Основные классы случайных процессов

А.Г. Трофимов

к.т.н., доцент, НИЯУ МИФИ

lab@neuroinfo.ru
http://datalearning.ru

Курс "Статистическая обработка временных рядов"

Сентябрь 2018

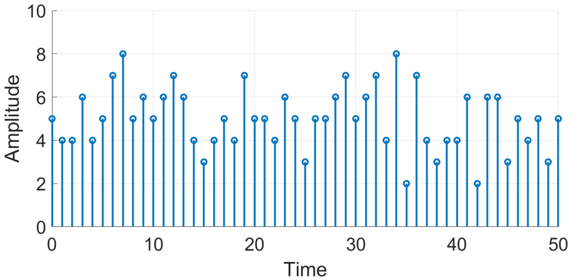
IID Process

IID Process

Independent and identically distributed (IID) process X is a process where all random variables $\{X(t), t \in \mathscr{T}\}$ are IID

A IID process can have any probability density function (e.g. Gaussian, Binomial, Poisson, etc.)

A sample function of Binomial $B(10, 0.5)$ IID time series



Properties of IID Processes

- IID processes are SSS
  IID process is k-order stationary for all $k = 1, 2, \ldots$:
  $$f(x,t) = f(x,t+\tau) = f(x), \quad \forall \tau, t \in \mathscr{T}$$
  $$f(x_1,x_2;t_1+\tau,t_2+\tau) = f(x_1,t_1+\tau)f(x_2,t_2+\tau)$$
  $$= f(x_1,t_1)f(x_2,t_2) = f(x_1,x_2;t_1,t_2), \quad \forall \tau, t_1, t_2 \in \mathscr{T}$$
  By induction, for all other orders of stationarity
- IID processes are ergodic
  The realization $x(1), x(2), \ldots$ of time series X can be viewed as independent sample of observations of the random variable with PDF $f(x)$. Therefore, for IID time series the time averaging is equivalent to the ensemble averaging

Bernoulli Process

Definition

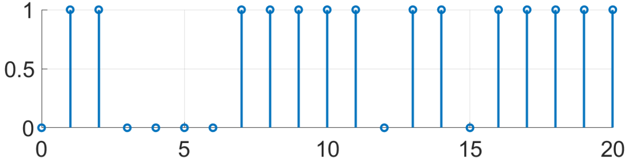
Bernoulli process is a IID process $\{X(t), t = 0, 1, ...\}$, where all

$$X(t) \sim B(1, p)$$

Bernoulli process is a discrete-time random process that takes only two values 0 and 1

Bernoulli process is a repeated coin flipping, possibly with an unfair coin (but with consistent unfairness)

A sample function of time series X ($p = 0.5$)



Discrete-time White Noise

Definition

Strongly (strict-sense) discrete-time white noise $\{X(t), t = 0, 1, ...\}$ is a discrete-time IID process with zero mean and finite variance σ^2

Definition

Weakly (wide-sense) discrete-time white noise X is a discrete-time process that satisfies conditions:

$$\mathrm{M}[X(t)] = 0, \quad \forall t = 0, 1, ...$$

$$\mathrm{D}[X(t)] = \sigma^2 < \infty, \quad \forall t = 0, 1, ...$$

$$cov(X(t), X(s)) = 0, \quad \forall t, s = 0, 1, ..., \quad t \neq s$$

Wide-sense whiteness skips the requirement of identity and relaxes the requirement of independency to the uncorrelatedness

Continuous-time White Noise

Definition

Strongly (strict-sense) continuous-time white noise $\{X(t), t \in (-\infty, \infty)\}$ is a continuous-time IID process with zero mean and infinite variance

Autocovariance of continuous-time white noise:

$$c(\tau) = \sigma^2 \delta(\tau), \quad \tau \in (-\infty, \infty)$$

where $\delta(\tau)$ is Dirac delta function:

$$\delta(\tau) = \begin{cases} +\infty, & \tau = 0, \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad \int_{-\infty}^{\infty} \delta(\tau) d\tau = 1$$

Infinite variance is related to constant spectral density of process X

A white noise is a mathematical abstraction, it cannot be physically realized since it has infinite variance

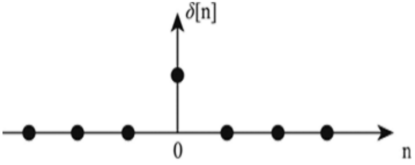
Autocovariance of White Noise

ACF of discrete-time white noise:

$$c(\tau) = \sigma^2 \delta[\tau], \quad \tau = 0, 1, \ldots$$

Discrete-time unit impulse (Kronecker delta)

$$\delta[\tau] = \begin{cases} 1, & \tau = 0, \\ 0, & \text{otherwise} \end{cases}$$

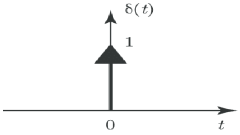


ACF of continuous-time white noise:

$$c(\tau) = \sigma^2 \delta(\tau), \quad \tau \in (-\infty, \infty)$$

Continuous-time unit impulse (Dirac delta function)

$$\delta(\tau) = \begin{cases} +\infty, & \tau = 0, \\ 0, & \text{otherwise} \end{cases}$$



White Noise. Notes (1)

- X is strongly white process $\Rightarrow$ X is weakly white process. The converse is not necessarily true
- Strongly white noise is a SSS process
- By definition, weakly white noise is a WSS process
  Weakly white noise is a serially uncorrelated, zero-mean and constant variance process
- The random variables $X(t)$ and $X(s)$ are independent for strongly white noise (or uncorrelated for weakly white noise) no matter how closely they are placed
- Sometimes the zero-mean requirement is relaxed to constant-mean requirement. In this case, white noise with zero mean is referred to as centered white noise
- Variance σ^2 is called intensity of white noise

White Noise. Notes (2)

- A weakly white noise process is linearly unpredictable as the noise samples at different instants of time are uncorrelated
- A strongly white process is unpredictable and sometimes is referred to as purely random process
- A white noise process can have any probability density function (e.g. Gaussian, Binomial, Poisson, etc.)
  If X is a white noise and $X(t) \sim N(0, \sigma^2)$, $t \in \mathscr{T}$, then X is called white Gaussian noise
- For Gaussian random variables uncorrelatedness and independence are equivalent. Therefore, for white Gaussian noise strict-sense whiteness and wide-sense whiteness are equivalent

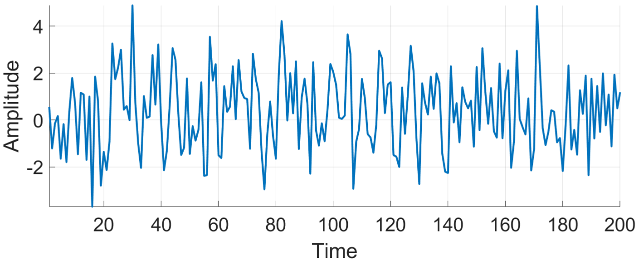
Example 1. Discrete-Time White Gaussian Noise (WGN)

For white Gaussian noise strict-sense whiteness and wide-sense whiteness are equivalent

All $\{X(t), t = 0, 1, ...\}$ are normally distributed IID random variables:

$$X(t) \sim N(0, \sigma^2)$$

A sample function of discrete-time WGN

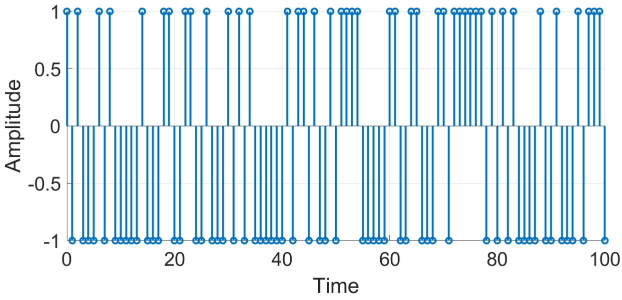


Example 2. White Bipolar Noise

All $\{X(t), t = 0, 1, ...\}$ are IID random variables with distribution:

$$X(t) = \begin{cases} -1, & \text{with probability } 1-p, \\ 1, & \text{with probability } p \end{cases}$$

A sample function of white bipolar noise ($p = 0.5$)

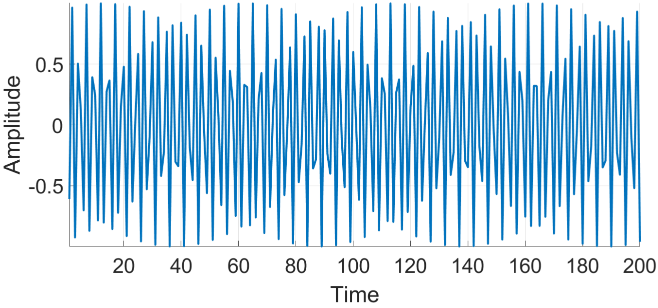


White bipolar noise is strongly white noise

Example 3. White Sinusoidal Noise

Consider a time series $X(t) = \sin(Vt)$, where V is a random variable, $V \sim R(0, 2\pi)$ and $t = 1, 2, ...$

A sample function of time series X



For each realization $\{x(t), t = 1, 2, ...\}$ the frequency v is constant over time

White Sinusoidal Noise. Expectation and Variance

Check if X is a weakly white time series

Expectation of time series X:

$$m(t) = \mathrm{M}[X(t)] = \mathrm{M}[\sin(Vt)] = \frac{1}{2\pi}\int\limits_0^{2\pi} \sin(vt)dv = \frac{-1}{2\pi t}\cos(vt)\Big|_0^{2\pi} = 0$$

Variance of time series X:

$$\sigma^2(t) = \mathrm{D}[X(t)] = \mathrm{D}[\sin(Vt)] = \frac{1}{2\pi}\int\limits_0^{2\pi} \sin^2(vt)dv$$

$$= \frac{1}{4\pi}\int\limits_0^{2\pi}(1-\cos(2vt))dv = \frac{1}{4\pi}\left(v - \frac{\sin(2vt)}{2t}\right)\Bigg|_0^{2\pi}$$

$$= \frac{1}{4\pi}(2\pi - 0) = \frac{1}{2}$$

White Sinusoidal Noise. Autocovariance

Autocovariance of time series X:

$$c(t, t+\tau) = \mathrm{M}[X(t)X(t+\tau)] = \frac{1}{2\pi}\int\limits_0^{2\pi} \sin(vt)\sin(vt+v\tau)dv$$

$$= \frac{1}{4\pi}\int\limits_0^{2\pi} (\cos(v\tau) - \cos(2vt+v\tau))dv$$

$$= \frac{1}{4\pi}\left(\frac{\sin(v\tau)}{\tau} - \frac{\sin(2vt+v\tau)}{2t+\tau}\right)\Bigg|_{v=0}^{v=2\pi} = 0, \quad \tau \neq 0$$

$m = 0$, $\sigma^2 = \frac{1}{2} < \infty$ and $c(\tau) = 0$ for all $\tau \neq 0$

$\Rightarrow$ X is weakly white noise

White Sinusoidal Noise. Stationarity

Check if X is a strongly white time series

For strongly white time series all $X(1), X(2), \ldots$ must be IID random variables

Consider random variables $X(1)$ and $X(2)$

Univariate distributions:

$$F(x_1, 1) = P(X(1) < x_1) = P(\sin V < x_1)$$

$$F(x_2, 2) = P(X(2) < x_2) = P(\sin 2V < x_2)$$

Bivariate distribution:

$$\begin{aligned} F(x_1, x_2; 1, 2) &= P(X(1) < x_1 \& X(2) < x_2) \\ &= P(\sin V < x_1 \& \sin 2V < x_2) \end{aligned}$$

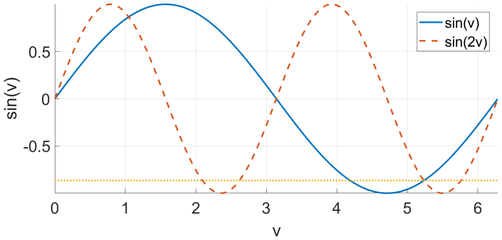
White Sinusoidal Noise. Stationarity

Assume $x_1 = x_2 = -\frac{\sqrt{3}}{2}$:

$$F(x_1, 1) = P\left(\sin V < -\sqrt{3}/2\right) > 0$$

$$F(x_2, 2) = P\left(\sin 2V < -\sqrt{3}/2\right) > 0$$

$$F(x_1, x_2; 1, 2) = P\left(\sin V < -\sqrt{3}/2 \;\&\; \sin 2V < -\sqrt{3}/2\right) = 0$$



$X(1)$ and $X(2)$ are dependent $\Rightarrow$ X is not strongly white noise

Gaussian Process

Gaussian Process

A random process X is called Gaussian process if all its finite order distributions are Gaussian, i.e. the random vector $(X(t_1), ..., X(t_k))$ has multivariate normal distribution

$$(X(t_1), ..., X(t_k)) \sim N_k(m, \Sigma)$$

for all k and all $t_1, ..., t_k \in \mathscr{T}$

The PDF of multivariate normally distributed vector $(X_1, ...X_k)$:

$$f_{X_1 ... X_k}(x) = \frac{1}{\sqrt{(2\pi)^k \det \Sigma}} \exp\left(-\frac{1}{2}(x-m)^T \Sigma^{-1}(x-m)\right)$$

where $m = (m_1, ..., m_k)^T$ is expectation, Σ is covariance matrix of vector $(X_1, ...X_k)$

Properties of Gaussian Processes

- Gaussian random processes generalize Gaussian random vectors to infinite dimensions
- Gaussian process $\{X(t), t \in \mathscr{T}\}$ can be completely defined by its second-order statistics: expectation $m(t)$, $t \in \mathscr{T}$, and autocovariance function $c(t_1, t_2)$, $t_1, t_2 \in \mathscr{T}$
- For Gaussian processes WSS and SSS properties are equivalent:

$$\text{WSS} \Leftrightarrow \text{SSS}$$

- Any linear transformations of Gaussian time series X:

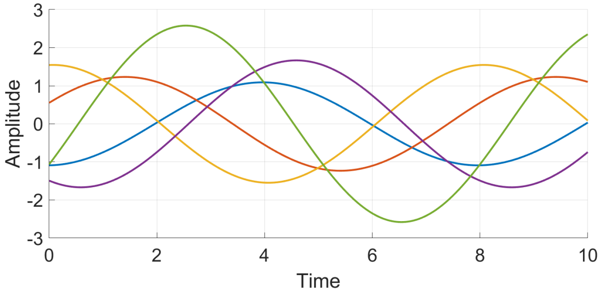
$$Y(t) = \sum_{i=-k}^{k} h_i X(t+i), \quad t = ..., -1, 0, 1, ...$$

results to a Gaussian time series Y

Gaussian Processes. Example 1

Process $X(t) = A\cos\left(\frac{\pi t}{4}\right) + B\sin\left(\frac{\pi t}{4}\right)$, where A and B are independent random variables, $A \sim N(0,1)$, $B \sim N(0,1)$ and $t \in \mathscr{T}$

Some realizations of time series X



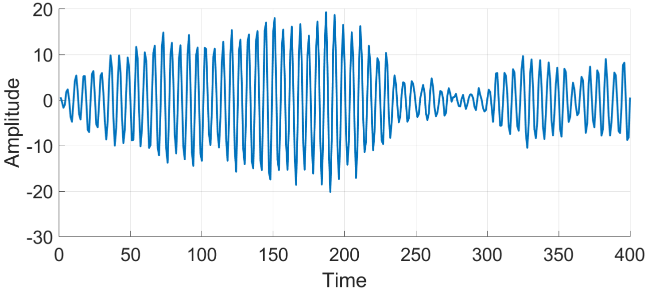
For each time moment $X(t) \sim N(0,1)$, $t \in \mathscr{T}$

Prove that time series X is stationary if $\mathscr{T} = \{0, 1, 2, ...\}$!

Gaussian Processes. Example 2

Time series $X(t) = X(t-1) - X(t-2) + U(t)$, $X(1) = U(1)$, $X(2) = U(2)$, where U is a Gaussian white noise $N(0,1)$ and $t = 1, 2, \dots$

A sample function of time series X

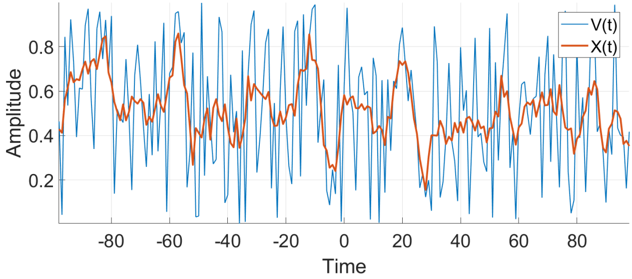


Prove that time series X is non-stationary!

Gaussian Processes. Example 3

Time series $X(t) = \frac{1}{5}\sum_{i=-2}^{2} V(t+i)$, where $\{V(t), t=-1,0,1,...\}$ is a uniform white noise $R(-1,1)$ and $t = 1,2,...$

A sample function of time series V and X



Prove that time series X is stationary!
Is time series X a Gaussian white noise?

Normalization of Non-Gaussian Time Series

Central limit theorem (CLT) establishes that, in some situations, when independent random variables are added, their properly normalized sum tends toward a normal distribution even if the original variables themselves are not normally distributed

A time series Y constructed as linear transformation of non-Gaussian time series X tends to be Gaussian:

$$Y(t) = \sum_{i=-k}^{k} h_i X(t+i)$$

Particularly, a time series Y constructed as moving average of non-Gaussian time series X tends to be Gaussian:

$$Y(t) = \frac{1}{2k+1} \sum_{i=-k}^{k} X(t+i)$$

Process with Independent Increments

Definition

A process X is said to be a process with independent increments if for any $k > 1$ and for any time moments $t_1 < t_2 < ... < t_k \in \mathscr{T}$ the random variables

$$X(t_1), X(t_2) - X(t_1), ..., X(t_k) - X(t_{k-1})$$

are mutually independent

Definition

A process X is said to be a process with stationary increments if for any $\tau \in \mathscr{T}$ and for any time moments $t_1 \leq t_2$, $t_1, t_2 \in \mathscr{T}$, the increments

$$X(t_2) - X(t_1) \quad \text{and} \quad X(t_2 + \tau) - X(t_1 + \tau)$$

have the same distributions

Process with Stationary Increments

- The distribution of the difference $X(t_2) - X(t_1)$ depends only on the length of the interval $(t_1, t_2]$ and not on the exact location of the interval on the time line
- Usually, it's assumed $X(0) = 0$
- Let $\tau = -t_1$:

$$\begin{aligned} X(t_2) - X(t_1) &\sim X(t_2 + \tau) - X(t_1 + \tau) \\ X(t_2) - X(t_1) &\sim X(t_2 - t_1) - X(0) \\ X(t_2) - X(t_1) &\sim X(t_2 - t_1) \end{aligned}$$

Definition

A process X is said to be a process with stationary increments if for any time moments $t_1 \leq t_2$, $t_1, t_2 \in \mathscr{T}$, the increment $X(t_2) - X(t_1)$ has the same distribution as the random variable $X(t_2 - t_1)$

SII Time Series

If the increments of the process X are stationary and independent then the process X is referred to as process with stationary, independent increments (SII-process)

Theorem (Criterion of SII time series)

The time series $\{X(t), t = 0, 1, ...\}$ has stationary independent increments iff there exists an IID time series $\{U(t), t = 1, 2, ...\}$ such that

$$X(t) = \sum_{i=1}^{t} U(i), \quad t = 1, 2, ...$$

For time series with stationary independent increments:

$$X(0) = 0$$

$$X(t) = X(t-1) + U(t), \quad t = 1, 2, ...$$

Criterion of SII Time Series. Proof

Sufficient condition:

Suppose that $\{U(t), t = 1, 2, ...\}$ is an IID time series and $X(t) = \sum\limits_{i=1}^{t} U(i)$, $t = 1, 2, ...$ Then

$$X(t_2) - X(t_1) = \sum_{i=t_1+1}^{t_2} U(i) = \sum_{i=1}^{t_2-t_1} U(i) = X(t_2 - t_1)$$

Therefore, X is SII time series

Necessary condition:

Suppose that $\{X(t), t = 0, 1, ...\}$ is process with stationary independent increments. Let

$$U(t) = X(t) - X(t-1), \quad t = 1, 2, ...$$

Therefore, $\{U(t), t = 1, 2, ...\}$ is a sequence of IID random variables and $X(t) = U(t) + X(t-1) = ... = \sum_{i=1}^{t} U(i)$

Expectation and Covariance of SII Time Series

Theorem

Suppose that $\{X(t), t = 0, 1, ...\}$ is a time series with stationary independent increments. Then there exist constants m and σ^2 such that the expectation and covariance function of X are

$$\mathrm{M}[X(t)] = mt, \quad t = 0, 1, ...$$

$$c(t, s) = cov(X(t), X(s)) = \sigma^2 \min\{t, s\}, \quad t, s = 0, 1, ...$$

If $t < s$:

$$c(t, s) = cov(X(t), X(s)) = \sigma^2 t, \quad t = 0, 1, ...$$

Variance of SII time series:

$$\mathrm{D}[X(t)] = cov(X(t), X(t)) = \sigma^2 t, \quad t = 0, 1, ...$$

Expectation and Covariance of SII Time Series. Proof

Time series X is SII $\Rightarrow$ exists an IID time series $\{U(t), t = 1, 2, ...\}$ such that

$$X(t) = \sum_{i=1}^{t} U(i), \quad t = 1, 2, ...$$

$$\mathrm{M}[X(t)] = \mathrm{M}\left[\sum_{i=1}^{t} U(i)\right] = \sum_{i=1}^{t} \mathrm{M}[U(i)] = mt$$

$$\begin{aligned} cov(X(t), X(s)) &= cov(X(t), X(t) + (X(s) - X(t))) \\ &= cov(X(t), X(t)) + cov(X(t), X(s) - X(t)) = \mathrm{D}[X(t)] \\ &= \mathrm{D}\left[\sum_{i=1}^{t} U(i)\right] = \sum_{i=1}^{t} \mathrm{D}[U(i)] = \sigma^2 t, \quad t < s \end{aligned}$$

Therefore, m and σ^2 are expectation and variance of random variables $U(t)$, $t = 1, 2, ...$

Counting Process

Definition

Counting process $\{X(t), t \in [0, \infty)\}$ is a random process with non-negative, integer and non-decreasing values:

$$1)\ X(t) \in \{0, 1, 2, ...\}, \quad t \in [0, \infty)$$
$$2)\ X(s) \leq X(t) \quad \text{for} \quad s \leq t$$

Usually, it's assumed $X(0) = 0$

$X(t)$ shows the number of events up to time t

$X(t) - X(s)$ shows the number of events occurred during the interval $(s, t]$

The occurrence of each event is referred to as an "arrival"

Bernoulli Counting Process

Definition

Bernoulli counting process $\{Y(t), t = 0, 1, ...\}$ is a discrete-time process defined as:

$$Y(0) = 0$$

$$Y(t) = \sum_{i=1}^{t} X(i), \quad t = 1, 2, ...$$

where X is a IID Bernoulli process, $X(t) \sim B(1, p)$, $t = 1, 2, ...$

By definition, Bernoulli counting process is SII time series

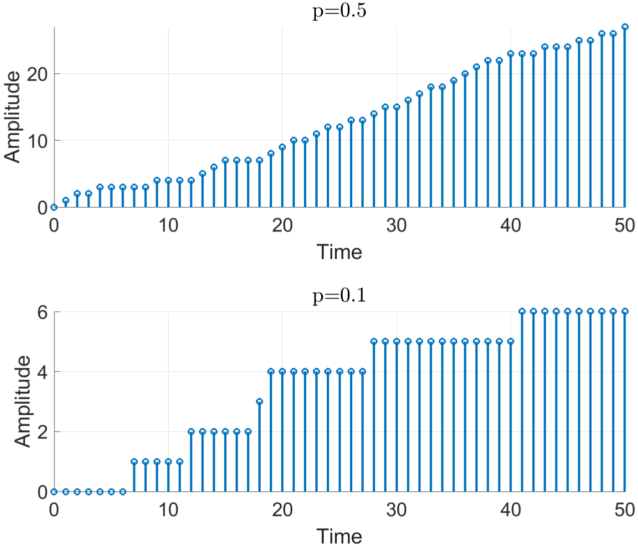
If $p = 0.5$, then the Bernoulli process is called symmetric

Expectation and covariance:

$$m(t) = \mathrm{M}[Y(t)] = t\mathrm{M}[X(t)] = tp, \quad t = 0, 1, ...$$

$$cov(t, s) = t\mathrm{D}[X(t)] = tp(1 - p), \quad t < s$$

Bernoulli Counting Process. Illustrations



Poisson Process

Definition

The counting process $\{N(t), t \in [0, \infty)\}$ is called a Poisson process with rate (or intensity) $\lambda > 0$ if all the following conditions hold:

1) N is SII process

2) $N(0) = 0$

3) The number of arrivals in any interval of length $\tau > 0$ has Poisson distribution $Pois(\lambda\tau)$

The Poisson process is one of the most widely-used counting processes

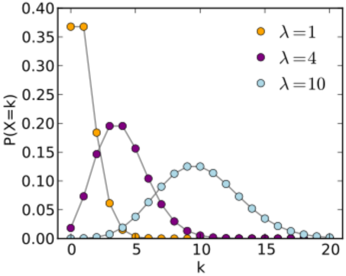
It is usually used in scenarios where we are counting the occurrences of certain events that appear to happen at a certain rate, but completely at random (without a certain structure)

Poisson Distribution

Poisson distribution is discrete probability distribution with support $\{0, 1, 2, ...\}$ and probabilities defined as:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, ...$$

The Poisson distribution $Pois(\lambda)$ can be viewed as a limiting case of the binomial distribution $B(n, p)$ as the number of trials $n \to \infty$ and the expected number of successes $np = \lambda$ remains fixed



Shrunk Bernoulli Counting Process

Consider a Bernoulli counting process $\{N_\Delta(t), t = 0, \Delta, 2\Delta, ...\}$, $\Delta \ll 1$, where time scale was shrunk:

$$\{0, 1, 2, ...\} \to \{0, \Delta, 2\Delta, ...\}$$

The random variable $N_\Delta(t)$ is a number of arrivals passed up to time moment $t = n(t)\Delta$, where $n(t)$ is the number of events passed up to time moment t, $n(t) \in \{0, 1, 2, ...\}$:

$$n(t) = \left[\frac{t}{\Delta}\right]$$

For Bernoulli counting process:

$$N_\Delta(t) = \sum_{i=1}^{n(t)} X(i), \quad t = \Delta, 2\Delta, ...$$

where $X(i) \sim B(1, p)$ is outcome of i-th event, $i = 1, 2, ...$

Shrunk Bernoulli Counting Process

To keep arrival rate constant, assume the probability of arrivals:

$$p = P(X(i) = 1) = \lambda\Delta, \quad i = 1, 2, ...$$

where $\lambda > 0$

The expected number of successes passed up to time t:

$$n(t)p = \left[\frac{t}{\Delta}\right]\lambda\Delta \approx \lambda t$$

doesn't depend on Δ

As $\Delta \to 0$ the distribution of random variable $N_\Delta(t)$ will approach to Poisson distribution:

$$N_\Delta(t) \sim B(n(t), \lambda\Delta) = B\left(\left[\frac{t}{\Delta}\right], \lambda\Delta\right) \to Pois(\lambda t)$$

Poisson Process as a Limit of Bernoulli Counting Process

Expectation and covariance of counting process N_Δ:

$$m(t) = \mathrm{M}[N_\Delta(t)] = \mathrm{M}[X(t)]n(t) = \left[\frac{t}{\Delta}\right]\lambda\Delta \approx \lambda t, \quad t = 0, \Delta, 2\Delta, ...$$

$$cov(t,s) = \mathrm{D}[X(t)]n(t) = \left[\frac{t}{\Delta}\right]\lambda\Delta(1-\lambda\Delta) \approx \lambda(1-\lambda\Delta)t \approx \lambda t, \ t < s$$

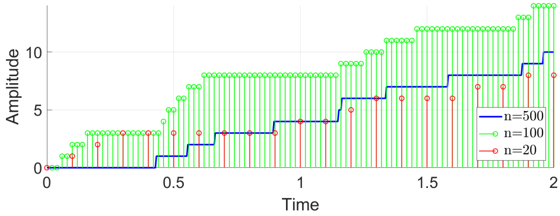
It's shown that the process $\{N_\Delta(t), t = 0, \Delta, 2\Delta, ...\}$ converges to the Poisson process $\{N(t), t \in [0,\infty)\}$ with intensity λ at $\Delta \to 0$:

$$N(t) = \lim_{\Delta\to 0} N_\Delta(t)$$

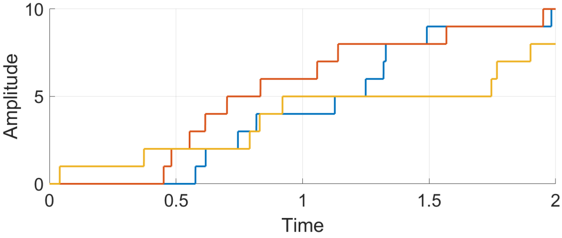
Poisson process can be viewed as Bernoulli counting process with infinitely small probable but infinitely frequent events

Poisson Process. Illustrations

Sample paths of processes N_Δ



Sample paths of process N



Properties of Poisson Process

- Poissonity
  The number of points (arrivals) in each interval $(t, t+\tau]$ has a Poisson distribution: $N(t+\tau) - N(t) \sim N(\tau) \sim Pois(\lambda\tau)$
- Stationarity
  The distribution of number of points in each finite interval $(t, t+\tau]$ depends only on interval's length τ
- Independency
  The number of points in disjoint intervals are independent random variables
- Homogeneity
  The average density of the points located in each finite interval $(t, t+\tau]$ is constant and equal to process' intensity λ
- Orderliness
  Probability of any of two points coinciding in the same time is zero

Second Definition of Poisson Process

Definition

The counting process $\{N(t), t \in [0, \infty)\}$ is called a Poisson process with rate (or intensity) $\lambda > 0$ if all the following conditions hold:

1) N is SII process
2) $N(0) = 0$
3) The probabilities of number of arrivals in any interval $(t, t+\Delta]$, $\Delta \to 0$, are

$$P(N(t+\Delta) - N(t) = 0) = 1 - \lambda\Delta + o(\Delta)$$
$$P(N(t+\Delta) - N(t) = 1) = \lambda\Delta + o(\Delta)$$
$$P(N(t+\Delta) - N(t) \geq 2) = o(\Delta)$$

It can be shown that the orderliness implies Poissonity and vice versa

First Arrival of Poisson Process

Let $\{N(t), t \in [0, \infty)\}$ be a Poisson process with rate λ

The probability of k arrivals in time interval $(0, t]$ has Poisson distribution $Pois(\lambda t)$:

$$P(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \quad k = 0, 1, 2, ...$$

Let X_1 be the time of the first arrival

The probability of no arrival in time interval $(0, t]$:

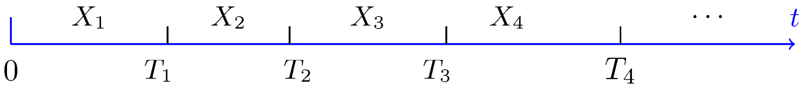
$$P(N(t) = 0) = P(X_1 > t) = e^{-\lambda t}$$

Therefore, the CDF of random variable X_1:

$$F_{X_1}(t) = P(X_1 < t) = 1 - P(X_1 \geq t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Interarrival Times of Poisson Process

Let X_2 be the time elapsed between the first and the second arrival



The probability of no arrival in time interval $(s, s+t]$ given that the first arrival was at time s:

$$P(X_2 > t | X_1 = s) = P(N(s+t) - N(s) = 0) = P(N(t) = 0) = e^{-\lambda t}$$

We conclude that $X_1 \sim Ex(\lambda t)$, $X_2 \sim Ex(\lambda)$ and random variables X_1 and X_2 are independent

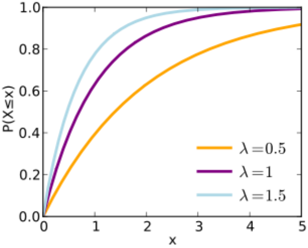
Similarly all interarrival times $X_1, X_2, \ldots$ of the Poisson process $\{N(t), t \in [0, \infty)\}$ are independent exponentially distributed random variables $X_i \sim Ex(\lambda)$

Exponential Distribution

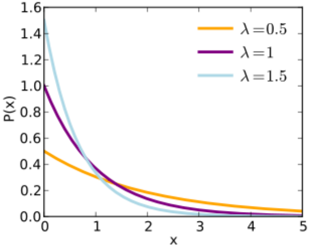
Exponential distribution is continuous probability distribution with support $[0, \infty)$ and CDF (PDF) defined as:

$$F_X(t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0 \\ 0, & \text{otherwise} \end{cases} \qquad f_X(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Exponential distribution CDF



Exponential distribution PDF



The Waiting Time Paradox

The situation:
You arrive at the bus stop, ready to catch your bus: a line that advertises arrivals every 10 minutes. You glance at your watch and note the time... and when the bus finally comes 11 minutes later, you wonder why you always seem to be so unlucky

Naïvely, you might expect that if buses are coming every 10 minutes and you arrive at a random time, your average wait would be something like 5 minutes

But in reality:
When waiting for a bus that comes on average every 10 minutes, your average waiting time will be 10 minutes

The Waiting Time Paradox. Explanation

The bus arrival process is an example of Poisson process

The probability of no arrival in time interval $(0, t_0 + t]$ given that there was no arrival in time interval $(0, t_0]$:

$$P(N(t_0+t) = 0 \mid N(t_0) = 0) = \frac{P(N(t_0+t) = 0 \;\&\; N(t_0) = 0)}{P(N(t_0) = 0)}$$

$$= \frac{P(N(t_0+t) = 0)}{P(N(t_0) = 0)} = \frac{e^{-\lambda(t_0+t)}}{e^{-\lambda t_0}} = e^{-\lambda t}$$

Therefore, the waiting time X given that there was no arrival in time interval $(0, t_0]$ has exponential distribution $X|_{N(t_0)=0} \sim Ex(\lambda)$ and the average waiting time is $\mathrm{M}[X] = 1/\lambda$

It doesn't matter how long you were waiting till time t_0, the probability of no arrival in next t minutes depends only on t

Arrival Times of Poisson Process

Let T_n be the time of n-th arrival:

$$T_n = \sum_{i=1}^{n} X_i$$

where $X_i \sim Ex(\lambda)$ is i-th interarrival time

It's known that the sum of exponentially distributed random variables follows Erlang distribution (that is Gamma distribution with integer parameter n):

$$T_n \sim Erlang(n, \lambda)$$

The Poisson process is related to probability distributions:

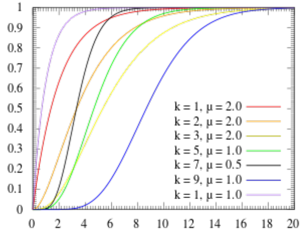
- Poisson $Pois(\lambda\tau)$
- Exponential $Ex(\lambda)$
- Erlang $Erlang(n, \lambda)$

Erlang Distribution

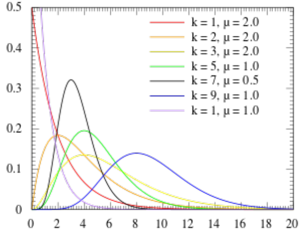
Erlang distribution is continuous probability distribution with support $[0, \infty)$ and PDF defined as:

$$f_X(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}, \quad n = 1, 2, \dots$$

Erlang distribution CDF



Erlang distribution PDF



Compound Poisson Process

Let $\{N(t), t \in [0, \infty)\}$ be a Poisson process with rate λ and and suppose that i-th arrival of the Poisson process has an associated random variable Y_i, $i = 1, 2, ...$

Definition

A compound Poisson process with rate λ is a process $\{Z(t), t \in [0, \infty)\}$ such that $Z(t) = \sum_{i=1}^{N(t)} Y_i$, where $Y_1, Y_2, ...$ are IID random variables and also independent of the process N

If all $Y_i = 1$, then Z is a simple Poisson process

If all $Y_i \in \{1, 2, ...\}$, then Z is a Poisson process which has the feature that two or more events occur in a very short time

Examples: random amount of money spent by a customer, random amount of time spent by a visitor on website, etc.

Applications of Poisson Processes

Poisson process is one of the most important tools to model the natural phenomena

- Number of calls to a phone number
- Arrival of email messages
- Visits to a website
- Traffic accidents at an intersection
- The points scored by a football team
- Occurrences of natural disasters
- Number of radioactive particles emitted during radioactive decay
- Claims received by insurance company
- Customers arrive at a shopping mall
- Earthquakes at a particular location
- ...

Simple Random Walk

Definition

Simple random walk $\{Y(t), t = 0, 1, ...\}$ is a discrete-time process defined as:

$$Y(0) = 0$$

$$Y(t) = \sum_{i=1}^{t} X(i), \quad t = 1, 2, ...$$

where $\{X(t), t = 1, 2, ...\}$ is a IID bipolar process, $X(t) \in \{-1, 1\}$, $P(X(t) = 1) = p$ and $P(X(t) = -1) = 1 - p$

By definition, simple random walk is SII time series

Expectation and covariance:

$$m(t) = \mathrm{M}[Y(t)] = t\mathrm{M}[X(t)] = (2p - 1)t, \quad t = 0, 1, ...$$

$$cov(t, s) = t\mathrm{D}[X(t)] = 4p(1 - p)t, \quad t < s$$

Simple Random Walk. Illustration

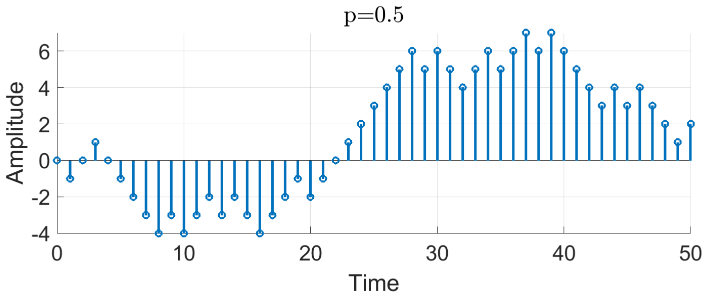
The increments for Bernoulli counting process:

$$Y(t+1) - Y(t) \in \{0, 1\}$$

The increments for simple random walk:

$$Y(t+1) - Y(t) \in \{-1, 1\}$$

A sample function of simple random walk



2D Random Walk

Consider 2D random process $\{(X(t), Y(t)), t = 0, 1, ...\}$, where processes X and Y are independent simple random walks

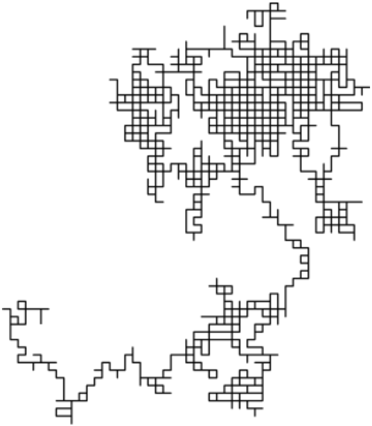
At any time moment $t = 0, 1, ...$ the increment of 2D random walk is one of four equally probable vectors:

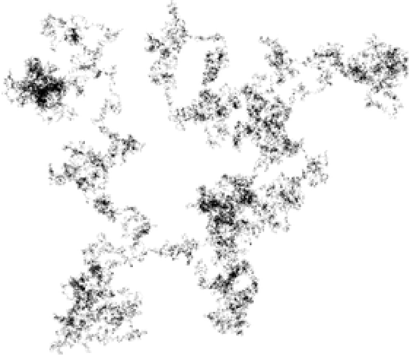
$$(-1, -1),\ (-1, 1),\ (1, -1),\ (1, 1)$$

To visualize 2D random walk one can imagine a person walking randomly around a city. The city is effectively infinite and arranged in a square grid of sidewalks. At every intersection, the person randomly chooses one of the four possible routes (including the one originally traveled from)

2D random walk is a random walk on the set of all points in the plane with integer coordinates

2D Random Walk. Illustration





https://upload.wikimedia.org/wikipedia/commons/c/cb/Random_walk_25000.svg

Wiener Process

Definition

Wiener process $\{W(t), t \in [0, \infty)\}$ is a continuous-time process with the following properties:

1) W is SII process
2) $W(0) = 0$
3) Increments $W(t+\tau) - W(t)$ have distribution $N(0, \tau)$ for all $t \in [0, \infty)$, $\tau > 0$
4) $W(t)$ is a continuous function of time on $[0, \infty)$

Wiener process is a model of Brownian motion

It describes the position of a Brownian particle in one dimension, starting at an arbitrary time which we designate as $t = 0$, with the initial position designated as $w = 0$

Standartization of Random Walk

Consider a symmetric simple random walk $\{Y(n), n = 0, 1, ...\}$ on the lattice of integers:

$$Y(0) = 0, \quad Y(n) = \sum_{i=1}^{n} \xi(i), \quad n = 1, 2, ...$$

where $\{\xi(n), n = 1, 2, ...\}$ is a IID bipolar process and

$$p = P(\xi(n) = 1) = P(\xi(n) = -1) = 0.5, \quad n = 1, 2, ...$$

Expectation: $m(n) = \mathrm{M}[Y(n)] = (2p - 1)n = 0, \quad n = 0, 1, ...$

Variance: $c(n, n) = \mathrm{D}[Y(n)] = 4p(1 - p)n = 4\frac{1}{2}\left(1 - \frac{1}{2}\right) n = n$

The Central Limit Theorem asserts that

$$\frac{Y(n)}{\sqrt{n}} \sim N(0, 1), \quad n \to \infty$$

Wiener Process as a Limit of Scaled Random Walk

Consider a continuous-time process $\{W_n(t), t \in [0, \infty)\}$, where random variable $W_n(t)$ is defined as:

$$W_n(t) = \frac{Y([nt])}{\sqrt{n}}$$

where $[nt]$ is the largest integer less than nt

The random variable $W_n(t)$ can be viewed as an amplitude of simple random walk with step $\sqrt{n}$ after $[nt]$ steps

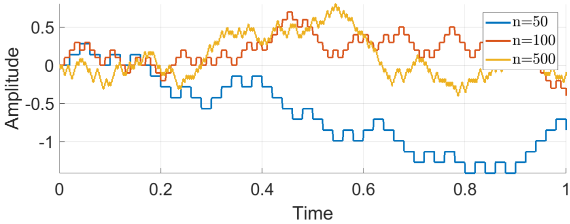
It's shown that the process $\{W_n(t), t \in [0, \infty)\}$ converges to Wiener process $\{W(t), t \in [0, \infty)\}$ at $n \to \infty$ (Donsker's theorem):

$$W(t) = \lim_{n\to\infty} W_n(t) = \lim_{n\to\infty} \frac{Y([nt])}{\sqrt{n}}$$

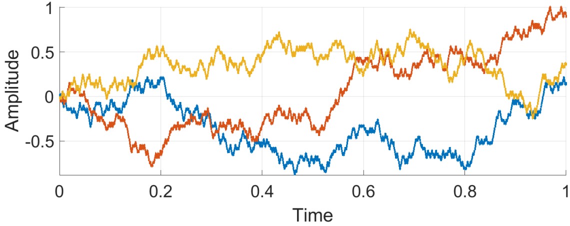
Wiener process can be viewed as random walk with infinitely small but infinitely frequent steps

Wiener Process. Illustrations

Sample paths of processes W_n



Sample paths of process W



Properties of Wiener Process

- Random variable $W(t) \sim N(0,t)$ for all $t \in [0,\infty)$
  By definition, $W(t+\tau) - W(t) \sim N(0,\tau)$ and $W(0) = 0$

  Assume $t = 0 \Rightarrow W(\tau) \sim N(0,\tau)$ for all $\tau \in [0,\infty)$
- Expectation, covariance and correlation:

$$\mathrm{M}[W(t)] = 0$$

$$cov(W(t), W(s)) = \min(t,s)$$

$$\rho(W(t), W(s)) = \frac{cov(W(t), W(s))}{\sigma(t)\sigma(s)} = \frac{\min(t,s)}{\sqrt{ts}} = \sqrt{\frac{\min(t,s)}{\max(t,s)}}$$

- Wiener process is temporally and spatially homogeneous
  If we "restart" Wiener process at a fixed time s, and shift the origin to $W(s)$, then we have another Wiener process

Gaussianity of Wiener Process

Consider distribution of random vector $(W(t_1), W(t_2))$, $t_1 < t_2$

W is SII process $\Rightarrow W(t_1)$ and $W(t_2 - t_1)$ are independent random variables and $W(t_2 - t_1)$ is distributed as $W(t_2) - W(t_1)$:

$$W(t_1) \sim N(0, t_1)$$

$$W(t_2) - W(t_1) \sim N(0, t_2 - t_1)$$

The joint CDF of random vector $(W(t_1), W(t_2))$:

$$\begin{aligned} &F_{t_1,t_2}(x_1, x_2) = P(W(t_1) < x_1 \;\&\; W(t_2) < x_2) \\ &= P(W(t_1) < x_1 \;\&\; (W(t_1) + W(t_2 - t_1) < x_2)) \\ &= P(W(t_1) < x_1 \;\&\; W(t_2 - t_1) < x_2 - x_1) = F_{t_1}(x_1) F_{t_2 - t_1}(x_2 - x_1) \end{aligned}$$

Therefore, 2D and any finite dimensional distributions of Wiener process (by induction) are multivariate normal distributions, i.e. Wiener process is a Gaussian process

Generalizations of Wiener Process

Definition

A Wiener process with drift m and variance σ^2 is a process $\{X(t), t \in [0, \infty)\}$ defined as

$$X(t) = mt + \sigma W(t)$$

where $\{W(t), t \in [0, \infty)\}$ is a standard Wiener process

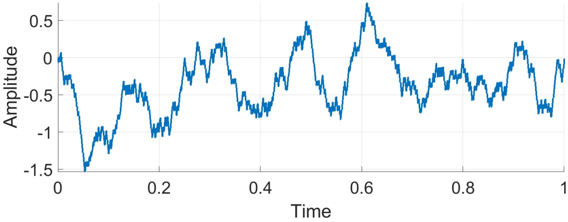
Definition

A n-dimensional Wiener process is a random process $(W_1, ..., W_n)$ with the following properties:

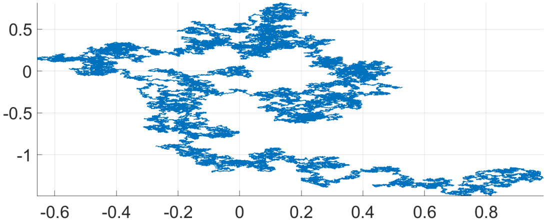
1) Each W_i is a one-dimensional Wiener process, $i = 1, ..., n$
2) Processes W_i and W_j are independent, $i, j = 1, ..., n$, $i \neq j$

Generalizations of Wiener Process. Illustrations

Wiener process with drift $m = 1$ and $\sigma = 2$



2D Wiener process



Wiener Process and White Gaussian Noise

The increment of Wiener process W at time t:

$$\Delta W(t) = W(t + \Delta t) - W(t) \sim N(0, \Delta t)$$

The derivative of Wiener process W at time t:

$$V(t) = \frac{dW(t)}{dt} = \lim_{\Delta t \to 0} \frac{\Delta W(t)}{\Delta t}$$

$$\mathrm{M}[V(t)] = \mathrm{M}\left[\frac{dW(t)}{dt}\right] = \lim_{\Delta t \to 0} \frac{0}{\Delta t} = 0$$

$$\mathrm{D}[V(t)] = \mathrm{D}\left[\frac{dW(t)}{dt}\right] = \frac{\mathrm{D}[dW(t)]}{dt^2} = \lim_{\Delta t \to 0} \frac{\Delta t}{(\Delta t)^2} = \infty$$

For any time moment t the derivatives $V(t)$ are normally distributed IID random variables with zero-mean and infinite variance, therefore, the derivative of Wiener process is continuous-time Gaussian white noise

Additive White Gaussian Noise and Brown Noise

Let $y(t)$ is a true value of observed variable (e.g. output of some dynamical system), $\tilde{y}(t)$ is a measured value of $y(t)$

Noise is unwanted (and, in general, unknown) modifications that a signal may suffer during capture, storage, transmission, processing, or conversion

Additive white Gaussian noise

$$\tilde{y}(t) = y(t) + \varepsilon(t)$$

$$\varepsilon(t) \sim N(0, \sigma^2)$$

Additive Brown noise

$$\tilde{y}(t) = y(t) + \nu(t)$$

$$\dot{\nu}(t) = \varepsilon(t)$$

$$\nu(0) = 0$$

$$\varepsilon(t) \sim N(0, \sigma^2)$$

Applications of Wiener Processes

The Wiener process is a natural model of Brownian motion

It describes a random continuous motion of a particle, subjected to the influence of a large number of chaotically moving molecules of the liquid

- Brownian noise (i.e. noise due to the Brownian motion) appears in many areas (e.g. Brownian noise of diaphragm in ultrasensitive pressure sensors, in electrical circuits, etc.)
- Brownian motion in the stock market
- Swimming of microorganisms in a pool of liquid
- Brownian agents to model collective social processes (online communities, etc.)
- Brownian motion in medical imaging (modelling of random textures, etc.)
- ...

Poisson Process vs Wiener Process

Property	Poisson Process	Wiener Process
Index set	$[0, \infty)$	$[0, \infty)$
State space	$\{0, 1, 2, ...\}$	$(-\infty, \infty)$
Parameters	Intensity $\lambda > 0$	Drift m, variance σ^2
IID process	No	No
SII process	Yes	Yes
Counting process	Yes	No
Gaussianity	No	Yes
WSS process	No	No
SSS process	No	No
Expectation $m(t)$	λt	mt
Variance $\sigma^2(t)$	λt	$\sigma^2 t$
Covariance $cov(t, s)$	$\lambda t, t < s$	$\sigma^2 t, t < s$