18-660: Numerical Methods for Engineering Design and Optimization

Xin Li
Department of ECE
Carnegie Mellon University
Pittsburgh, PA 15213

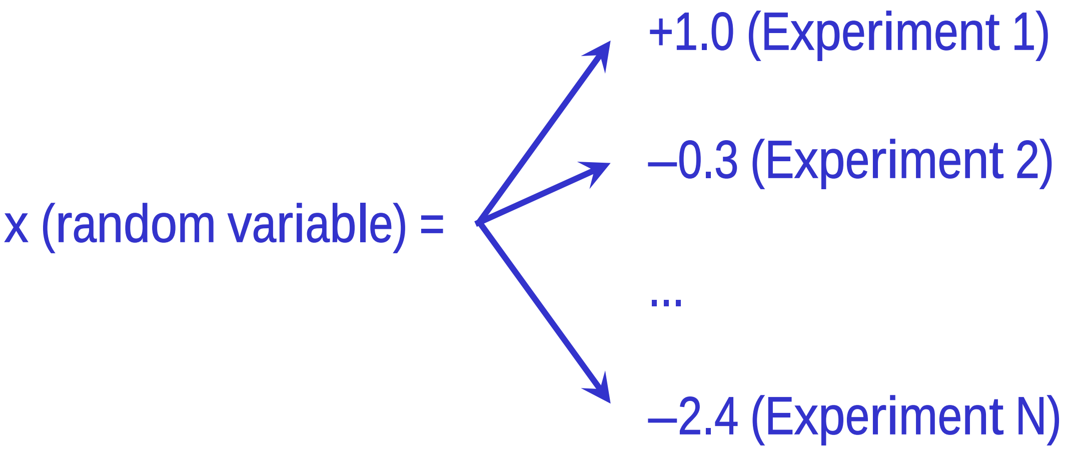
Electrical & Computer ENGINEERING

Overview

- Monte Carlo Analysis
  - Random variable
  - Probability distribution
  - Random sampling

Random Variables

- A random variable is a real-valued function of the outcome of the experiment



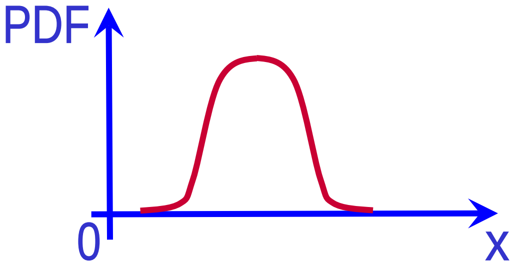
We get different results from different experiments (i.e., the output is random)

Probability Distribution

- A continuous random variable x is defined by its probability distribution function

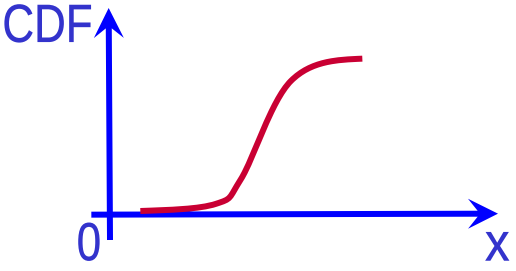
- Probability density function (PDF)
  - $\text{pdf}_x(t_x)$ denotes the probability per unit length near $x = t_x$

$$\int_a^b pdf_x(\tau_x)\cdot d\tau_x = P(a \le x \le b)$$



- Cumulative distribution function (CDF)
  - $\text{cdf}_x(t_x)$ equals the probability of $x \le t_x$

$$cdf_x(t_x) = \int_{-\infty}^{t_x} pdf_x(\tau_x)\cdot d\tau_x = P(x \le t_x)$$



Expectation

- Given a random variable x and a function f(x), the expectation of f(x) is the weighted average of the possible values of f(x)

$$E[f(x)] = \int_{-\infty}^{+\infty} f(\tau_x) \cdot pdf_x(\tau_x) \cdot d\tau_x$$

- A useful equation for expected value calculation

$$E[f(x)+g(x)] = E[f(x)] + [g(x)]$$

Mean, Variance and Standard Deviation

- Mean

$$\boxed{E[x] = \int_{-\infty}^{+\infty} \tau_x \cdot pdf_x(\tau_x) \cdot d\tau_x}$$

- Variance

$$\boxed{VAR[x] = E\left[(x - E[x])^2\right] = \int_{-\infty}^{+\infty} [\tau_x - E(x)]^2 \cdot pdf_x(\tau_x) \cdot d\tau_x}$$

- Standard deviation

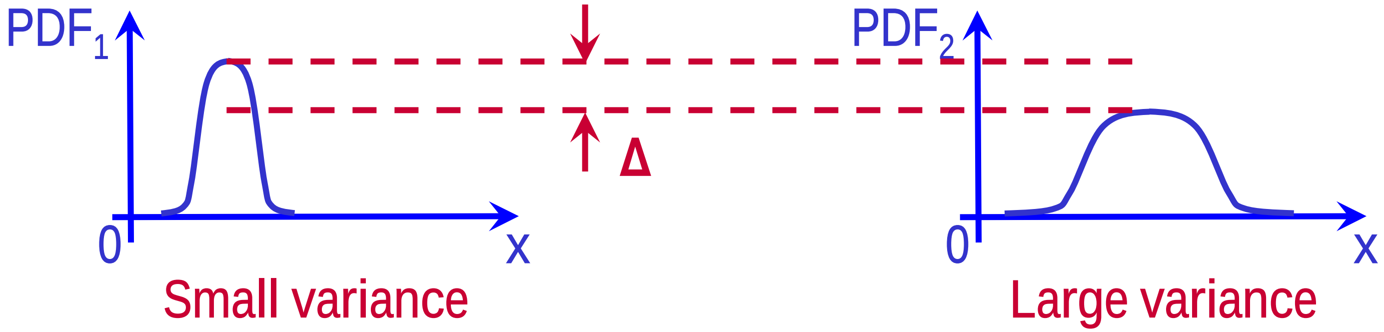
$$\boxed{STD[x] = \sqrt{VAR[x]}}$$

VAR[x] is always positive!

Mean, Variance and Standard Deviation

- Mean measures the “average position” of x

- Variance measures the “spread” of the distribution



Moments and Central Moments

- k-th order moment

$$\boxed{E\left[x^k\right]=\int_{-\infty}^{+\infty}\tau_x^k\cdot pdf_x\left(\tau_x\right)\cdot d\tau_x}$$

  - Mean is the first order moment

- k-th order central moments

$$\boxed{E\left[\left(x-E[x]\right)^k\right]=\int_{-\infty}^{+\infty}\left[\tau_x-E\left(x\right)\right]^k\cdot pdf_x\left(\tau_x\right)\cdot d\tau_x}$$

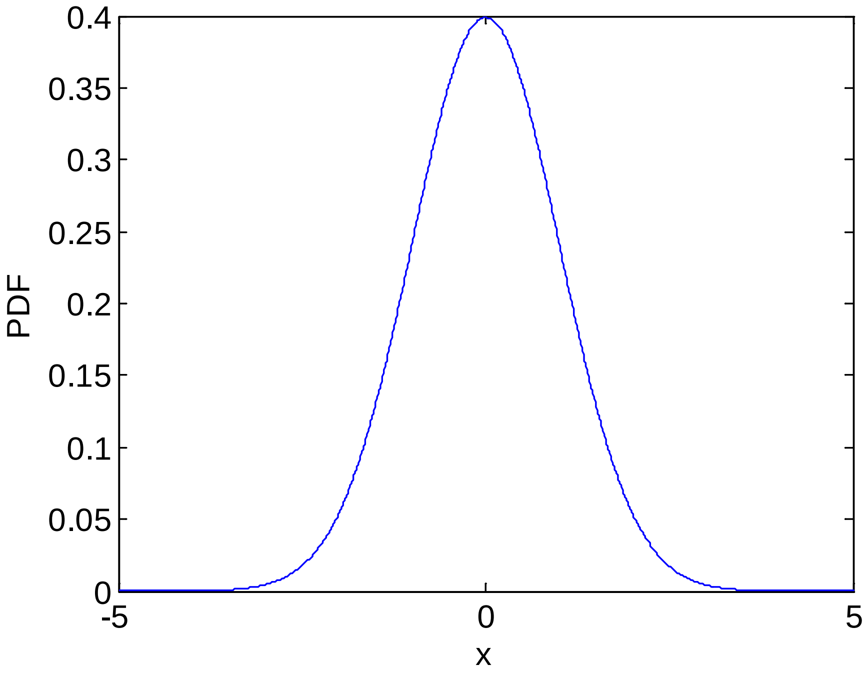
  - Variance is the second order central moment

Normal Distribution

- A random variable x is Normal if

$$pdf_x(t) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(t-\mu)^2}{2\sigma^2}}$$

  - μ: mean
  - σ: standard deviation
  - Denoted as N(μ, σ^2)



Standard Normal distribution

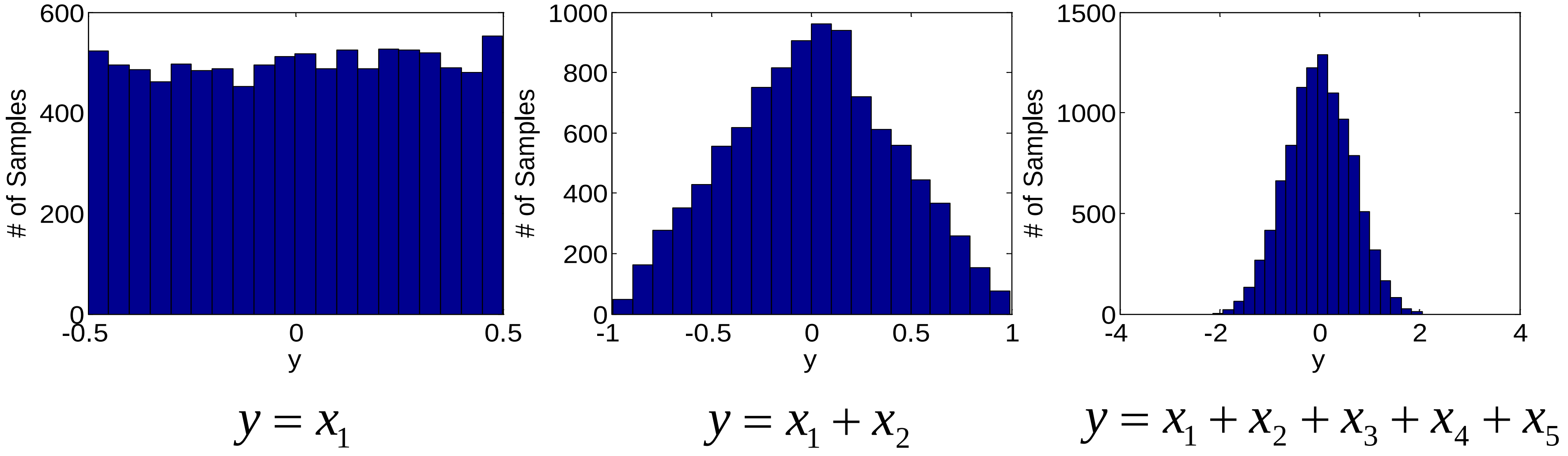
- If **μ** = 0 and **σ** = 1, it is called standard Normal distribution

Why is Normal distribution important to us?

Normal Distribution

- Many physical variations are Normal

- Central limit theorem: the variation caused by a large number of independent random factors is “almost” Normal



$$y = x_1$$

$$y = x_1 + x_2$$

$$y = x_1 + x_2 + x_3 + x_4 + x_5$$

Assume that all x_i’s are independent and have the same uniform distribution

Multiple Random Variables

- Two continuous random variables x and y are defined by their joint probability distribution
- Joint probability density function

$$\boxed{\int_c^d \int_a^b pdf_{x,y}(\tau_x, \tau_y) \cdot d\tau_x \cdot d\tau_y = P(a \le x \le b, c \le y \le d)}$$

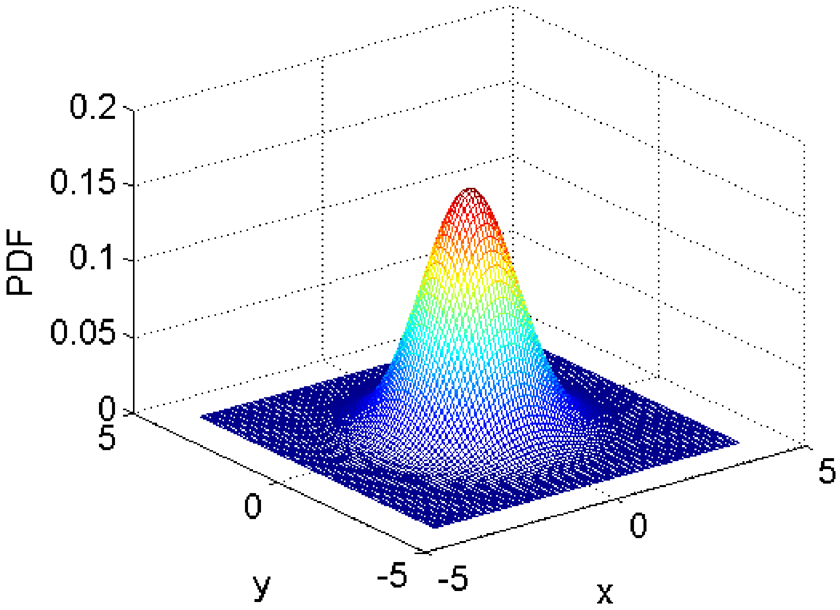
- Joint cumulative distribution function

$$\boxed{cdf_{x,y}(t_x, t_y) = P(x \le t_x, y \le t_y) = \int_{-\infty}^{t_x} \int_{-\infty}^{t_y} pdf_{x,y}(\tau_x, \tau_y) \cdot d\tau_x \cdot d\tau_y}$$

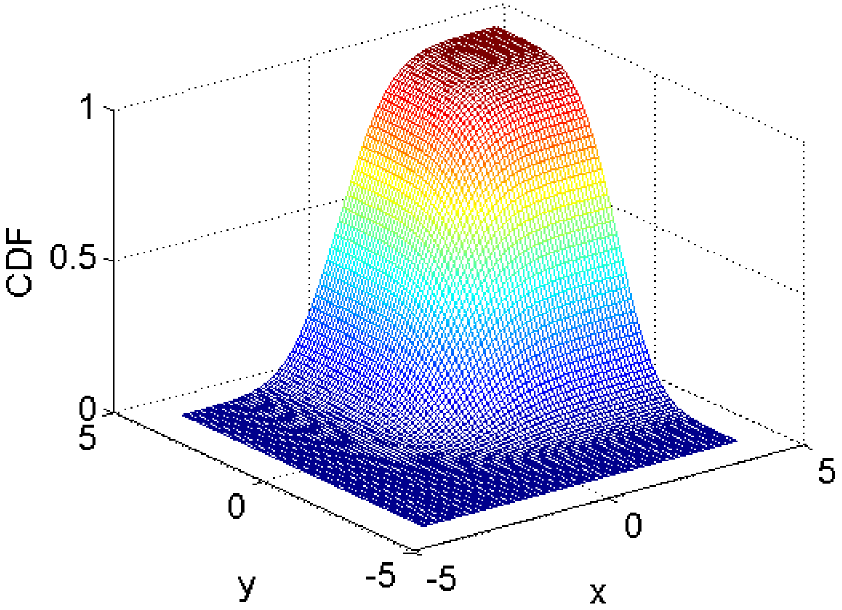
Applicable to more than two random variables

Joint Probability Distribution

- Example: bivariate Normal distribution



Joint probability density function



Joint cumulative distribution function

Marginal Distribution Function

- Marginal probability density function

$$pdf_x(t_x) = \int_{-\infty}^{+\infty} pdf_{x,y}(t_x, \tau_y) \cdot d\tau_y$$

$$pdf_y(t_y) = \int_{-\infty}^{+\infty} pdf_{x,y}(\tau_x, t_y) \cdot d\tau_x$$

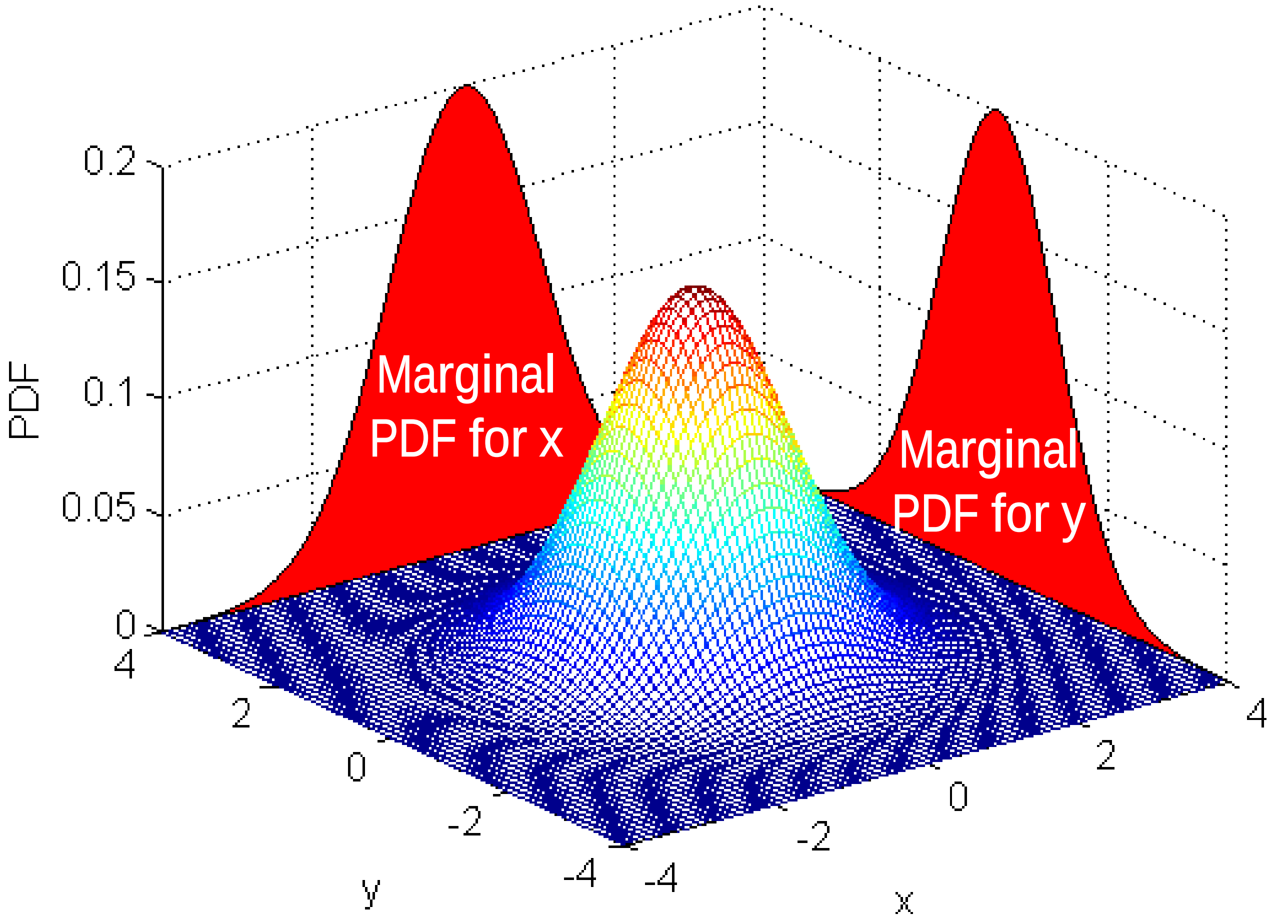
- Marginal cumulative distribution function

$$cdf_x(t_x) = P(x \le t_x, y \le +\infty) = \lim_{t_y \to +\infty} cdf_{x,y}(t_x, t_y)$$

$$cdf_y(t_y) = P(x \le +\infty, y \le t_y) = \lim_{t_x \to +\infty} cdf_{x,y}(t_x, t_y)$$

Marginal Distribution Function

- Example: bivariate Normal distribution



Covariance and Correlation

- Covariance

$$\boxed{COV[x,y] = E[(x - E[x]) \cdot (y - E[y])]}$$

  - If COV[x,y] = 0, then x and y are uncorrelated

- Covariance matrix

$$\boxed{\Sigma = \begin{bmatrix} COV[x,x] & COV[x,y] \\ COV[y,x] & COV[y,y] \end{bmatrix}}$$

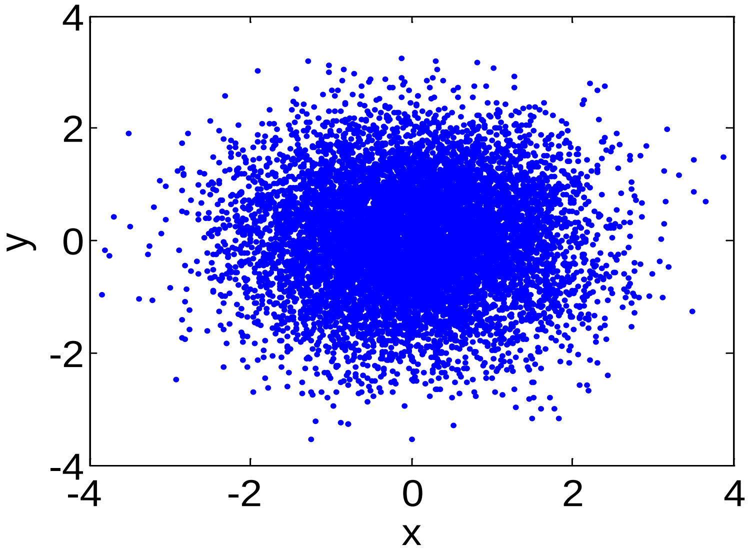
  - Σ is always symmetric
  - Diagonal components are corresponding to variance values
  - Σ is diagonal if x and y are uncorrelated

Covariance and Correlation

- Correlation (normalized covariance)

$$\boxed{COR[x,y] = \frac{COV[x,y]}{STD[x]\cdot STD[y]}}$$

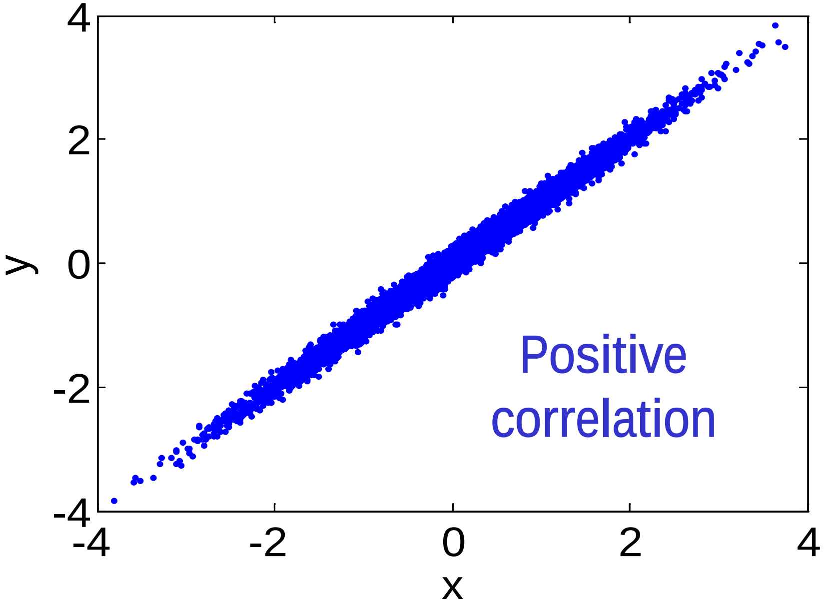
  - Correlation between two random variables can be visualized by scatter plot



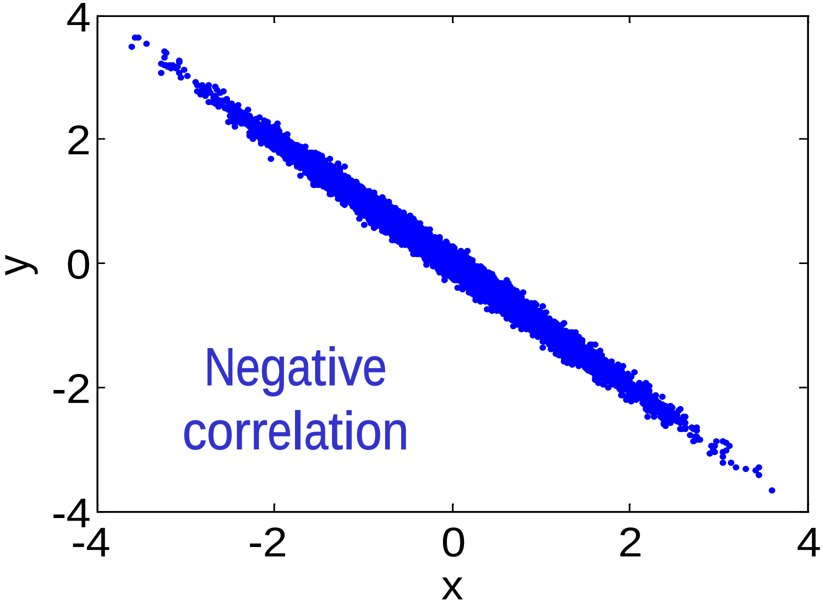
Zero correlation

Covariance and Correlation

- Example: correlated random variables



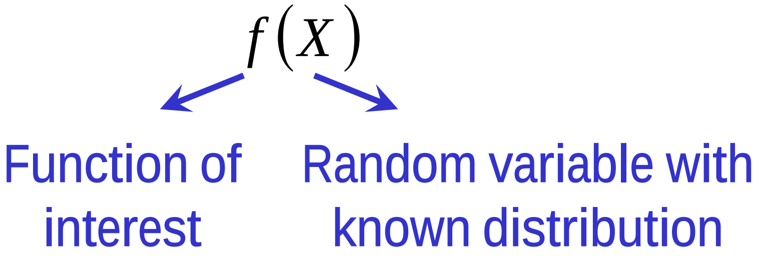
Positive correlation



Negative correlation

Monte Carlo Analysis

- Problem definition
  - Find probability distribution and/or moments of



$$f(X)$$

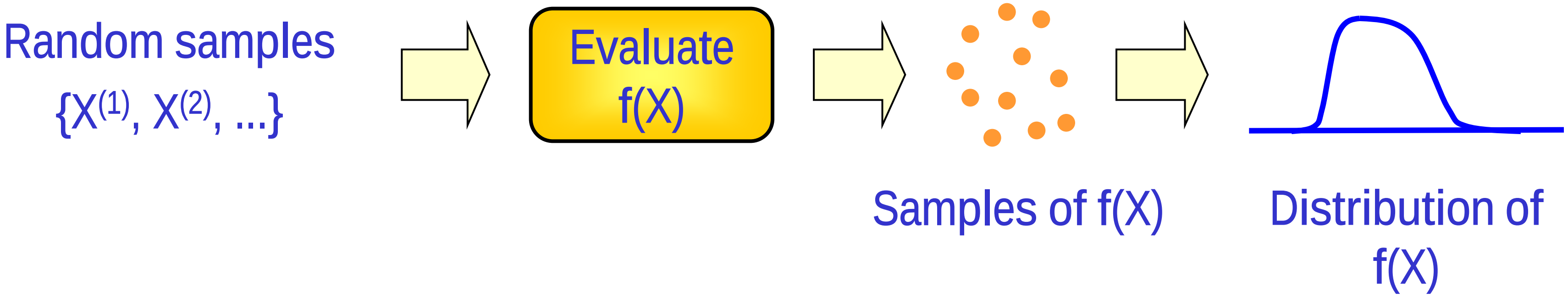
Function of interest

Random variable with known distribution

- In general, the distribution and/or moments of f cannot be calculated analytically, because
  - f(X) is nonlinear
  - f(X) may not have closed-form expression (we can only numerically calculate f for a given X value)

Monte Carlo Analysis

- Monte Carlo analysis for f(X)
  - Randomly select M samples for X
  - Evaluate function f(X) at each sampling point
  - Estimate distribution of f using these M samples



Random samples

$\{X^{(1)}, X^{(2)}, \ldots\}$

Evaluate f(X)

Samples of f(X)

Distribution of f(X)

Monte Carlo Analysis Example

- Example: estimate the probability distribution of

$$y = \exp(x)$$

  - x ~ N(0,1) (standard Normal distribution)

Monte Carlo Analysis Example

- Step 1: draw random samples for x

Samples	1	2	3	4	5	6	...
x	-0.4326	-1.6656	0.1253	0.2877	-1.1465	1.1909	...

M random samples for x

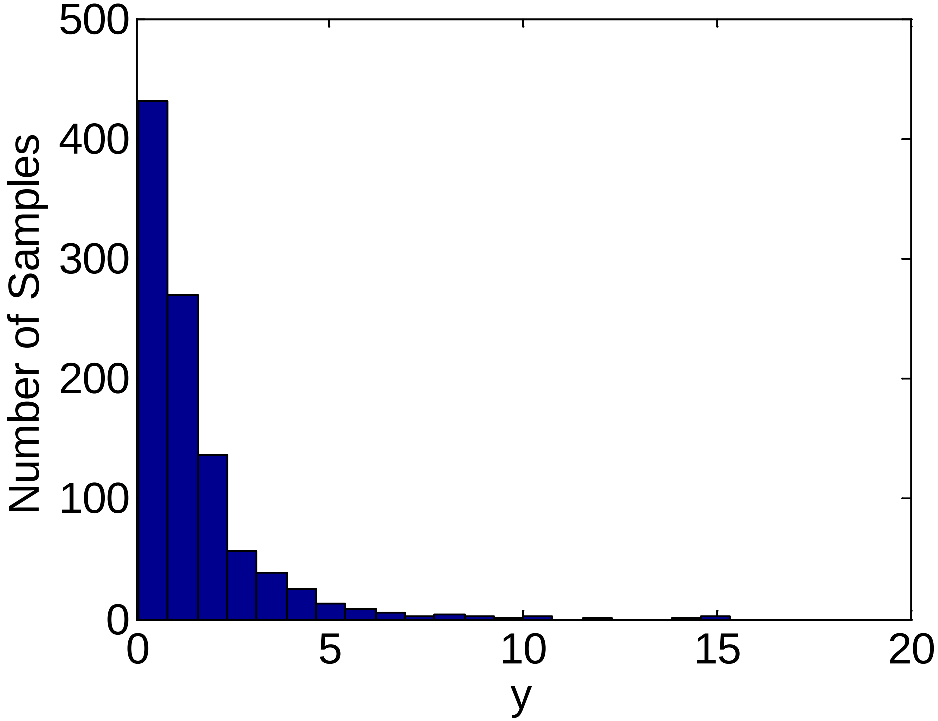
- Step 2: calculate y at each sampling point

Samples	1	2	3	4	5	6	...
y	0.6488	0.1891	1.1335	1.3333	0.3178	3.2901	...

M random samples for y

Monte Carlo Analysis Result

- Monte Carlo result is typically represented by a histogram
    - A big table of data is not intuitive



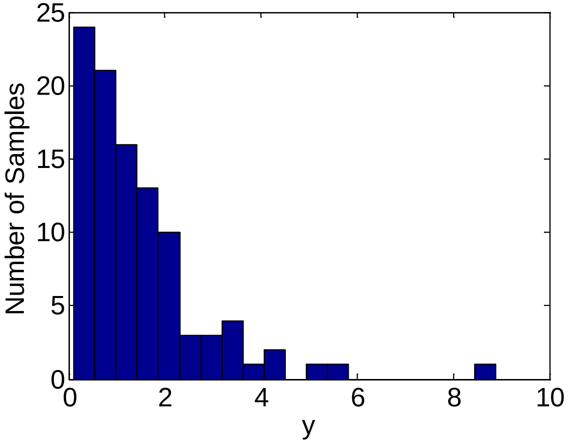
Histogram of y based on 1000 random samples

QUESTION: how accurate is Monte Carlo analysis?

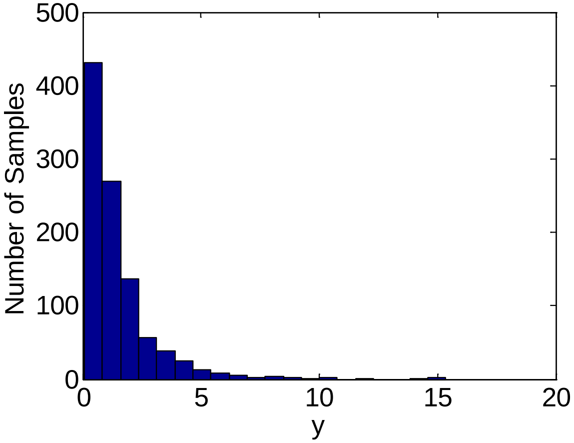
Monte Carlo Analysis Accuracy

- Monte Carlo analysis is not deterministic
  - We cannot get identical results when running MC twice
  - The analysis error is not deterministic

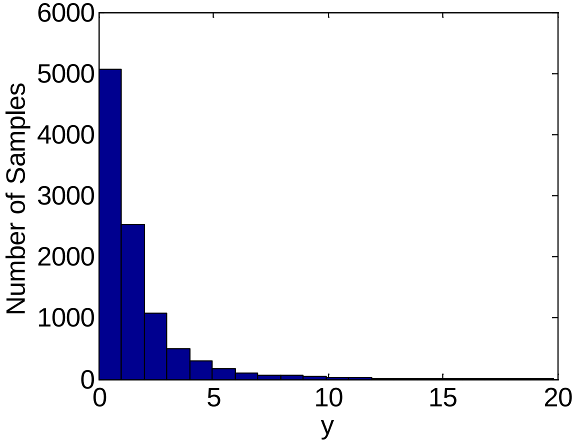
- Monte Carlo accuracy depends on the number of samples
  - Examples: histogram of y



100 samples



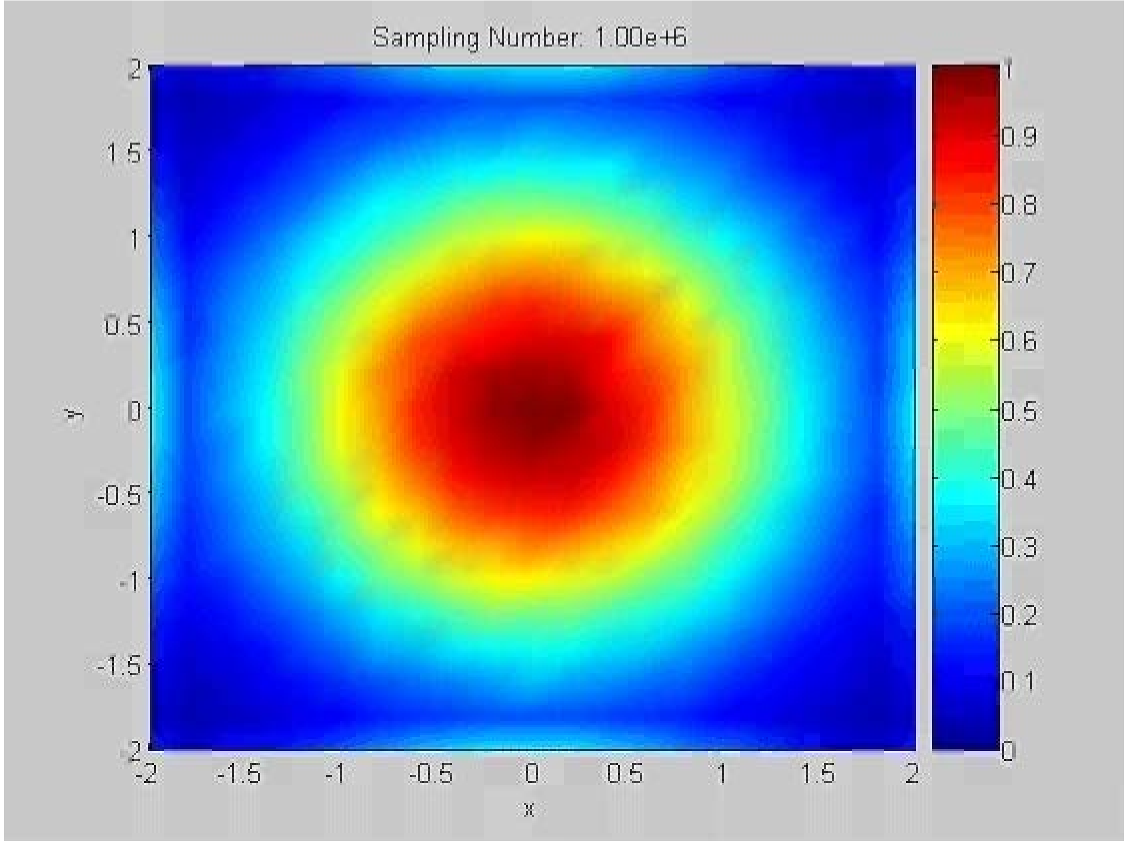
1000 samples



10000 samples

Monte Carlo Analysis Accuracy

- Example: bivariate Normal distribution
  - x and y are independent and jointly standard Normal



Monte Carlo Analysis Accuracy

- Statistical methods exist to analyze Monte Carlo accuracy

- Example: Monte Carlo accuracy analysis

$$x \sim N(0,1)$$ Standard Normal distribution

  - Estimate the mean value μ_x by Monte Carlo analysis
  - Our question: how accurate is the estimated μ_x (dependent on the number of Monte Carlo samples)?

Monte Carlo Analysis Accuracy

- Monte Carlo analysis for the mean value **μ_x**
  - Randomly draw M sampling points $\{x^{(1)}, x^{(2)}, \ldots, x^{(M)}\}$
  - Estimate μ_x by the following equation

$$\mu_x = \frac{x^{(1)} + x^{(2)} + \cdots + x^{(M)}}{M}$$

Called an estimator

- Assumptions in our accuracy analysis
  - Each $x^{(i)}$ is random and satisfies standard Normal distribution – it is randomly created for x ~ N(0,1)
  - All $x^{(i)}$'s are mutually independent – samples from a good random number generator should be independent

- μ_x is a function of $\{x^{(1)}, x^{(2)}, \ldots, x^{(M)}\}$, which is a random variable

Monte Carlo Analysis Accuracy

- Mean of μ_x

$$E\{\mu_x\} = E\left\{\frac{x^{(1)} + x^{(2)} + \cdots + x^{(M)}}{M}\right\} = \frac{E\{x^{(1)}\} + E\{x^{(2)}\} + \cdots + E\{x^{(M)}\}}{M} = 0$$

$E\{x^{(i)}\} = 0$

- Variance of μ_x

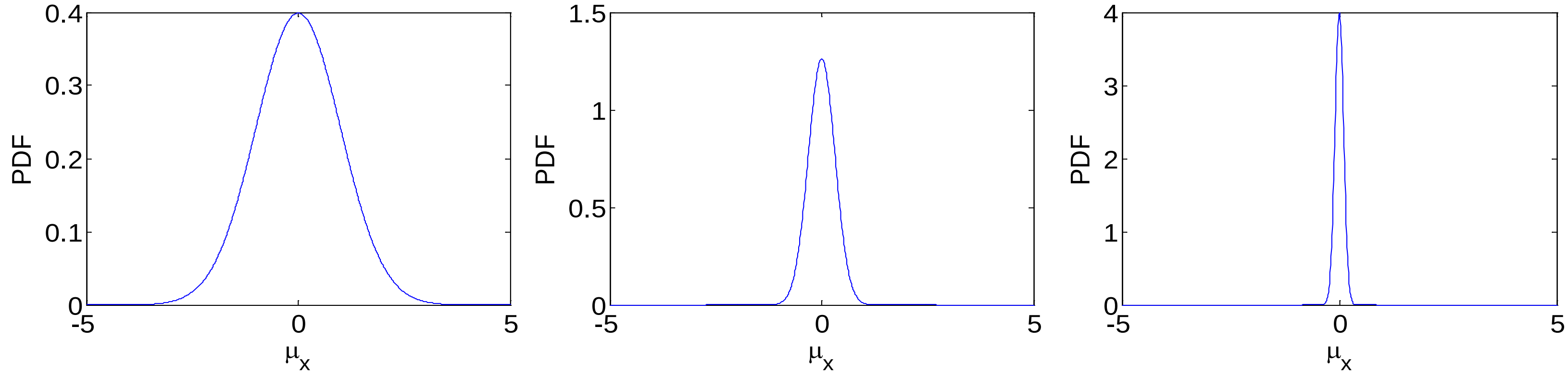
$$E\{\mu_x^2\} = E\left\{\left[\frac{x^{(1)} + x^{(2)} + \cdots + x^{(M)}}{M}\right]^2\right\} = \frac{E\{x^{(1)2}\} + E\{x^{(2)2}\} + \cdots + E\{x^{(M)2}\}}{M^2} = \frac{1}{M}$$

$x^{(i)}$'s are independent

$E\{x^{(i)2}\} = 1$

Monte Carlo Analysis Accuracy

- $E\{\mu_x\} = 0$
  - u_x is an unbiased estimator
  - Otherwise, if the estimator mean is not equal to the actual mean, it is called a biased estimator
- $E\{\mu_x^2\} = 1/M$
  - Variance decreases as M increases
  - Distributions of μ_x for different M values



1 sample 10 samples 100 samples

μ_x is a Normal distribution N(0, 1/M)

Monte Carlo Analysis Accuracy

- "Average" estimation accuracy is better when using larger M

- In this μ_x example
  - If we require that ±3 sigma of μ_x is within [-0.1, 0.1]

$$\frac{3}{\sqrt{M}} \le 0.1 \quad \Rightarrow \quad M \ge 900$$

  - If we require that ±3 sigma of μ_x is within [-0.01, 0.01]

$$\frac{3}{\sqrt{M}} \le 0.01 \quad \Rightarrow \quad M \ge 90000$$

Monte Carlo Analysis Accuracy

- Accuracy is improved by 10x if the number of samples is increased by 100x
- 1K ~ 10K sampling points are typically required to achieve reasonable accuracy
- However, even if you use 10K sampling points, an accurate result is not guaranteed!
  - Monte Carlo analysis is random, and you can be unlucky (e.g., going beyond ±3 sigma range)

Summary

- Monte Carlo analysis
  - Random variable
  - Probability distribution
  - Random sampling