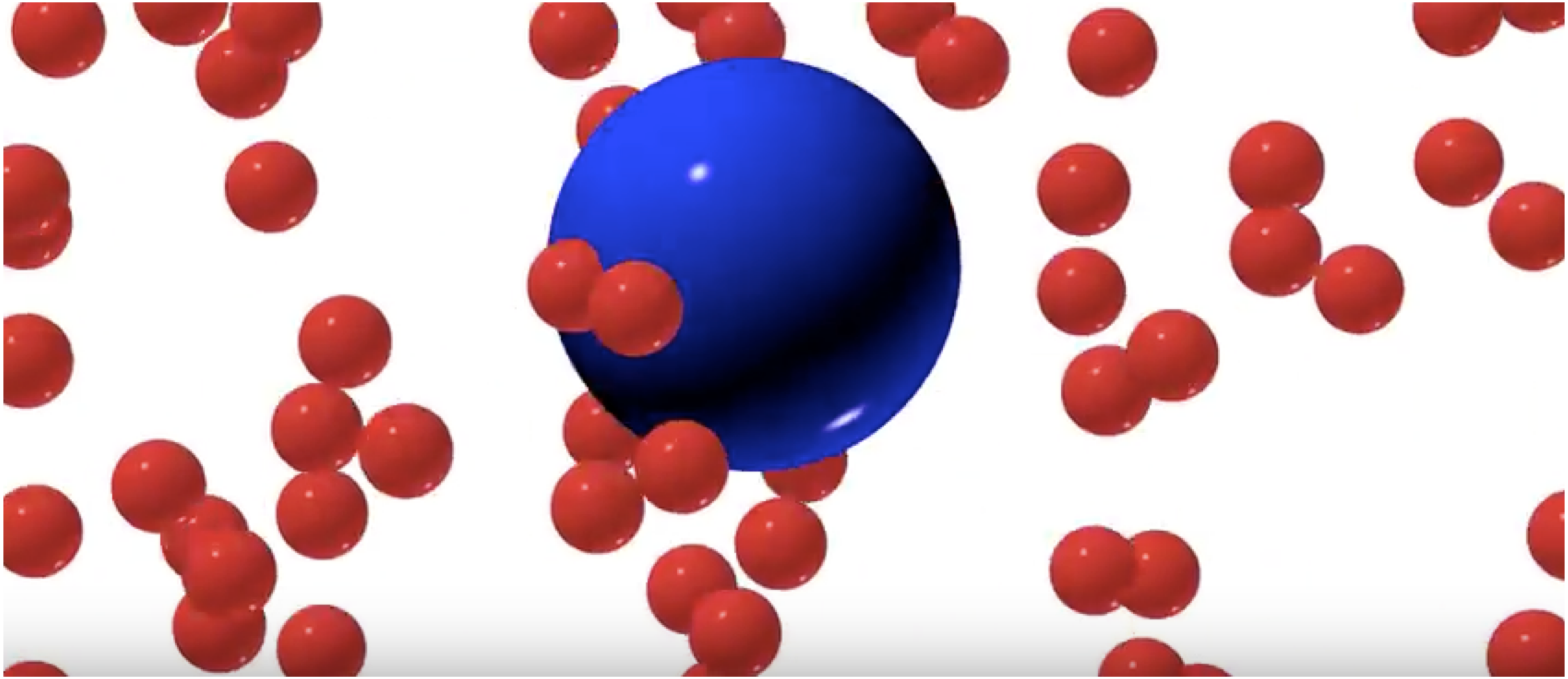
The Physics of Energy

Luca Gammaitoni

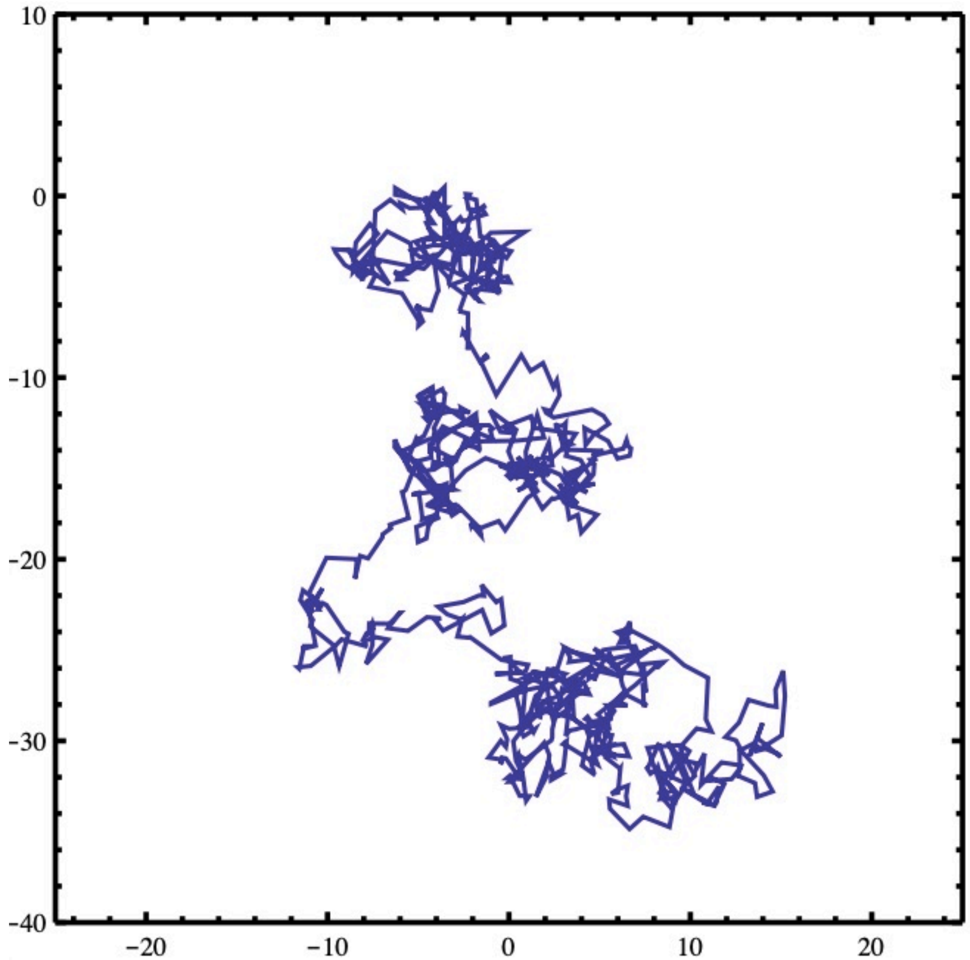
Corso di Laurea in Fisica, 2020-2021

Brownian motion



https://www.youtube.com/watch?v=6VdMp46ZIL8

Highly irregular motion trajectory

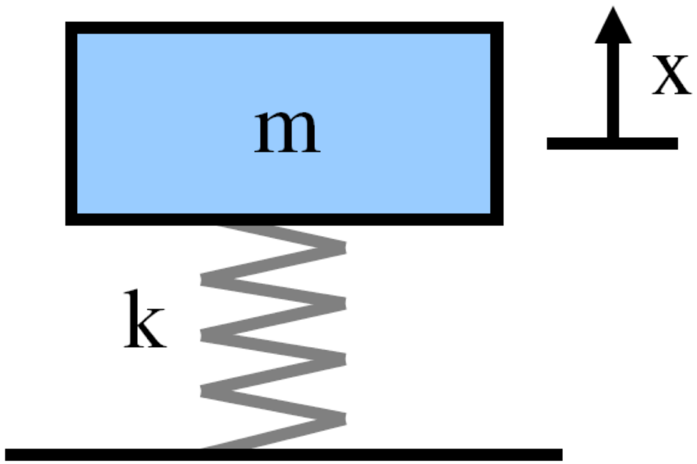


Random data vs deterministic data

Deterministic

Deterministic data are those can be described by an explicit mathematical relationship

Deterministic



$$x(t) = X \cos\left(\sqrt{\frac{k}{m}}t\right)$$

Non deterministic

- There is no way to predict an exact value at a future instant of time

- These data are random in character and must be described in terms of probability statements and statistical averages

In practical terms

The decision of whether physical data are deterministic or random is usually based on the ability of reproduce the data by controlled experiments

In practical terms

- If the experiment can be repeated producing identical data, within the limits of experimental error -> deterministic

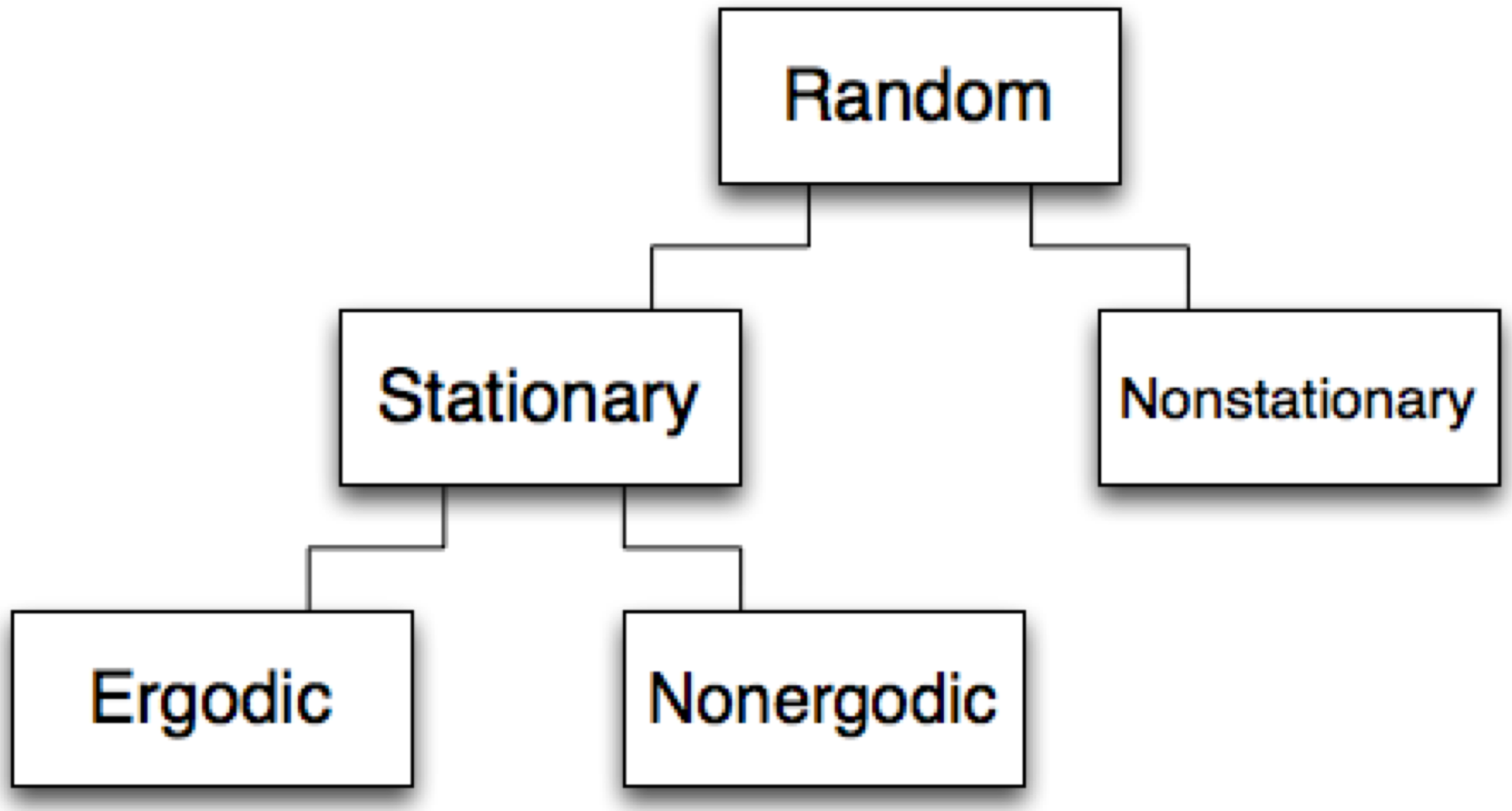
- If an experiment cannot be designed that will produce identical results when repeated -> non deterministic (random)

Terminology

- A single time history representing a random phenomena is called a *sample function* or a *sample record*
- The collection of all *sample function* that a random phenomenon might have produced is called *random process* or *stochastic process*
- A *sample record* of data may be thought of as one physical realization of a *random process*

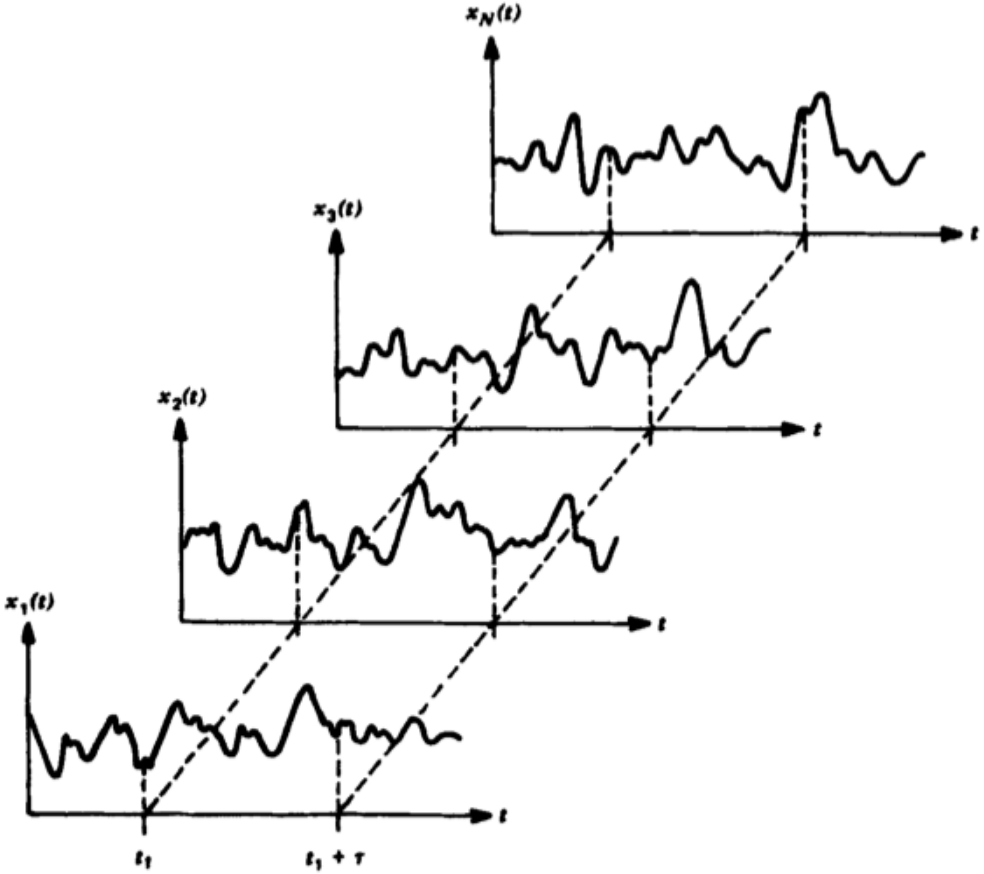
Classification of random data

Random data classification



Statistical properties

Considering a collection of *sample functions*:



Statistical properties

Considering a collection of *sample functions*:

- mean value of a *random process* at some time t_i can be computed averaging all instantaneous values of each sample
- correlation between values at two different times is the average of the product of instantaneous values at time t_i and t_i+tau

Statistical properties

$$\mu_x(t_i) = \lim_{N \to +\infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_i)$$

$$R_{xx}(t_i, t_i + \tau) = \lim_{N \to +\infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_i) x_k(t_i + \tau)$$

Stationary vs Nonstationary

- When *mean value* or *autocorrelation* vary as time t_i varies, the *random process* is said to be *nonstationary*
- When *mean value* and *autocorrelation* do not vary as time t_i varies, the *random process* is said to be *weakly stationary or* stationary in a wide sense

Weakly stationary

For weakly stationary random process, the mean value is constant and the autocorrelation function is dependent only on the time displacement tau.

$$\mu_x(t_i) = \mu_x$$

$$R_{xx}(t_i, t_i + \tau) = R_{xx}(\tau)$$

Weakly stationary

In most case it is possible to describe the properties of a stationary *random process* by computing time averages over specific *sample function*

$$\mu_x(k) = \lim_{T\to\infty} \frac{1}{T}\int_0^T x_k(t)\,dt$$

$$R_{xx}(\tau, k) = \lim_{T\to\infty} \frac{1}{T}\int_0^T x_k(t)x_k(t+\tau)\,dt$$

Ergodic random data

- If mean value and autocorrelation function do not differ over different *sample functions* the random process is said to be *ergodic*

$$\mu_x(k) = \mu_x$$

$$R_{xx}(\tau, k) = R_{xx}(\tau)$$

- Only stationary random process can be ergodic

Ergodic random data

- Ergodic random process are an important class of random processes
- All properties of ergodic random process can be determined by a single sample function
- Fortunately, in practice, random data representing stationary physical phenomena are generally ergodic

Nonstationary random data

- The properties of nonstationary random process are generally time-varying function
- In practice it is often not feasible to obtain a sufficient number of sample records to permit an accurate measurement of properties of the ensemble
- This has tend to impede the development of practical techniques for measuring and analyzing nonstationary random data

Stationary sample records

Data in the form of sample records are referred to be stationary or nonstationary

$$\mu_x(t_i, k) = \frac{1}{T}\int_{t_i}^{t_i+T} x_k(t)\,dt$$

$$R_{xx}(t_i, t_i+\tau, k) = \frac{1}{T}\int_{t_i}^{t_i+T} x_k(t)x_k(t+\tau)\,dt$$

Stationary sample records

- A single time series is referred to be stationary if properties computed over short time intervals do not vary significantly from one interval to the next
- If the sample properties vary significantly as the starting time t_i varies the individual sample record is said to be nonstationary

Stationary sample records

- A *sample record* obtained from an ergodic *random process* will be stationary
- *Sample records* from nonstationary *random process* will be nonstationary
- Hence if an ergodic assumption is justified verification of stationarity of a single *sample record* will justify an assumption of stationarity and ergodicity for the *random process*

Analysis of random data

Analysis of random data

Since no explicit mathematical equation can be written, statistical procedures must be used to define the descriptive properties of the data

Basic descriptive properties

- Probability density functions
- Mean and mean square values
- Autocorrelation functions
- Power spectral density functions

Joint statistical properties

- Joint probability density functions
- Cross-correlation functions
- Cross-spectral density functions
- Frequency response functions
- Coherence functions

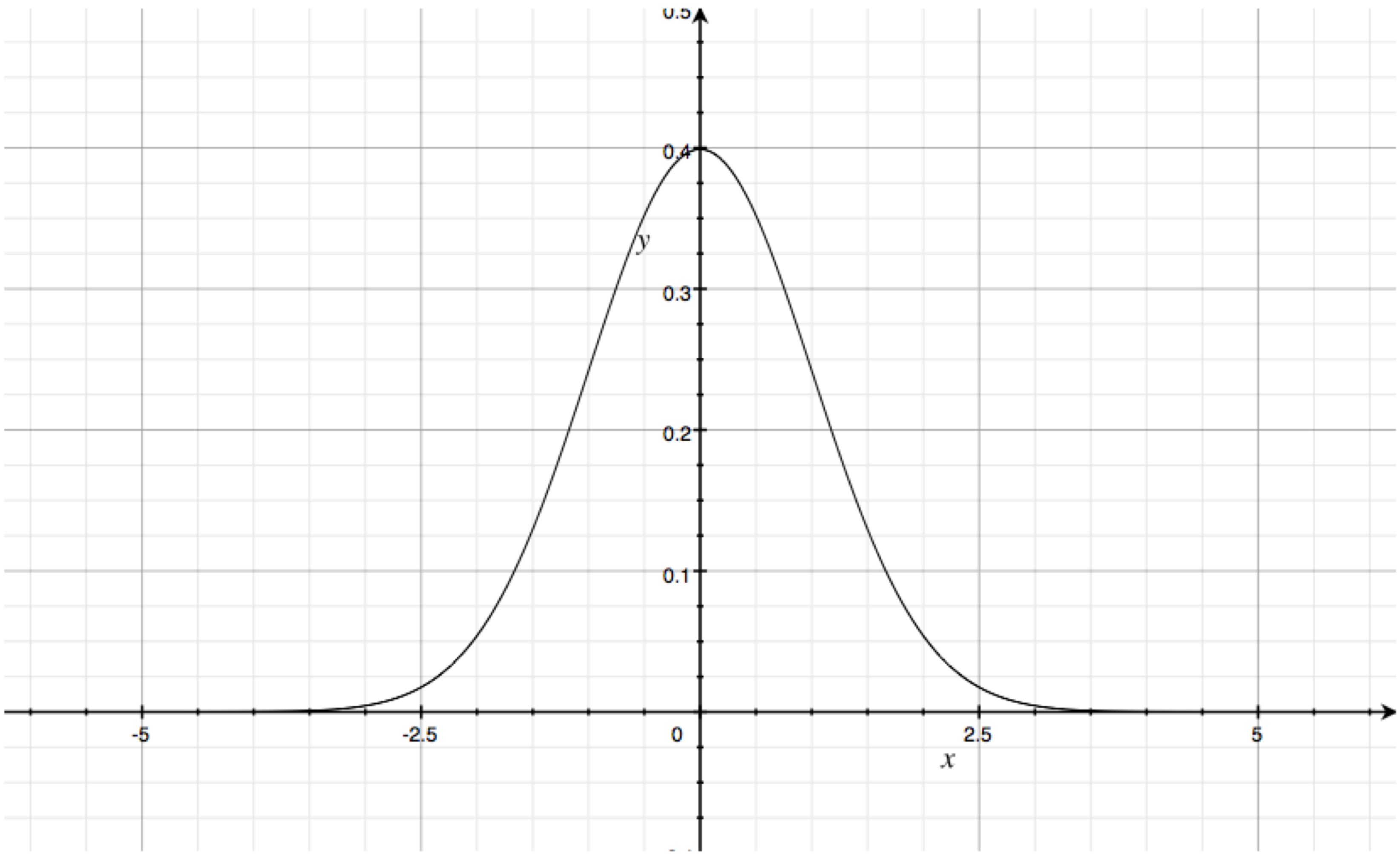
Probability density function

- Is a function that describes the relative likelihood for a random variable to take on a given value

- The probability of the random variable falling within a particular range of values is given by the integral of this variable's density over that range

Probability density function: example

PDF(x)



Probability density function

$$\Pr[a \leq X \leq b] = \int_a^b f_X(x)\, dx$$

$$CDF_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)\, du$$

Time domain analysis

Expected value

$$\mathrm{E}[X] = \sum_{i=1}^{\infty} x_i \, p_i,$$

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x) \, dx$$

Arithmetic mean is an estimator of the expected value of a random process

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\mathrm{Var}(\bar{x}) = \frac{\sigma^2}{n}$$

Expected value

$$\mathrm{E}[X+c] = \mathrm{E}[X] + c$$

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$

$$\mathrm{E}[aX] = a\,\mathrm{E}[X]$$

$$\mathrm{E}[XY] = \int\int xy\,j(x,y)\,dx\,dy$$

$$\mathrm{Cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]$$

if x and y are independent ?

$$\mathrm{E}[XY] = \int\int xy\,j(x,y)\,dx\,dy = \int\int xy f(x) g(y)\,dy\,dx$$

$$= \left[\int x f(x)\,dx\right]\left[\int y g(y)\,dy\right] = \mathrm{E}[X]\,\mathrm{E}[Y]$$

Variance

$$\begin{aligned}\operatorname{Var}(X) &= \operatorname{E}\left[(X-\mu)^2\right] \\ &= \operatorname{E}\left[X^2 - 2X\operatorname{E}[X] + (\operatorname{E}[X])^2\right] \\ &= \operatorname{E}\left[X^2\right] - 2\operatorname{E}[X]\operatorname{E}[X] + (\operatorname{E}[X])^2 \\ &= \operatorname{E}\left[X^2\right] - (\operatorname{E}[X])^2\end{aligned}$$

$$S_n^2(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad s_n^2(X) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Chebyshev's inequality

$\forall a > 0$

$$\mathbb{P}(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$$

$$\mathbb{P}(|X - \mu| \geq a\sigma) \leq \frac{1}{a^2}$$

no more than $1/a^2$ of the distribution's values can be more than "*a*" standard deviations away from the mean

Autocorrelation

$$C_{xx}(\tau) = \mathrm{E}[(X_t - \mu)(X_{t+\tau} - \mu)]$$

$$R_{xx}(\tau) = \frac{\mathrm{E}[(X_t - \mu)(X_{t+\tau} - \mu)]}{\sigma^2}$$

$$\langle x(t)x(t-\tau)\rangle = \lim_{T->\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)x(t-\tau)dt$$

Often the autocovariance is called autocorrelation even if this normalization has not been performed and vice-versa