PHY408

Lecture10: Stochastic Processes

March 22, 2023

Random Variables

1. Random variable X: when measured or inspected, has a value that is not necessarily repeatable or predictable. Realization of X.
2. characterize random variable X: probability density function (pdf), from which mean (expected value) and variance (scatter about the mean) can be extracted.
3. PDF of X: $f(x)$ gives the probability of X taking on values between $[x, x + dx]$ as $dp = f(x)dx$.

$$\int_{\mathbb{R}} f(x)dx = 1$$

$$\mu_x = \int_{\mathbb{R}} xf(x)dx$$

$$V_x = \int_{\mathbb{R}} (x - \mu_x)^2 f(x)dx$$

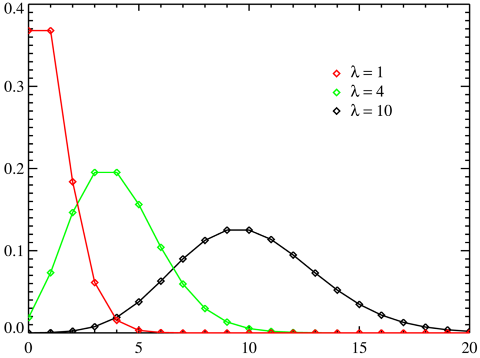
$$\mu_q = \int_{\mathbb{R}} q(x)f(x)dx$$

Examples of Analytical PDF

1. Poisson distribution for discrete random variable n:

$$f(n;\lambda) = \frac{\lambda^n e^{-\lambda}}{n!}, \quad n \in \mathbb{N} \tag{1}$$

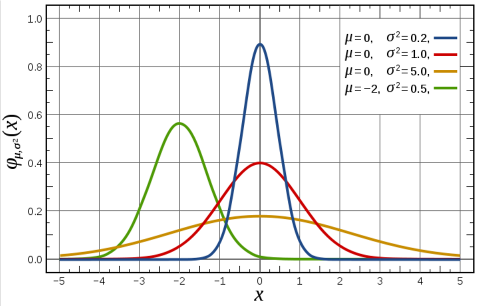
gives probability of exact n occurrences within a fixed time interval if the expected number of occurrences is λ (may not be an integer)



Normal

Normal distribution (also known as Gaussian distribution or bell curve)

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{2}$$

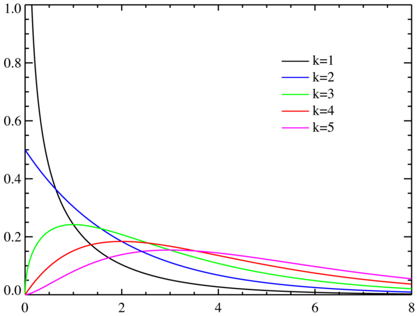


Chi-square

χ^2-distribution of degree k: $S_k = X_1^2 + X_2^2 + \cdots X_k^2$, X_i independent with standard normal distribution $N(0,1)$

$$f(s) = \frac{1}{\Gamma(k/2)2^{k/2}} s^{k/2-1} e^{-s/2} \tag{3}$$

1. $k=1$, $f(s) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{s}}e^{-s/2}$; $k=2$, $f(s) = \frac{1}{2}e^{-s/2}$



2. $E(S_k) = k$, $Var(S_k) = 2k$. Based on CLT: $S_k \sim N(k, \sqrt{2k})$.

Multivariate PDFs

Join PDF's of two or more variables: $f(x, y)$

1. covariance (a deterministic number): dependence between x and y.

$$C_{xy} = \iint (x - \mu_x)(y - \mu_y) f(x, y) dxdy \tag{4}$$

   Define $C_{xy}^{max} = \sqrt{V_x V_y}$ and correlation coefficient:

$$\rho_{xy} = \frac{C_{xy}}{C_{xy}^{max}} = \frac{C_{xy}}{\sqrt{V_x V_y}} \tag{5}$$

   and $\rho_{xy} \in [-1, 1]$.

2. if X, Y are independent: $f(x, y) = f_x(x) f_y(y)$, covariance $C_{xy} = 0$

Stochastic process and autocorrelation

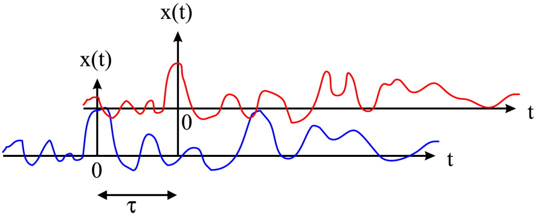
1. also named random process: counterpart of deterministic process; given as an ordered set of random variables $\{X_1, X_2, \cdots X_N, \cdots\}$ or $\{X(t)\}$ with interdependency.
2. a process whose realizations are time series; instead of being a deterministic $x(t)$, its future evolution is described by PDFs with some outcomes more probable than others; used to describe a physical process which is wholly or in part controlled by chance mechanism
3. examples: stock market, exchange rate, random movement such as Brownian motion or random walks.
4. Autocorrelation function: correlation of random variables at two different times (interchangeable with autocovariance)

$$A(t, \tau) =< X(t)X(t+\tau) > \tag{6}$$

Note $A(t, \tau)$ is deterministic.

Stationary Process

1. Stationary process (ergodic): $A(t, \tau)$ is indep. of t
$$A(\tau) = < X(t)X(t+\tau) > \tag{7}$$
reminds us of LTI systems.
2. average over realizations is equivalent to average over time
$$A(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)x(t+\tau)dt \tag{8}$$
3. This is the auto-correlation function of a realization $x(t)$ (deterministic) at the limit of $T \to \infty$.



Stationary Process

$$A(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)dt \tag{9}$$

For a time series x_i:

$$A_j = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N} x_k x_{k+j} \tag{10}$$

Practical assumptions:

- stochastic processes have zero mean
- random variables have normal PDF

Cross-correlation

Cross-correlation of two signals $f(t)$ and $g(t)$ is defined as

$$c_{fg}(\tau) = \int_{-\infty}^{\infty} f^*(t)g(t+\tau)dt \tag{11}$$

1. shift $g(t)$ by τ with respect to $f(t)$, multiply it to $f^*(t)$ and integrate over time; reminiscent of convolution
2. $c_{fg}(\tau) \neq c_{gf}(\tau)$, $c_{fg}(\tau) = c_{gf}(-\tau)$
3. $f(t) \Leftrightarrow F(\omega)$, therefore define $p(t) = f^*(t) \Leftrightarrow F^*(-\omega)$; $g(t) \Leftrightarrow G(\omega)$, let $h(t) = g(-t) \Leftrightarrow G(-\omega)$, then $g(t+\tau) = h(-\tau - t)$ and we have

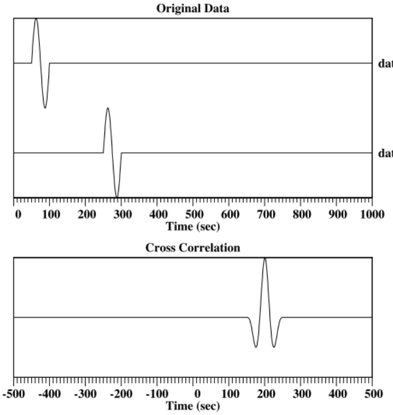
$$c_{fg}(\tau) = \int_{-\infty}^{\infty} p(t)h(-\tau - t)dt \tag{12}$$

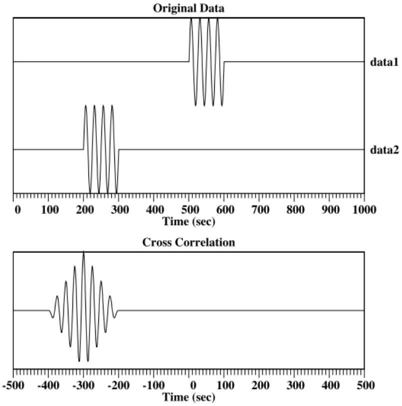
   i.e. $c_{fg}(-\tau) = p(\tau) * h(\tau)$. In the frequency domain, this translates to $C_{fg}(-\omega) = F^*(-\omega)G(-\omega)$, or

$$C(\omega) = F^*(\omega)G(\omega) \tag{13}$$

Cross-correlation examples

Cross-correlation functions can be used to measure the similarity of two signals under time translation.





Auto-correlation

If $f(t) = g(t)$, then cross-correlation is simplified to the auto-correlation of $f(t)$ itself:

$$a(\tau) = \int_{-\infty}^{\infty} f^*(t) f(t+\tau) dt \tag{14}$$

1. provides information on how similar the signal is to itself under time translation.
2. $C(\omega) = F^*(\omega)F(\omega) = |F(\omega)|^2$, i.e. proportional to energy spectrum
3. sometimes for periodic signal or ergodic signal, auto-correlation is defined as:

$$a(\tau) = \frac{1}{T} \int_0^T f^*(t) f(t+\tau) dt \tag{15}$$

then $A(\omega) = \lim_{T\to\infty} \frac{1}{T}|F(\omega)|^2$, is the power spectrum.

Power spectrum of a stationary process

Recall that

$$F_T(\omega) = \int_{-T/2}^{T/2} x(t) e^{-i\omega t} dt. \tag{16}$$

Therefore, the power spectrum can be expressed as follows:

$$\begin{aligned}
P(\omega) &= \lim_{T\to\infty} \frac{1}{T} < |F_T(\omega)|^2 > \\
&= \lim_{T\to\infty} \frac{1}{T} < \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} x(t)x(t') e^{-i\omega t} e^{i\omega t'} dt' dt > \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} < x(t)x(t') > e^{-i\omega (t-t')} dt' dt \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} A(\tau) e^{-i\omega \tau} dt' d\tau \\
&= \lim_{T\to\infty} \int_{-T/2}^{T/2} A(\tau) e^{-i\omega\tau} d\tau = \int_{-\infty}^{\infty} A(\tau) e^{-i\omega\tau}\, d\tau
\end{aligned} \tag{17}$$