is information (just as in the case of conditional probability). In our probability space, $(\Omega, \mathcal{F}, \mathbb{P})$ information is represented by the filtration $\mathcal{F}$; hence a conditional expectation with respect to the (usual information) filtration seems a natural choice.



process   always refer to

$$\mathbb{E}_m [W_n | \mathcal{F}_m] \, ; \; n > m$$

$T < t$

t   T

is the expected value of the random variable conditional upon the filtration set $\mathcal{F}$. In general W will be adapted to the filtration $\mathcal{F}$.

Adapted (Measurable) Process

$$\mathbb{E}_t [S_T | \mathcal{F}_t] \qquad t < T$$

stock price

A stochastic process S_t is said to be adapted to the filtration $\mathcal{F}_t$ (or measurable with respect to $\mathcal{F}_t$, or $\mathcal{F}_t$ -adapted) if the value of S_t at time t is known given the information set $\mathcal{F}_t$.

$\int_{-\infty}^{\infty}$ $\int_{\mathbb{R}}$

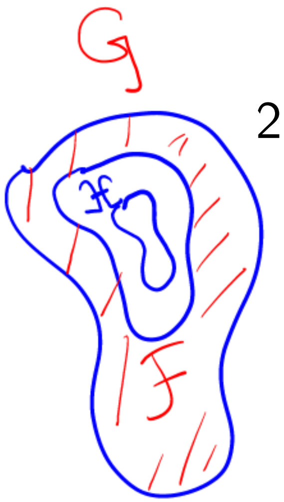
Conditional expectations have the following useful properties: If X, Y are integrable random variables and α, β are constants then we list the following results:

X, Y $\qquad$ $\mathbb{E}[|X|] < \infty$

$\mathbb{E}[|Y|] < \infty$

1. **Linearity**:

$$\mathbb{E}[(\alpha X + \beta Y) | \mathcal{F}] = \alpha \mathbb{E}[X | \mathcal{F}] + \beta \mathbb{E}[Y | \mathcal{F}]$$



2. **Tower Property (i.e. Iterated Expectations)**: if $\mathcal{F} \subset \mathcal{G}$

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{F}] = \mathbb{E}[X | \mathcal{F}]$$

This property states that if taking iterated expectations with respect to several levels of information, we may as well take a single expectation subject to the smallest set of available information. The moral being "the smallest filtration always wins".

$$\mathbb{E}\Big[\mathbb{E}\big[\mathbb{E}[X | \mathcal{G}] \big| \mathcal{F}\big] \Big| \mathcal{H}\Big] = \mathbb{E}[X | \mathcal{H}]$$

3. As a special case of the Tower property, we have

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{F}\right]\right] = \mathbb{E}\left[X\right]$$

since "no filtration" is always a smaller information set than any filtration

4. **Taking Out What Is Known**: if X is $\mathcal{F}$-measurable, then the value of X is known once we know $\mathcal{F}$. Therefore,

$\mathcal{G}$ $\quad$ $\mathcal{G} \subset \mathcal{F}$ $\quad$ $\mathcal{G} \subseteq \mathcal{F}$ $\quad$ X_t $\quad$ $\mathcal{F}_t$

$$\mathbb{E}_t\left[X_t|\mathcal{F}_t\right] = X_t$$

5. **Taking Out What Is Known (2)**: by extension, if X is $\mathcal{F}$-measurable but not Y,

$$\mathbb{E}\left[XY|\mathcal{F}\right] = X\,\mathbb{E}\left[Y|\mathcal{F}\right]$$

6. **Independence**: if X is independent from $\mathcal{F}$, then knowing $\mathcal{F}$ is useless to predict the value of X. Hence,

$$\boxed{\mathbb{E}\left[X|\mathcal{F}\right] = \mathbb{E}\left[X\right]}$$

X stock price

$\mathcal{F}$ filtration for I.R.

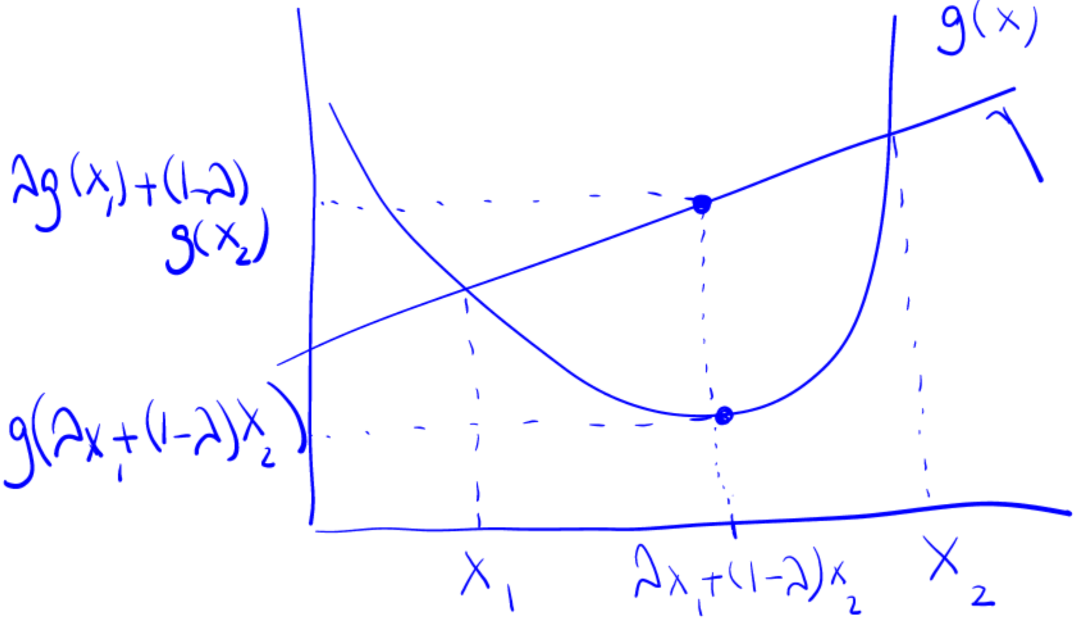
7. **Positivity:** if $X \geq 0$ then $\boxed{\mathbb{E}[X|\mathcal{F}] \geq 0}$.

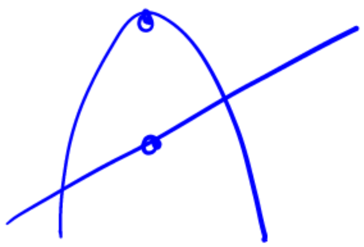
8. **Jensen's Inequality**: First formally define *convexity*. A function $g : I \to \mathbb{R}$ is a convex function if for $x_1,\ x_2$, in I and any $\lambda \in [0, 1]$

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2)$$

If f is a convex function, then

$$f(\mathbb{E}[X|\mathcal{F}]) \leq \mathbb{E}[f(X)|\mathcal{F}]$$





$$\boxed{\Omega_1 \longrightarrow \Omega_2}$$

Changing Probability Measure

You have seen in the Binomial Model lecture that there is more than just one probability measure.

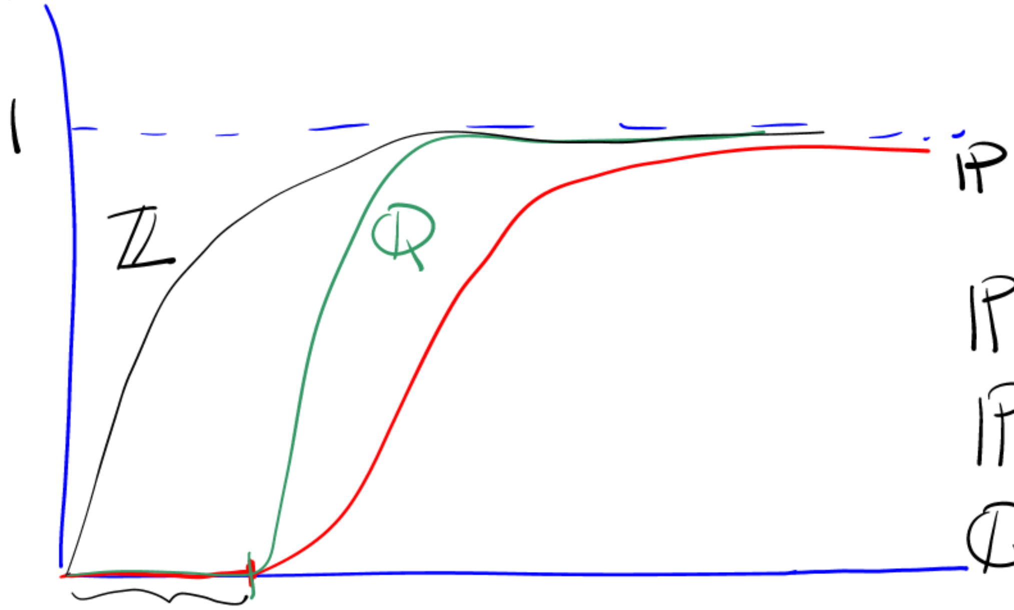
$\Omega_1 = (p, 1-p) \qquad \Omega_2 = (q, 1-q)$

Indeed, the lecture introduced you to the distinction between the "real" or "physical" probability measure, which we encounter every day on our Bloomberg or Reuters screen, and the so-called "risk-neutral" measure, which is used for pricing.

Probability measures are by no means unique. We will see in the next lecture that the powerful arsenal of martingale techniques enables us, under certain assumptions, to change measure and transpose our problem subject to the real world measure into an equivalent problem formulated as a martingale under a different measure.

For now, we just outline the rules that allow us to define equivalent measures.

3 Measures:



$\mathbb{P} \sim \mathbb{Q}$

$\mathbb{P} \nsim \mathbb{Z}$

$\mathbb{Q} \nsim \mathbb{Z}$

Itô Product Rule:

2 stochastic processes X_t, Y_t. Consider $F = XY$

SDE $dF = d(XY)$

Consider $F = XY$ $\quad X \to X + dX \quad Y \to Y + dY$

$dF = F(X+dX, Y+dY) - F(X,Y)$

$$dF = \frac{\partial F}{\partial X} dX + \frac{\partial F}{\partial Y} dY + \frac{1}{2}\frac{\partial^2 F}{\partial X^2} dX^2 + \frac{1}{2}\frac{\partial^2 F}{\partial Y^2} dY^2 + \frac{\partial^2 F}{\partial X \partial Y} dX dY$$

$F = XY$ $\quad \frac{\partial F}{\partial X} = Y \; ; \; \frac{\partial^2 F}{\partial X^2} = 0 \; ; \; \frac{\partial F}{\partial Y} = X \; ; \; \frac{\partial^2 F}{\partial Y^2} = 0 \; ; \; \frac{\partial^2 F}{\partial X \partial Y} = 1 = \frac{\partial^2 F}{\partial Y \partial X}$

$$\boxed{dF = Y\,\underline{dX} + X\,\underline{dY} + dX\,dY}$$

$\rightarrow dX = a\,dt + b\,dW_1$

$\rightarrow dY = c\,dt + e\,dW_2$

$dX\,dY$

$$\boxed{dW_1\,dW_2 = \rho\,dt}$$

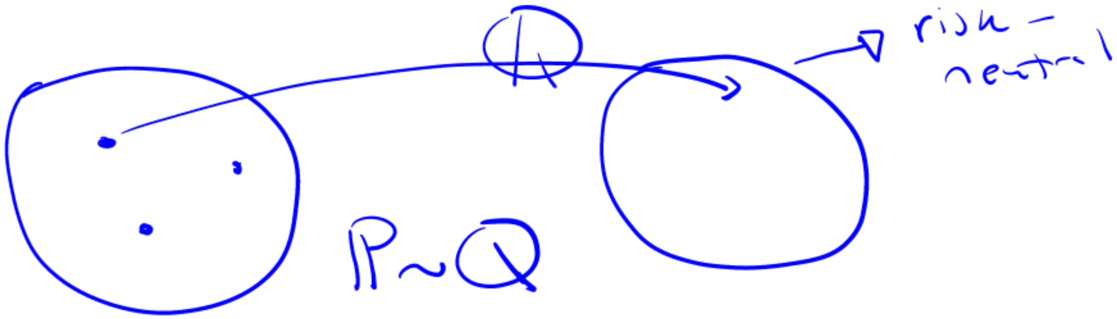
$$G = \frac{X}{Y}$$

$$d\left(\frac{X}{Y}\right)$$

physical $\leftarrow \mathbb{P}$

physical

Equivalent Measure



$\mathbb{P} \sim \mathbb{Q}$

If two measures $\mathbb{P}$ and $\mathbb{Q}$ share the same sample space Ω and if $\mathbb{P}(A) = 0$ implies $\mathbb{Q}(A) = 0$ for all subset A, we say that $\mathbb{Q}$ is **absolutely continuous** with respect to $\mathbb{P}$ and denote this by $\boxed{\mathbb{Q} << \mathbb{P}}$.

$\mathbb{P} << \mathbb{Q}$

The key point is that all impossible events under $\mathbb{P}$ remain impossible under $\mathbb{Q}$. The probability mass of the possible events will be distributed differently under $\mathbb{P}$ and $\mathbb{Q}$. In short "it is alright to tinker with the probabilities as long as we do not tinker with the (im)possibilities"

If $\mathbb{Q} << \mathbb{P}$ and $\mathbb{P} << \mathbb{Q}$ then the two measures are said to be **equivalent**, denoted by $\mathbb{P} \sim \mathbb{Q}$.

real prob. risk-neutral

$\frac{1}{4}$ $\mathbb{P}$

$\mathbb{P}$ $\mathbb{P}$

$\boxed{\begin{array}{c} \mathbb{P} << \mathbb{Q} \\ \mathbb{Q} >> \mathbb{P} \end{array}}$

$\mathbb{Q}$ $1/6$

$\mathbb{P} \sim \mathbb{Q}$