Probability Spaces

A formal definition of probability involves the specification of:

- a sample space

  The sample space, S, is the set of all possible outcomes, plus the null outcome. Each element in S is called a sample point.

- a field (or algebra)

  The field, or algebra, is a set of subsets of S closed under complementation and union (recall Venn diagrams).

- and a probability measure.

  A probability measure obeys these axioms:

  1. $P(S) = 1$ (implies probabilities less than one)
  2. $P(A) \geq 0$
  3. For two mutually exclusive events:

$$P(A \cup B) = P(A, B) = P(A) + P(B)$$

Two events are said to be statistically independent if:

$$P(A \cap B) = P(A)P(B)$$

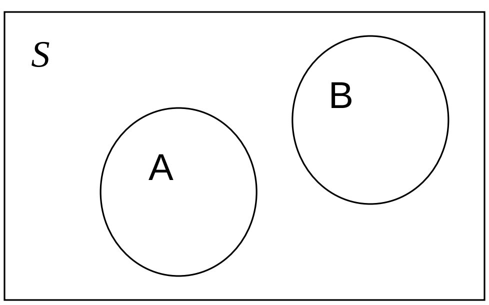
The conditional probability of B given A is:

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

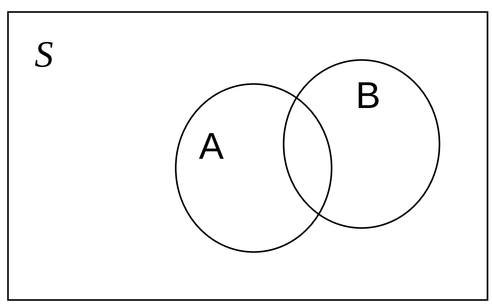
Hence,

$$P(B \cap A) = P(B|A)P(A)$$

Mutually Exclusive

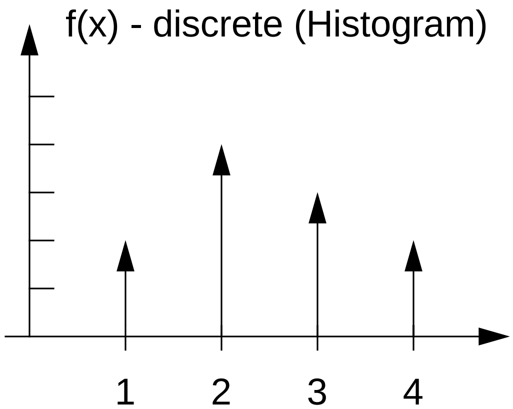


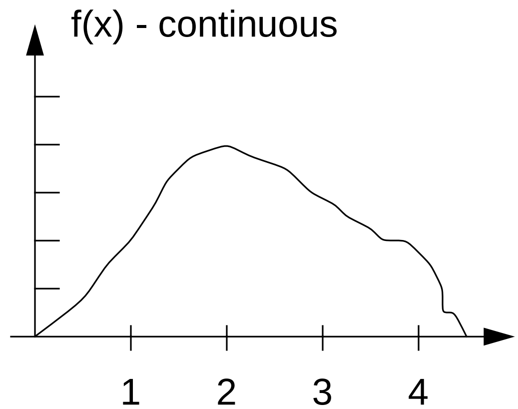
$$P(B \cap A) = P(B|A)P(A)$$



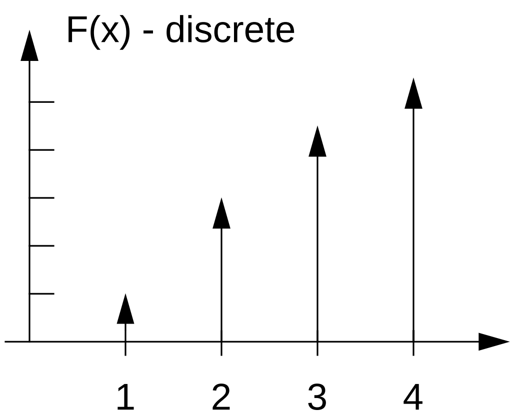
Functions of Random Variables

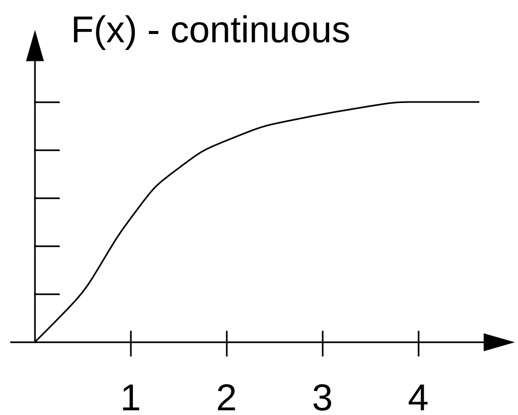
Probability Density Functions:



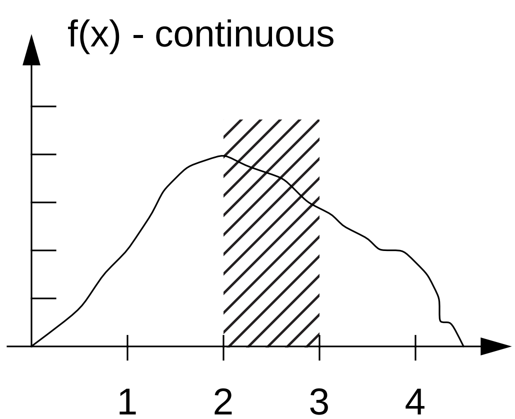


Cumulative Distributions:





Probability of Events:



$$P(2 < x \leq 3) = \int_{2}^{3} f(x)dx = F(3) - F(2)$$

Important Probability Density Functions

Uniform (Unix rand function):

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x \le b \\ 0 & elsewhere \end{cases}$$

Gaussian:

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

Laplacian (speech signal amplitude, durations):

$$f(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left\{\frac{-\sqrt{2}|x|}{\sigma}\right\}$$

Gamma (durations):

$$f(x) = \frac{\sqrt{k}}{2\sqrt{\pi|x|}} \exp\{-k|x|\}$$

We can extend these concepts to N-dimensional space. For example:

$$P(A \in A_x | B \in A_y) = \frac{P(A, B)}{P(B)} = \frac{\int_{A_x}\int_{A_y} f(x, y)dydx}{\int_{A_y} f(y)dy}$$

Two random variables are statistically independent if:

$$P(A, B) = P(A)P(B)$$

This implies:

$$f_{xy}(x, y) = f_x(x)f_y(y) \quad \text{and} \quad F_{xy}(x, y) = F_x(x)F_y(y)$$

Expectations and Moments

The statistical average of a scalar function, $g(x)$, of a random variable is:

$$E[g(x)] = \sum_{i=1}^{\infty} x_i P(x = x_i) \quad \text{and} \quad E[g(x)] \equiv \int_{-\infty}^{\infty} g(x) f(x) dx$$

The central moment is defined as:

$$E[(x-\mu)^i] = \int_{-\infty}^{\infty} (x-\mu)^i f(x) dx$$

The joint central moment between two random variables is defined as:

$$E[(x-\mu_x)^i (y-\mu_y)^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu_i)^i (y-\mu_y)^k f(x,y) dx dy$$

We can define a correlation coefficient between two random variables as:

$$\rho_{xy} = \frac{c_{xy}}{\rho_x \rho_y}$$

We can extend these concepts to a vector of random variables:

$$\bar{x} = [x_1, x_2, \ldots, x_N]^T$$

$$f_{\bar{x}}(\bar{x}) = \frac{1}{\sqrt[N/2]{2\pi}\sqrt{|\boldsymbol{C}|}} \exp\left\{-\frac{1}{2}(\bar{x}-\bar{\mu})^T \boldsymbol{C}_{\bar{x}}^{-1} (\bar{x}-\bar{\mu})\right\}$$

What is the difference between a random vector and a random process?

What does wide sense stationary mean? strict sense stationary?

What does it mean to have an ergodic random process?

How does this influence our signal processing algorithms?

Correlation and Covariance of WSS Random Processes

For a signal, $x(n)$, we can compute the following useful quantities:

Autocovariance:

$$\begin{aligned} c(i, j) &= E[(x(n-i)-\mu_i)(x(n-j)-\mu_j)] \\ &= E[x(n-i)x(n-j)]-\mu_i\mu_j \\ &= \frac{1}{N}\sum_{n=0}^{N-1} x(n-i)x(n-j) - \left(\frac{1}{N}\sum_{n=0}^{N-1} x(n-i)\right)\left(\frac{1}{N}\sum_{n=0}^{N-1} x(n-j)\right) \end{aligned}$$

If a random process is wide sense stationary:

$$c(i, j) = c(|i-j|, 0)$$

Hence, we define a very useful function known as the autocorrelation:

$$r(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)x(n-k)$$

If $x(n)$ is zero mean WSS:

$$c(i, j) = r(|i-j|)$$

What is the relationship between the autocorrelation and the spectrum:

$$DFT\{r(k)\} = |X(k)|^2$$

For a linear time-invariant system, $h(n)$:

$$DFT\{r_y(k)\} = |DFT\{h(n)\}|^2 DFT\{r_x(k)\}$$

The notion of random noise is central to signal processing:

white noise?

Gaussian white noise?

zero-mean Gaussian white noise?

colored noise?

Therefore, we now embark upon one of the last great mysteries of life:

How do we compare two random vectors?