Stochastic process for macro

Tianxiao Zheng
SAIF

1. Stochastic process

The state of a system $\{X_t\}$ evolves probabilistically in time. The joint probability distribution is given by $\Pr(X_{t_1}, t_1; X_{t_2}, t_2; X_{t_3}, t_3; ...; X_{t_m}, t_m; ...)$. $(t_1 > t_2 > ...t_m > ...)$

Under Markov assumption, *i.e.* knowledge of the most recent condition determines the future, the conditional probability distribution is given by

$$\Pr(X_{t_1}, t_1 | X_{t_2}, t_2; X_{t_3}, t_3; ...) = \Pr(X_{t_1}, t_1 | X_{t_2}, t_2)$$

It is not difficult to see that under the Markov assumption

$$\begin{aligned} \Pr(X_{t_1}, t_1; X_{t_2}, t_2; X_{t_3}, t_3; ...; X_{t_m}, t_m) =& \Pr(X_{t_1}, t_1 | X_{t_2}, t_2) \Pr(X_{t_2}, t_2 | X_{t_3}, t_3) ... \\ & \Pr(X_{t_{m-1}}, t_{m-1} | X_{t_m}, t_m) \Pr(X_{t_m}, t_m) \end{aligned} \tag{1}$$

provided that $t_1 > t_2 > ...t_{m-1} > t_m$.

For stochastic processes, we have the following identity (law of total probability),

$$\Pr(X_{t_1} = x_{t_1} | X_{t_3} = x_{t_3}) = \sum_{x_{t_2}} \Pr(X_{t_1} = x_{t_1} | X_{t_2} = x_{t_2}; X_{t_3} = x_{t_3}) \Pr(X_{t_2} = x_{t_2} | X_{t_3} = x_{t_3})$$

For Markov process, this can be simplified. If $t_1 > t_2 > t_3$

$$\Pr(X_{t_1} = x_{t_1} | X_{t_3} = x_{t_3}) = \sum_{x_{t_2}} \Pr(X_{t_1} = x_{t_1} | X_{t_2} = x_{t_2}) \Pr(X_{t_2} = x_{t_2} | X_{t_3} = x_{t_3})$$

This equation is called **Chapman-Kolmogorov** equation. If x is continuous, the sum should be changed to integral and the Chapman-Kolmogorov equation is given by

$$\Pr(x_{t_1}, t_1 | x_{t_3}, t_3) = \int dx_{t_2} \Pr(x_{t_1}, t_1 | x_{t_2}, t_2) \Pr(x_{t_2}, t_2 | x_{t_3}, t_3)$$

2. Discrete state space: Markov chain

If time is discrete $[t = 0, 1, ..., T\ (T \leq \infty)]$, the Markov process in this case can be described by a transition matrix

$$P_{ij} = \Pr(X_{t+1} = e_j | X_t = e_i)$$

If the transition matrix does not depend on time, x_t is called time-invariant discrete-time discrete-space Markov process. This kind of Markov process plays an important role in the study of economic dynamics.

2.1. Properties of the transition matrix

- $\sum_{j=1}^{n} P_{ij} = 1$. A matrix satisfies this property (and $P_{ij} \geq 0$ for all i, j) is called a stochastic matrix.
- A stochastic matrix defines the probabilities of moving from each value of the state to any other in one period. The probability of moving from one value of the state to any other in two periods is determined by P^2 as a result of the Chapman-Kolmogorov equation

$$\Pr(X_{t+2} = e_j | X_t = e_i) = \sum_{h=1}^{n} \Pr(X_{t+2} = e_j | X_{t+1} = e_h) \Pr(X_{t+1} = e_h | X_t = e_i) = \sum_{h=1}^{n} P_{ih} P_{hj} = P_{ij}^2$$

- By iterating the equation, we have the $\Pr(X_{t+k} = e_j | X_t = e_i) = P_{ij}^k$. Therefore, in the case of Markov chain, the Chapman-Kolmogorov equation is simply

$$\begin{aligned} \Pr(X_{t+m+k} = e_j | X_t = e_i) &= \sum_{h} \Pr(X_{t+m+k} = e_j | X_{t+k} = e_h) \Pr(X_{t+k} = e_h | X_t = e_i) \\ &= \sum_{h=1}^{n} P_{ih}^k P_{hj}^m = P_{ij}^{m+k} \end{aligned}$$

2.2. Time evolution

x_t at time t can be described by a vector $\pi_t = [\pi_{t,1}, \pi_{t,2}, ..., \pi_{t,n}]'$, where $\pi_{t,j}$ $(j = 1, 2, ..., n)$ is the probability of x_t being e_j so that $\sum_{j=1}^{n} \pi_{t,j} = 1$. As a result, π_t evolves as

$$\pi'_{t+1} = \pi'_t P$$

Example 2.1. *Regime shift*

Two state representing a boom and a recession, respectively. The transition matrix is given by

$$P = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}, \quad p, q \in (0,1)$$

The m-step transition matrix is given by

$$P^m = \frac{1}{2-p-q} \begin{bmatrix} 1-q & 1-p \\ 1-q & 1-p \end{bmatrix} + \frac{(p+q-1)^m}{2-p-q} \begin{bmatrix} 1-p & p-1 \\ q-1 & 1-p \end{bmatrix}$$

Example 2.2. *Cyclical moving subset*

$$P = \begin{bmatrix} 0 & P_1 \\ P_2 & 0 \end{bmatrix}$$

where P_1 is a $k \times (n-k)$ matrix and P_2 is a $(n-k) \times k$ matrix. The m-step transition matrix is given by

$$P^{2m} = \begin{bmatrix} (P_1P_2)^m & 0 \\ 0 & (P_1P_2)^m \end{bmatrix}, \quad P^{2m+1} = \begin{bmatrix} 0 & (P_1P_2)^m P_1 \\ (P_1P_2)^m P_2 & 0 \end{bmatrix}$$

Example 2.3. *Random walk*

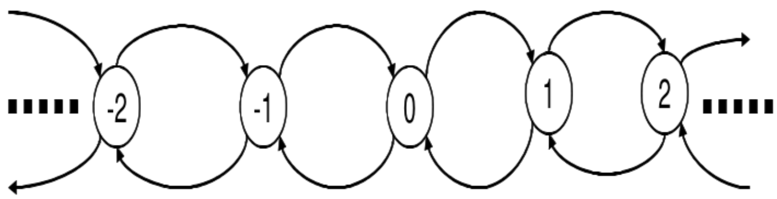

Fig. 1. Random walk

2.3. Convergence

Does π_t converge to a distribution? if the answer is yes, then does the convergent distribution depend on the initial distribution π_0?

If the state space is finite dimensional, we introduce the following theorem to answer the above questions.

Theorem 2.1. *Let P be a stochastic matrix for which $P_{ij}^t > 0$ for any (i,j) and some value of $t \geq 1$. Then P has a unique stationary distribution, and the process is asymptotically stationary, i.e. for all initial distribution π_0, $P^\infty \pi_0$ converges to the same π_∞. In this case, $\lim_{t\to\infty} P_{ij}^t = \pi_{\infty j}$ for all i,j.*

(Proof: see Perron–Frobenius theorem)

Example 2.4. *Regime shift (again)*

In this example

$$P^\infty = \frac{1}{2-p-q} \begin{bmatrix} 1-q & 1-p \\ 1-q & 1-p \end{bmatrix}$$

Note that if $p, q \neq 0$, the condition of the theorem is satisfied. We will have one unique stationary distribution. To find this stationary distribution, notice that a stationary distribution must satisfy

$$\pi' = \pi' P$$

That is to find π *satisfying*

$$(I - P')\pi = 0$$

Therefore, a stationary distribution is an eigenvector (normalized to satisfy $\sum_{j=1}^{n} \pi_j = 1$*) associated with a unit eigenvalue of* P'*. As*

$$P = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}$$

we have $\pi = \frac{1}{2-p-q}[1-q, 1-p]'$*. Therefore, it is true that* $P_{ij}^{\infty} = \pi_j$ *for all* i, j*.*

The fact that P is a stochastic matrix guarantees that P has at least one unit eigenvalue. This is because $P \times [1, 1, ..., 1]' = [1, 1, ..., 1]'$. So there is at least one such eigenvector π. However, this π may not be unique since P can have a repeated unit eigenvalues in general.

2.4. Characterization of stationary distribution

In practice, we often care about the first and second moment of stochastic variables. Below, we give a simple example.

Example 2.5. *Auto-correlation in the steady state*

There is a continuum of entrepreneurs. At each date some entrepreneurs are productive (productivity $x_t = \alpha$*); and the others are unproductive (*$x_t = \gamma \in (0, \alpha)$*). Each entrepreneur shifts stochastically between productive and unproductive states following a Markov process. The transition matrix is*

$$P = \begin{bmatrix} 1-\delta & \delta \\ n\delta & 1-n\delta \end{bmatrix}$$

The shifts of the productivity are exogenous and independent across entrepreneurs and over time.

- *Q1: Average productivity of entrepreneurs in the steady state*
  As shown before, the stationary distribution is $\pi = \frac{1}{n+1}[n, 1]'$*. Therefore, the average productivity is* $\frac{n}{n+1}\alpha + \frac{1}{n+1}\gamma$*.*
- *Q2: Under what condition such that the productivity of each entrepreneur is positively serially correlated in the steady state?*
  This is to calculate the autocorrelation function of x_s *(s represents a time index in the steady state). We start by calculating the standard deviation of* x_s*,* $\sigma(x_s)$*.*

$$\sigma(x_s) = \sqrt{\frac{n\alpha^2 + \gamma^2}{n+1} - \left(\frac{n\alpha + \gamma}{n+1}\right)^2} = \frac{\sqrt{n}(\alpha - \gamma)}{n+1}$$

According to the definition, $\text{corr}(x_s, x_{s+1}) = \text{Cov}(x_s, x_{s+1})/\sigma(x_s)/\sigma(x_{s+1})$*.*

$$\text{Cov}(x_s, x_{s+1}) = E(x_s x_{s+1}) - E(x_s)E(x_{s+1})$$

To calculate $E(x_s x_{s+1})$, we employ the fact that $E(x_s x_{s+1}) = E[E(x_s x_{s+1}|x_s)]$ (total expectation theorem). Therefore,

$$E(x_s x_{s+1}) = \frac{n}{n+1}\alpha[(1-\delta)\alpha + \delta\gamma] + \frac{1}{n+1}\gamma[(1-n\delta)\gamma + n\delta\alpha]$$

Note that in steady state, $E(x_{s+1}) = E(x_s)$,

$$\text{Cov}(x_s, x_{s+1}) = \frac{n(\alpha-\gamma)^2(1-\delta-n\delta)}{(n+1)^2}$$

Therefore, we have $\text{corr}(x_s, x_{s+1}) = 1 - \delta - n\delta$. (As expected, this should not depend on the value of α and γ) To have positively serially correlation in the steady stated, the probability of the productivity shifts is not too large: $\delta + n\delta < 1$.

3. Continuous state space

For discrete time, the time invariant Markov process can be described by the transition density

$$\mathcal{D}(x_{t+1}|x_t).$$

- Properties of the transition matrix:
  - $\mathcal{D}(x_{t+1}|x_t) > 0$ and $\int \mathcal{D}(x_{t+1}|x_t)dx_{t+1} = 1$
  - $\mathcal{D}(x_{t+2}|x_t) = \int dx_{t+1}\mathcal{D}(x_{t+2}|x_{t+1})\mathcal{D}(x_{t+1}|x_t)$
- Evolution of the system:

  The system at time t can be described by a probability density function $\pi_t(x_t)$. The initial density is given by $\pi_0(x_0)$. For all t, $\int dx_t\pi_t(x_t) = 1$. The evolution of the system is

  $$\pi_t(x_t) = \int dx_{t-1}\mathcal{D}(x_t|x_{t-1})\pi_{t-1}(x_{t-1})$$

  A stationary distribution satisfies

  $$\pi_\infty(s') = \int ds\mathcal{D}(s'|s)\pi_\infty(s)ds$$

3.1. Stochastic linear difference equations

Stochastic linear difference equations are an useful example of continuous-state Markov process. They are useful because (1) they are tractable; (2) they usually appear to represent an optimum or equilibrium outcome of agents' decision making; (3) they usually appear to represent the exogenous information flows impinging on an economy.

A state vector $x_t \in \mathbb{R}^n$ summarizes the information about the current position of a system. The initial distribution $\pi_0(x_0)$ is Gaussian with mean μ_0 and covariance matrix Σ_0. The transition density $D(x'|x)$ is Gaussian with mean Ax and covariance CC'. This specification can be represented in terms of the following equation

$$x_{t+1} = Ax_t + Cw_{t+1}, \tag{2}$$

where x_t is a $n \times 1$ state vector, A is a $n \times n$ matrix, C is a $n \times m$ matrix, and w_{t+1} is an i.i.d process satisfying $w_{t+1} \equiv \mathcal{N}(0, I)$.

3.1.1. Discussion of w_{t+1}

- In the continuous time Markov process, the Gaussian nature of w_{t+1} follows in the fact from the *continuity* of the process. Here, w represents the **Wiener process** and $dw(t)$ is Gaussian with mean 0 and variance dt.
- An important property of the Gaussian distribution: All moments of order higher than 2 can be expressible in terms of those of order 1 and 2.
- In practice, we usually focus on the first and second moments of x_t. The Gaussian assumption of w_{t+1} can be relaxed to the following

$$\mathbb{E}w_{t+1} = 0 \quad \text{for all } t$$

and

$$\mathbb{E}w_t w_{t-j}^T = \begin{cases} I, & \text{if } j = 0, \\ 0, & \text{if } j \neq 0. \end{cases}$$

A process satisfying the above two conditions is said to be a vector *white noise.* It is sufficient to justify the formulas that we report below for the second moments.

Example 3.1. *Vector auto regression*

Let z_t be an $n \times 1$ vector of random variables. We define a VAR of lag order 4 by a stochastic linear difference equation

$$z_{t+1} = \sum_{j=1}^{4} A_j z_{t+1-j} + C_y w_{t+1}.$$

We can represent the equation as follows

$$\begin{bmatrix} z_{t+1} \\ z_t \\ z_{t-1} \\ z_{t-2} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} z_t \\ z_{t-1} \\ z_{t-2} \\ z_{t-3} \end{bmatrix} + \begin{bmatrix} C_y \\ 0 \\ 0 \\ 0 \end{bmatrix} w_{t+1}$$

where A_j is an $n \times n$ matrix for each j. w_{t+1} is an $n \times 1$ white noise vector.

3.2. Characterization of stationary distribution

We care about the properties of x_t in the steady state. With Eq. 2, we are not working directly with the probability distribution of x_t as in the case of discrete state space. So what does stationarity means in this case? There are two important forms of stationarity:

- Strong stationarity
  The process x_t is strongly stationary if the joint distribution function of (x_{t_1+k}, x_{t_2+k}, ... ,x_{t_s+k}) is equal to the joint distribution function of (x_{t_1}, x_{t_2}, ..., x_{t_s}) for any finite set of time indices $\{t_1, t_2, \cdots, t_s\}$ and any integer k.
- Weak stationarity
  Weakly stationary can also be called covariance-stationary. It is when (1) the first moment of x_t is independent of t; (2) the second moment of x_t is finite for all t; (3) $\text{cov}(x_{t_1}, x_{t_2}) = \text{cov}(x_{t_1+h}, x_{t_2+h})$ for all t_1, t_2 and h.

We now proceed to find the first and second moments of x_t and then determine the condition under which x_t is covariance stationary. (In practice, the first and second moments are usually what people concern.) Taking expectation on both sides of Eq. 2,

$$\langle x_{t+1} \rangle = A \langle x_t, \rangle$$

where $\mathbb{E}(x_t) \equiv \langle x_t \rangle$. x_t possesses a stationary mean μ satisfying

$$(I - A)\mu = 0.$$

μ is an eigenvector associated with the single unit eigenvalue (this makes sure that μ is unique) of A. We will assume that all remaining eigenvalues of A are strictly less than 1 in modulus. Then the linear system $\langle x_{t+1} \rangle = A\langle x_t \rangle$ is asymptotically stable, implying that starting from any initial condition, $\langle x_t \rangle \to \mu$. This can be shown easily by representing A in Jordan form, $A = PJP^{-1}$. Here J is a Jordan matrix. Defining $\langle x_t \rangle = P\langle x_t \rangle^*$, we have $\langle x_{t+1} \rangle^* = J\langle x_t \rangle^*$. Therefore $\langle x_t \rangle^* = J^t \langle x_0 \rangle^*$. Because the eigenvalue structure of A, J^t converges to a matrix of zero except for the (1,1) element. As a result, $\langle x_t \rangle^*$ converges to a vector of zero except for the first element, which stays at $\langle x_0 \rangle_1^*$, and $\langle x_t \rangle$ converges to $\langle x_0 \rangle_1^* P_1$, where P_1 is the eigenvector corresponding to the unit eigenvalue.

For the second moments, we rewrite Eq. 2 as

$$x_{t+1} - \mu = A(x_t - \mu) + Cw_{t+1} \tag{3}$$

The covariance matrix is defined as $\Sigma_t = \mathbb{E}[(x_t - \mu)(x_t - \mu)^T]$. According to the above equation, its law of motion will be

$$\Sigma_{t+1} = A\Sigma_t A^T + CC^T$$

This is exactly the **discrete Lyapunov equation** in control theory. The following theorem tells us the existence and uniqueness of a stationary covariance matrix Σ_s.

Theorem 3.1. *If the linear system $\langle x_{t+1}\rangle = A\langle x_t\rangle$ is asymptotically stable, given any C, there exists a unique Σ_s satisfying*

$$\Sigma_s = A\Sigma_s A^T + CC^T$$

and Σ_s is positive definite.

The formal solution to Σ_s is given by

$$\Sigma_s = \sum_{j=0}^{\infty} A^j CC^T (A^T)^j \tag{4}$$

Numerically, it can be solved by a matlab subroutine `doublej.m`.

Iterating Eq. 3, we obtain

$$x_{t+j} - \mu = A^j(x_t - \mu) + Cw_{t+j} + ... + A^{j-1}Cw_{t+1}.$$

Therefore the autocovariance function of order j $\mathbb{E}[(x_{t+j} - \mu)(x_t - \mu)]$ is given by

$$\mathbb{E}[(x_{t+j} - \mu)(x_t - \mu)] = A^j \Sigma_s \tag{5}$$

If Σ_t achieves stationarity, so as the autocovariance function and it is given by $A^j\Sigma_s$.

Example 3.2. *AR(1) process*

In macroeconomics, the AR(1) process is most commonly used $y_t = \rho y_{t-1} + \sigma\varepsilon_t + b$, where b is a constant and ε_t are typically assumed to be uncorrelated $N(0,1)$ random variables.

$$\begin{bmatrix} 1 \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b & \rho \end{bmatrix} \begin{bmatrix} 1 \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma \end{bmatrix} \varepsilon_t \tag{6}$$

$A = \begin{bmatrix} 1 & 0 \\ b & \rho \end{bmatrix}$. *Its two eigenvalues are 1 and ρ. So the process will be weakly stationary if $|\rho| < 1$. The stationary mean is μ, and $A\mu = \mu$. So $\mu = [1, b/(1-\rho)]'$ and $E[y_t]$ in the steady state is $b/(1-\rho)$. This result can also be obtained by taking expectation on both sides of $y_t = \rho y_{t-1} + \sigma\varepsilon_t + b$ so that $E[y_t] = \rho E[y_{t-1}] + b$. Since $E[y_t]$ does not change in the steady state, stationary $E[y_t]$ is simply given by $b/(1-\rho)$.*

To find the stationary covariance $E[(y_t - E(y_t))(y_t - E(y_t))]$, we define $x_t = y_t - b/(1-\rho)$ so that $x_t = \rho x_{t-1} + \sigma\varepsilon_t$. According to Eq. 4, $E[(y_t - E(y_t))(y_t - E(y_t))] = E(x_t x_t) = \sum_{j=0}^{\infty} \rho^j \sigma\sigma\rho^j = \sigma^2/(1-\rho^2)$.

To find stationary covariance $E[(y_{t+k} - E(y_{t+k}))(y_t - E(y_t))]$, we use Eq. 5 and get $E[x_{t+k}x_t] = \rho^k E[(x_t x_t)]$. Therefore, $E[(y_{t+k} - E(y_{t+k}))(y_t - E(y_t))] = \rho^k\sigma^2/(1-\rho^2)$.

4. Ergodicity theory

Ergodic theory studies dynamical systems with an invariant distribution. It leads to a stunning result that under certain conditions, the time average of a function along the evolution path is equal to the space average. For stationary Markov process, we have the following theorem

Theorem 4.1. *If a time homogeneous Markov process has a unique stationary distribution, then it is ergodic.*

The property of ergodicity may give us a tremendous computational advantage to calculate moments of stochastic variables in the computer simulation.

For the theory of ergodicity, we refer the readers to *e.g.* chapter 4 in Miao, 2014 and references therein.

Appendix A. Basic concepts in stochastic processes

The modern formalism used by mathematicians to describe probability involves a number of concepts, predefined structures, and jargon. This modern formalism is *not* required to understand probability theory. Nevertheless, research work that is written in this modern language is not accessible unless you know the jargon. Unfortunately a considerable investment of effort is required to learn modern probability theory in its technical detail: significant groundwork is required to define the concepts with the precision demanded by mathematicians. Here we present the concepts and jargon of modern probability theory without the rigorous mathematical technicalities. There are many good textbooks that present the details of the modern formalism; we recommend, for example, the excellent and concise text by Williams, 1991.

Definition A.1. *Probability space*

A probability space is a mathematical object consisting of three elements:

- Ω*, sample space of possible outcome* ω
- $\mathcal{F}$*, The collection of events (Each possible subset of the samples space is called an event) is called the* σ-algebra*.*
- $\mathbb{P}$*, a measure that assigns probability values to those events. Associating a number (in our case the probability) with every element of a* σ*-algebra is called a* measure*. (It is from integration that the notion of a measure first sprung.)* $\int_\Omega \mathbb{P}(d\omega) = 1$

A.1. σ-algebra

Algebra, is a collection of objects, along with a number of mathematical operations that can be performed on the objects. To be an 'algebra' the collection must be closed under the operations. Operations can take one object, in which case they produce a second object, and are referred to as unary; Operations may also take two objects, in which case they produce a third object, and are referred to as binary.

The collection of events could form an algebra. There could be three operations associated with it:

- union (binary operation): $A \cup B$
- intersection: $A \cap B$
- complement (unary operation): $A^{\complement}$. The whole sample space is $\Omega = A \cup A^{\complement}$.

A σ-algebra ($\mathcal{F}$) is special kinds of family of events that satisfy three properties:

(1) $\Omega \in \mathcal{F}$,

(2) $\mathcal{F}$ is closed under complementation, $A \in \mathcal{F} \rightarrow A^{\complement} \in \mathcal{F}$,

(3) $\mathcal{F}$ is closed under countable union: if $\{A_i\}_{i=1}^{\infty}$ is a sequence of sets such that $A_i \in \mathcal{F}$, then $(\cup_{i=1}^{\infty} A_i) \in \mathcal{F}$.

Example A.1. *Borel σ-algebra*

Consider a real variable x, where x can take any value on the real line. The set of all values of x is the sample space, Ω. All the open intervals on the real line a σ-algebra. (or in N dimension, all the open N-dimensional cubes in $\mathbb{R}^N$ forms a σ-algebra.)

Definition A.2. *Random variable*

A random variable is a function whose domain is the set of events Ω and whose image is the real numbers. It is actually a function from the sample space, to a set of values.

X is a random variable that takes values on the real line, and is defined by the function $f(\omega)$. The probability that $-\infty < X < x$ (recall that this is the probability distribution for x) is obtained by

$$D(x) \equiv \Pr(X \in (-\infty, x]) = \mathbb{P}(f^{-1}(-\infty, x]))$$

if for any x, $f^{-1}(-\infty, x) \in \mathcal{F}$. We call X is $\mathcal{F}$ measurable. If X is $\mathcal{F}$-measurable, we can sensibly talk about the probability of X taking values in virtually any subset of the real line you can think of (the Borel sets). That is f^{-1} can map any Borel set to an element in $\mathcal{F}$.

Definition A.3. *Stochastic processes X_t*

A sequence of σ-algebras that describes what information we have access to at successive times is called a filtration

$$\{\mathcal{F}_t\}_{t=1}^{\infty} : \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 ... \subseteq \mathcal{F}.$$

Let a sequence of random variables x_t be $\mathcal{F}_t$-measurable for each t, which models a stochastic process. For example, consider an $\omega \in \Omega$, and choose an $\alpha \in \mathbb{R}$. Then for each t, the set $\{\omega : x_t(\omega) < \alpha\}$ will be a set included in the collection $\mathcal{F}_t$. Since $\mathcal{F}_t \subseteq \mathcal{F}$ for all t, the set also belongs to $\mathcal{F}$. Hence, we can assign probability to the set using the measure $\mathbb{P}$. This collection of objects is called a filtered probability space, or just filtered space, and usually written as $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$.

References

Miao, J., 2014. Economic Dynamics in Discrete Time. The MIT Press.

Williams, D., 1991. Probability with martingales. Cambridge University Press.