Sources of uncertainty

- Inherent stochasticity in the system being modeled (model of quantum mechanics)
- Incompleted observability - we do not have or cannot observe all the variables that drive the system (Economic models)
- Incomplete modeling - we use a moel that discards some observe information (simplified model of the brain)

Types of distributions

- Joint Probability Distribution
- Uniform Distribution

Marginal probability: the probability distribution over a subset of all variables
With discrete random variables:
$P(X = x) = \sum_y P(X = x, Y = y)$ With continuous random variables:
$P(X = x) = \int p(x, y)dy$

Conditional probability: the probability of some event, given that some other event has happened
$P(Y = y|X = x) = \frac{P(Y=y,X=x)/P(X=x)}{}$

Variance: the expectation of the squared deviation of a random variable from its mean.Meaning: measures how far random numbers drawn from a probability distribution P(x) are spread out from their average value.
$Var(f(x)) = \sigma^2 = E[(f(x) - E(f(x)))^2])$
Standard deviation (σ)
Covariance: a measure of how much two variables are **linearly** related to each other
$Cov(f(x), g(y)) = E[(f(x) - E[f(x)])(g(y) - E[g(y)])]$

- High absolute value - values are both far from their respective means at the same time
- Positive - both variables take on large values simultaneously
- Negative = variables take on large values at different times

The Covariance Matrix
the covariance matrix of a random vector x is an n x n matrix, such that:
$Cov(x)_{i,j} = Cov(x_i, x_j)$
The diagonal elements of the covariance matrix give the variance:

$Cov(x_i, x_i) = Var(x_i)$

Special Random variables

Distributions that are commonly found in real life data or machine learning applications

Bernoulli Distribution

A distribution over a single binary random variable

$P(X = 1) = \phi$

$P(X = 0) = 1 - \phi$

Multinoulli Distribution

A distribution over a single discrete variable with k different states (categorical distribution)

Gaussian Distribution

Also called the normal distribution - most common distribution over real numbers

$\mathcal{N}(x; \mu, \sigma^2) = \sqrt{\frac{1}{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(x - \mu)^2)$

Many complicated systems can be modelle using Gaussian distribution.

68%,95%,99.7% rules: 1σ, 2σ, 3σ

Exponential Distribution

Probability istribution with sharp point at x = 0

$p(x; \lambda) = \lambda * 1_{x>=0} \exp(-\lambda x)$

All negative value of x get probability 0

Laplace Distribution

Probability distribution with a sharp point at $x = \mu$

$Laplace(x; \mu, \gamma) = \frac{1}{2\gamma} exp(-\frac{|x-\mu|}{\gamma})$