Università di Pavia

Introduction to Stochastic processes

Eduardo Rossi

Stochastic Process: A stochastic process is an ordered sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$.

$\{Y_t(\omega), \omega \in \Omega, t \in \mathcal{T}\}$, such that for each $t \in \mathcal{T}$, $y_t(\omega)$ is a random variable on the sample space Ω, and for each $\omega \in \Omega$, $y_t(\omega)$ is a realization of the stochastic process on the index set $\mathcal{T}$ (that is an ordered set of values, each corresponds to a value of the index set).

Time Series: A time series is a set of observations $\{y_t, t \in \mathcal{T}_0\}$, each one recorded at a specified time t. The time series is a part of a realization of a stochastic process, $\{Y_t, t \in \mathcal{T}\}$ where $\mathcal{T} \supseteq \mathcal{T}_0$. An infinite series $\{y_t\}_{t=-\infty}^{\infty}$

The *unconditional mean*

$$\mu_t = E[Y_t] = \int Y_t f(Y_t) dY_t \tag{1}$$

Autocovariance function: The joint distribution of

$$(Y_t, Y_{t-1}, \ldots, Y_{t-h})$$

is usually characterized by the autocovariance function:

$$\begin{aligned}
\gamma_t(h) &= Cov(Y_t, Y_{t-h}) \\
&= E[(Y_t - \mu_t)(Y_{t-h} - \mu_{t-h})] \\
&= \int \ldots \int (Y_t - \mu_t)(Y_{t-h} - \mu_{t-h}) f(Y_t, \ldots, Y_{t-h}) dY_t \ldots dY_{t-h}
\end{aligned}$$

The *autocorrelation function*

$$\rho_t(h) = \frac{\gamma_t(h)}{\sqrt{\gamma_t(0)\gamma_{t-h}(0)}}$$

Weak (Covariance) Stationarity: The process Y_t is said to be *weakly stationary* or *covariance stationary* if the second moments of the process are time invariant:

$$
\begin{aligned}
E[Y_t] &= \mu < \infty \;\; \forall t \\
E[(Y_t - \mu)(Y_{t-h} - \mu)] &= \gamma(h) < \infty \;\; \forall t, h
\end{aligned}
$$

Stationarity implies $\gamma_t(h) = \gamma_t(-h) = \gamma(h)$.

Strict Stationarity The process is said to be strictly stationary if for any values of $h_1, h_2, \ldots, h_n$ the joint distribution of $(Y_t, Y_{t+h_1}, \ldots, Y_{t+h_n})$ depends only on the intervals $h_1, h_2, \ldots, h_n$ but not on the date t itself:

$$f(Y_t, Y_{t+h_1}, \ldots, Y_{t+h_n}) = f(Y_\tau, Y_{\tau+h_1}, \ldots, Y_{\tau+h_n}) \ \forall t, h \qquad (2)$$

Strict stationarity implies that all existing moments are time invariant.

Gaussian Process The process Y_t is said to be Gaussian if the joint density of $(Y_t, Y_{t+h_1}, \ldots, Y_{t+h_n})$, $f(Y_t, Y_{t+h_1}, \ldots, Y_{t+h_n})$, is Gaussian for any $h_1, h_2, \ldots, h_n$.

The statistical ergodicity theorem concerns what information can be derived from an average over time about the common average at each point of time.

Note that the WLLN does not apply as the observed time series represents just one realization of the stochastic process.

Ergodic for the mean. Let $\{Y_t(\omega), \omega \in \Omega, t \in \mathcal{T}\}$ be a weakly stationary process, such that $E[Y_t(\omega)] = \mu < \infty$ and $E[(Y_t(\omega) - \mu)^2] = \sigma^2 < \infty \;\; \forall t$. Let $\overline{y}_T = T^{-1}\sum_{t=1}^{T} Y_t$ be the time average. If $\overline{y}_T$ converges in probability to μ as $T \to \infty$, Y_t is said to be ergodic for the mean.

To be ergodic the memory of a stochastic process should fade in the sense that the covariance between increasingly distant observations converges to zero sufficiently rapidly.

For stationary process it can be shown that *absolutely summable autocovariances*, i.e. $\sum_{h=0}^{\infty} |\gamma(h)| < \infty$, are sufficient to ensure ergodicity.

Ergodic for the second moments

$$\widehat{\gamma}(h) = (T-h)^{-1} \sum_{t=h+1}^{T} (Y_t - \mu)(Y_{t-h} - \mu) \xrightarrow{P} \gamma(h) \tag{3}$$

Ergodicity focus on *asymptotic independence*, while stationarity on the *time-invariance* of the process.

Consider the stochastic process $\{Y_t\}$ defined by

$$Y_t = \begin{cases} u_0 & t = 0 \;\; with \;\; u_0 \sim N(0, \sigma^2) \\ Y_{t-1} & t > 0 \end{cases} \tag{4}$$

Then $\{Y_t\}$ is strictly stationary but not ergodic.

Proof Obviously we have that $Y_t = u_0$ for all $t \geq 0$. Stationarity follows from:

$$\begin{aligned} E[Y_t] &= E[u_0] = 0 \\ E[Y_t^2] &= E[u_0^2] = \sigma^2 \\ E[Y_t Y_{t-1}] &= E[u_0^2] = \sigma^2 \end{aligned} \tag{5}$$

Thus we have $\mu = 0$, $\gamma(h) = \sigma^2$ $\rho(h) = 1$ are time invariant.
Ergodicity for the mean requires:

$$\begin{aligned} \overline{y}_T &= T^{-1} \sum_{t=o}^{T-1} Y_t = T^{-1}(T u_0) \\ \overline{y}_T &= u_0 \end{aligned}$$

$$Y_t = Y_{t-1} + u_t$$

where $u_t \sim WN(0, \sigma^2)$. By recursive substitution,

$$\begin{aligned} Y_t &= Y_{t-1} + u_t \\ Y_t &= Y_{t-2} + u_{t-1} + u_t \\ &\vdots \\ &= Y_0 + u_t + u_{t-1} + u_{t-2} + \ldots + u_1 \end{aligned}$$

The mean is time-invariant:

$$\mu = E[Y_t] = E\left[Y_0 + \sum_{s=1}^{t} u_s\right] = Y_0 + \sum_{s=1}^{t} E[u_s] = 0. \tag{6}$$

But the second moments are diverging. The variance is given by:

$$
\begin{aligned}
\gamma_t(0) &= E[Y_t^2] = E\left[\left(Y_0 + \sum_{s=1}^{t} u_s\right)^2\right] = E\left[\left(\sum_{s=1}^{t} u_s\right)^2\right] \\
&= E\left[\sum_{s=1}^{t}\sum_{k=1}^{t} u_s u_k\right] = E\left[\sum_{s=1}^{t} u_s^2 + \sum_{s=1}^{t}\sum_{k=1}^{t} u_s u_k\right] \\
&= \sum_{s=1}^{t} E[u_s^2] + \sum_{s=1}^{t}\sum_{k=1,k\neq s}^{t} E[u_s u_k] = \sum_{s=1}^{t} \sigma^2 = t\sigma^2.
\end{aligned}
$$

The autocovariances are:

$$\begin{aligned}\gamma_t(h) &= E[Y_t Y_{t-h}] = E\left[\left(Y_0 + \sum_{s=1}^{t} u_s\right)\left(Y_0 + \sum_{k=1}^{t-h} u_k\right)\right] \\ &= E\left[\sum_{s=1}^{t} u_s \left(\sum_{k=1}^{t-h} u_k\right)\right] \\ &= \sum_{k=1}^{t-h} E[u_k^2] \\ &= \sum_{k=1}^{t-h} \sigma^2 = (t-h)\sigma^2 \quad \forall h > 0. \qquad (7)\end{aligned}$$

Finally, the autocorrelation function $\rho_t(h)$ for $h > 0$ is given by:

$$\rho_t^2(h) = \frac{\gamma_t^2(h)}{\gamma_t(0)\gamma_{t-h}(0)} = \frac{[(t-h)\sigma^2]^2}{[t\sigma^2][(t-h)\sigma^2]} = 1 - \frac{h}{t} \quad \forall h > 0 \qquad (8)$$

A white-noise process is a weakly stationary process which has zero mean and is uncorrelated over time:

$$u_t \sim WN(0, \sigma^2) \tag{9}$$

Thus u_t is WN process $\forall t \in \mathcal{T}$:

$$\begin{aligned} E[u_t] &= 0 \\ E[u_t^2] &= \sigma^2 < \infty \\ E[u_t u_{t-h}] &= 0 \;\; h \neq 0, t-h \in \mathcal{T} \end{aligned} \tag{10}$$

If the assumption of a constant variance is relaxed to $E[u_t^2] < \infty$, sometimes u_t is called a weak WN process.

If the white-noise process is normally distributed it is called a *Gaussian white-noise process*:

$$u_t \sim NID(0, \sigma^2). \tag{11}$$

The assumption of normality implies strict stationarity and serial independence (unpredictability). A generalization of the NID is the IID process with constant, but unspecified higher moments.

A process u_t with independent, identically distributed variates is denoted IID:

$$u_t \sim IID(0, \sigma^2) \tag{12}$$

Martingale The stochastic process x_t is said to be *martingale* with respect to an information set (σ-field), $\mathcal{I}_{t-1}$, of data realized by time $t-1$ if

$$\begin{aligned} E[|x_t|] &< \infty \\ E[x_t|\mathcal{I}_{t-1}] &= x_{t-1} \text{ a.s.} \end{aligned} \tag{13}$$

Martingale Difference Sequence The process $u_t = x_t - x_{t-1}$ with $E[|u_t|] < \infty$ and $E[u_t|\mathcal{I}_{t-1}] = 0$ for all t is called a *martingale difference sequence*, MDS.

Let $\{x_t\}$ be an m.d. sequence and let $g_{t-1} = \sigma(x_{t-1}, x_{t-2}, \ldots)$ be any measurable, integrable function of the lagged values of the sequence. Then $x_t g_{t-1}$ is a m.d., and

$$Cov[g_{t-1}, x_t] = E[g_{t-1}x_t] - E[g_{t-1}]E[x_t] = 0$$

This implies that, putting $g_{t-1} = x_{t-j} \;\; \forall j > 0$

$$Cov[x_t, x_{t-j}] = 0$$

M.d. property implies uncorrelatedness of the sequence.

White Noise processes may not be m.d. because the conditional expectation of an uncorrelated process can be a nonlinear.

An example is the bilinear model (Granger and Newbold, 1986)

$$x_t = \beta x_{t-2}\epsilon_{t-1} + \epsilon_t, \;\; \epsilon_t \sim i.i.d(0, 1)$$

The model is w.n. when $0 < \beta < \frac{1}{\sqrt{2}}$.

The conditional expectations are the following nonlinear functions:

$$E[x_t|\mathcal{I}_{t-1}] = -\sum_{i=1}^{\infty}(-\beta)^{i+1}x_{t-i}\left(\prod_{j=2}^{i+1}x_{t-j}\right)$$

Uncorrelated, zero mean processes:

- White Noise Processes
  1. IID processes
     (a) Gaussian White Noise processes
- Martingale Difference Processes

Innovation An innovation $\{u_t\}$ against an information set $\mathcal{I}_{t-1}$ is a process whose density $f(u_t|\mathcal{I}_{t-1})$ does not depend on $\mathcal{I}_{t-1}$.

Mean Innovation $\{u_t\}$ is a mean innovation with respect to an information set $\mathcal{I}_{t-1}$ if $E[u_t|\mathcal{I}_{t-1}] = 0$

If the zero-mean process Y_t is wide sense stationary (implying $E(Y_t^2) < \infty$) it has the representation

$$Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} + v_t$$

where $\psi_0 = 1, \sum_{j=0}^{\infty} \psi_j^2 < \infty$

$$\epsilon_t \sim WN(0, \sigma^2)$$

$$E(v_t \epsilon_{t-j}) = 0, \forall j$$

and there exist constants $\alpha_0, \alpha_1, \alpha_2, \ldots,$ such that $Var(\sum_{j=0}^{\infty} \alpha_j v_t) = 0$.

The distribution of v_t is singular, since

$$v_t = -\sum_{j=1}^{\infty} (\alpha_j/\alpha_0) v_{t-j}$$

with probability 1, and hence is perfectly predictable one-step ahead.
(*Deterministic process*).
If $v_t = 0$, Y_t is called a purely *non-deterministic process*.

Approximate the infinite lag polynomial with the ratio of two finite-order polynomials $\phi(L)$ and $\theta(L)$:

$$\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j \cong \frac{\theta(L)}{\phi(L)} = \frac{1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q}{1 - \phi_1 L - \phi_2 L^2 \ldots - \phi_p L^p}$$

Time Series Models

p	q	**Model**	**Type**
$p > 0$	$q = 0$	$\phi(L)Y_t = \epsilon_t$	AR(P)
$p = 0$	$q > 0$	$Y_t = \theta(L)\epsilon_t$	MA(Q)
$p > 0$	$q > 0$	$\phi(L)Y_t = \theta(L)\epsilon_t$	ARMA(p,q)