



Statistical signal processing (5CTA0)

Lecture 2, part A

Lecturer: Simona Turco

Electrical Engineering, Signal Processing Systems group

Part 1: Random variables and Random Signals

Part 1

Random Variables and Random Signals

Lecture 1: Probability and Random Variables

Lecture 2: Random vectors, Random processes
and random signals

Lecture 3: Rational signal models

Part 1: Random variables and Random Signals

Part 1

Random Variables and Random Signals

Lecture 2: Random vectors, Random processes
and random signals

Part A: Pairs of random variables

Part B: Random vectors, random processes and
random signals

Pairs of random variables

Lecture 2, Part A

Outline

- Functions of random variables
- Pairs random variables
 - Joint probability distribution and expected values
- Conditional probability distributions
- Marginalization
- Central limit theorem

Functions of random variables

- What happens when I manipulate a random variable?
 - The probability distribution as well as the statistical properties may change

If X is a random variable and g is a function, then

DISCRETE RVs:
$$E[g(X)] = \sum_{x \in S_X} g(x) p_X(x)$$

CONTINUOUS RVs:
$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) p_X(x) dx$$

Functions of random variables

- What happens when I manipulate a random variable?
 - The probability distribution as well as the statistical properties may change

Example: $g(x) = x^m$  $E[X^m] = \sum_{x \in S_X} x^m p_X(x)$

Example: $g(x) = Y = aX + b$

a, b deterministic constants

$$E[g(x)] = E[aX + b] = aE[X] + b = a\mu_x + b$$

$$\text{var}[g(x)] = \text{var}[aX + b] = a^2 \text{var}[X]$$

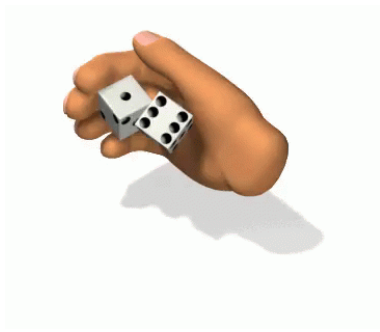


Pairs of random variables: introduction

- A stochastic process may involve multiple random variables
 - Multiple random variables associated to the same stochastic process are called **joint** random variables
 - The **probability model** of joint random variables contains properties of the individual random variables and the relationships among them

Pairs of random variables: example

Random process: rolling two dice

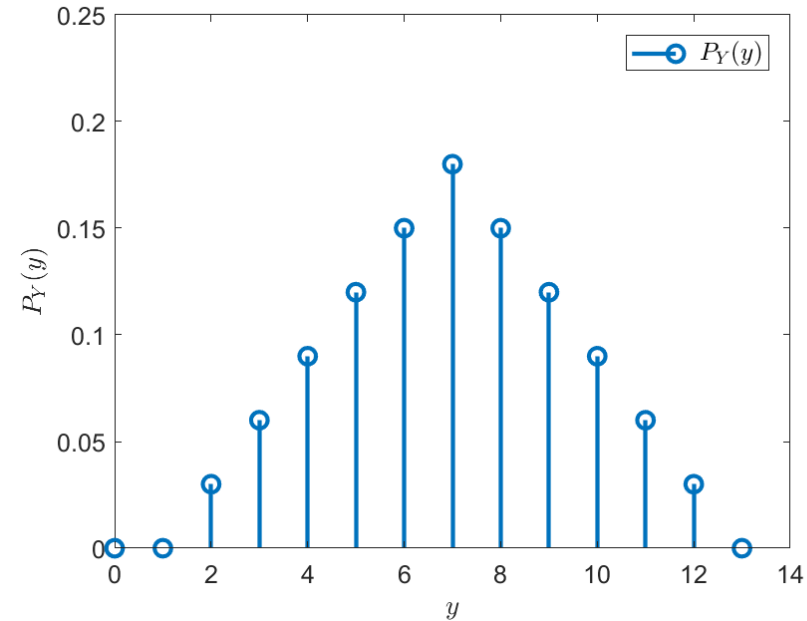


Y = [sum of the upward faces]

D_1 = number on first die

D_2 = number on second die

$$Y = D_1 + D_2$$



Pairs of random variables: CDF

The joint CDF for a pair of random variables X and Y is given by

$$P_{X,Y}(x, y) = \Pr[X \leq x, Y \leq y]$$

Properties

1. $0 \leq P_{X,Y}(x, y) \leq 1$
2. $P_{X,Y}(\infty, \infty) = 1$ and $P_{X,Y}(x, -\infty) = P_{X,Y}(-\infty, y) = 0$
3. $P_X(x) = P_{X,Y}(x, \infty)$ and $P_Y(y) = P_{X,Y}(\infty, y)$
4. CDF is non-decreasing, i.e., for $x \leq x_0$ and $y \leq y_0$ it holds $P_{X,Y}(x, y) \leq P_{X,Y}(x_0, y_0)$

Pairs of random variables: PMF and PDF

The joint PMF for a pair of **discrete** random variables X and Y is given by

$$p_{X,Y}(x, y) = \Pr[X = x, Y = y]$$

The joint PDF for a pair of **continuous** random variables X and Y is given by

$$p_{X,Y}(x, y) = \frac{\partial^2 P_{X,Y}(x, y)}{\partial x \partial y}$$

Statistical description of random vectors

The **covariance** is a measure of the dependence between two random variables

$$\text{Cov}[X, Y] = E[(X - \mu_x)(Y - \mu_y)] =$$

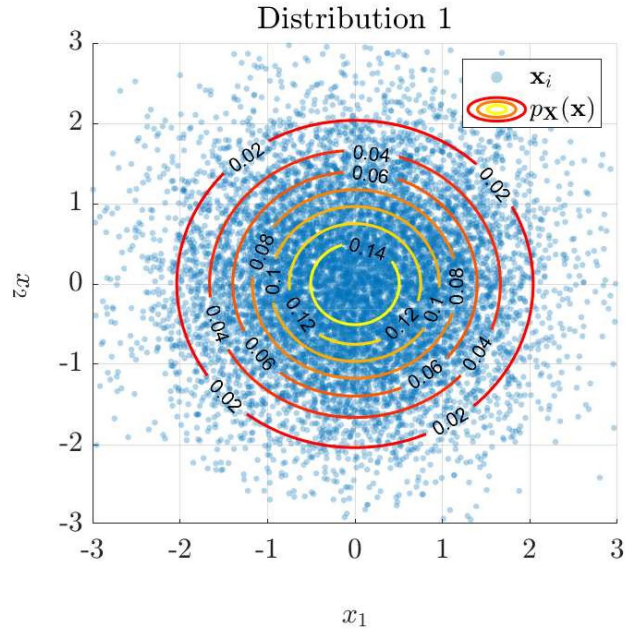
DISCRETE RV

$$= \sum_{y \in S_y} \sum_{x \in S_x} (x - \mu_x)(y - \mu_y) p_{X,Y}(x, y)$$

CONTINUOUS RV

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) p_{X,Y}(x, y) dx dy$$

Example

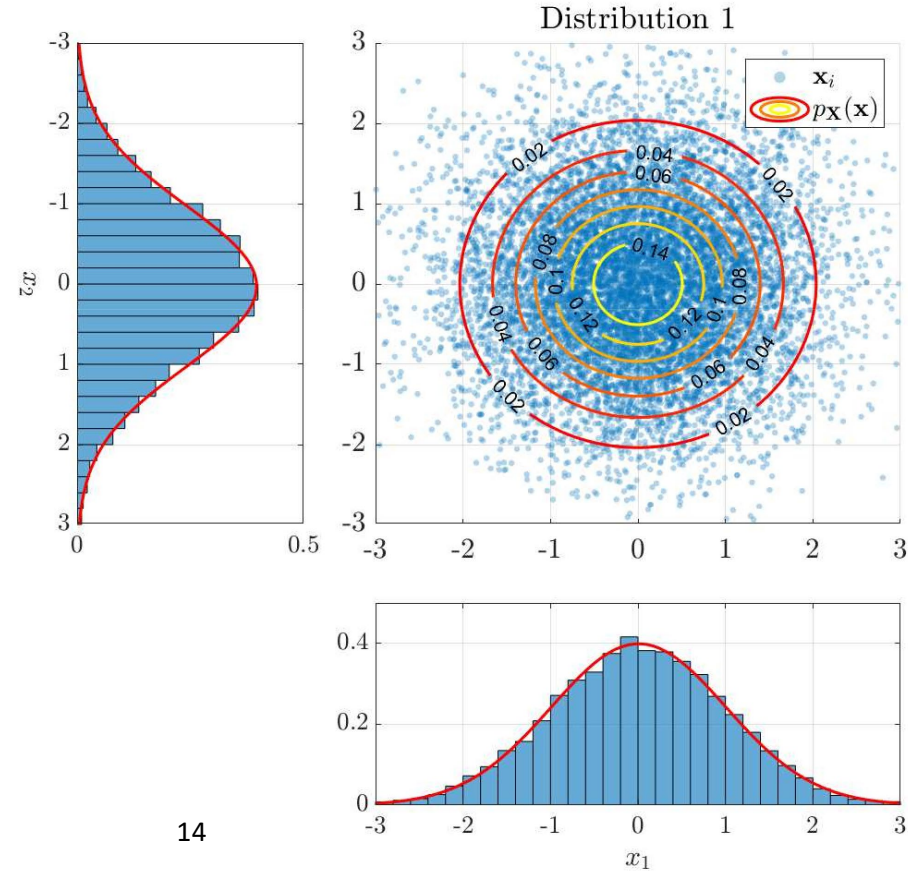


$$X_1 \sim N(0,1)$$

$$X_2 \sim N(0,1)$$

$$p^1_{X_1, X_2}(x_1, x_2)$$

Example

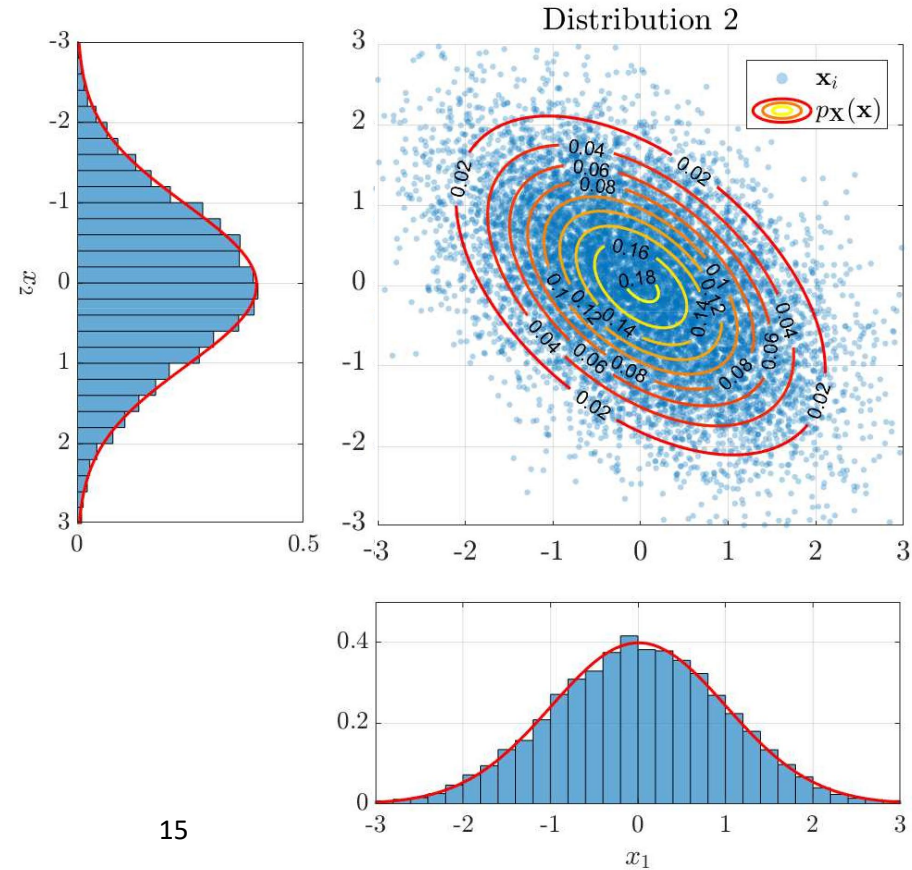


$$X_1 \sim N(0,1)$$

$$X_2 \sim N(0,1)$$

$$p^1_{X_1, X_2}(x_1, x_2)$$

Example

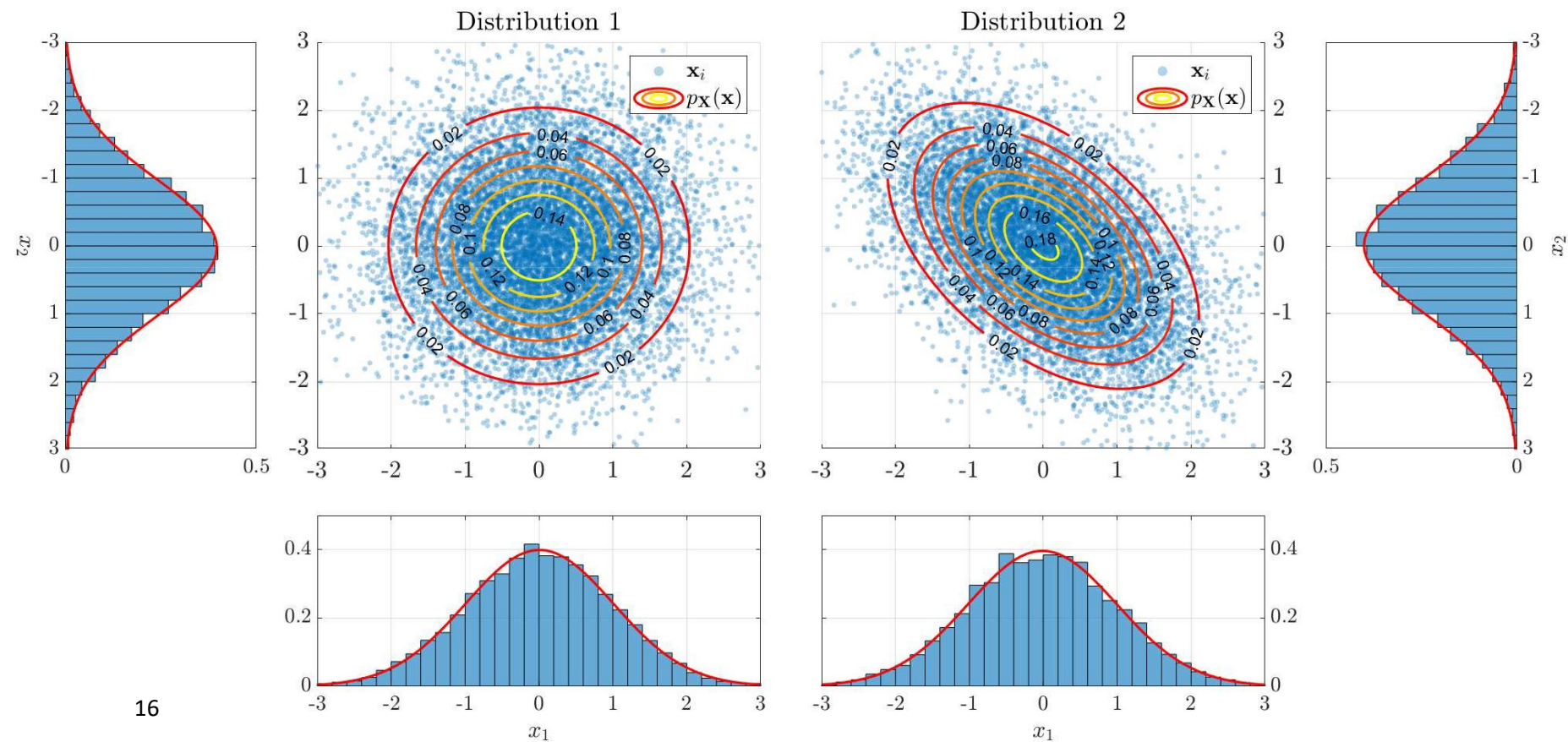


$$X_1 \sim N(0,1)$$

$$X_2 \sim N(0,1)$$

$$p^2_{X_1, X_2}(x_1, x_2)$$

Example





Covariance and correlation

The **covariance** is a measure of the dependence between two random variables

$$\begin{aligned}\text{Cov}[X, Y] &= E[(X - \mu_X)(Y - \mu_Y)] = \\ &= E[XY] - \mu_X \mu_Y = r_{X,Y} - \mu_X \mu_Y\end{aligned}$$

With $r_{x,y}$ called **correlation** $r_{X,Y} = E[XY]$

X, Y Real-valued random variables



Covariance and correlation (general)

The **covariance** is a measure of the dependence between two random variables

$$\begin{aligned}\text{Cov}[X, Y] &= E[(X - \mu_X)(Y - \mu_Y)^*] = \\ &= E[XY^*] - \mu_X \mu_Y^* = r_{X,Y} - \mu_X \mu_Y^*\end{aligned}$$

With $r_{x,y}$ called **correlation** $r_{X,Y} = E[XY^*]$

X, Y Complex-valued random variables

Uncorrelated and orthogonal random variables

Two random variables are uncorrelated if

$$\text{Cov}[X, Y] = 0$$

And orthogonal if

$$r_{X,Y} = 0$$

Correlation coefficient

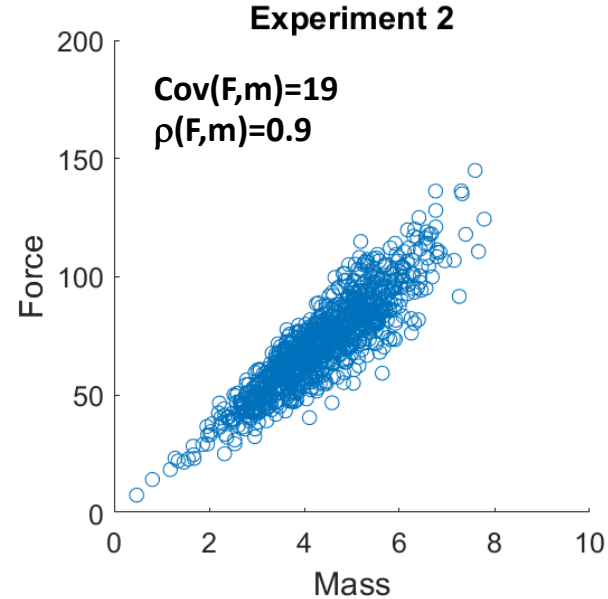
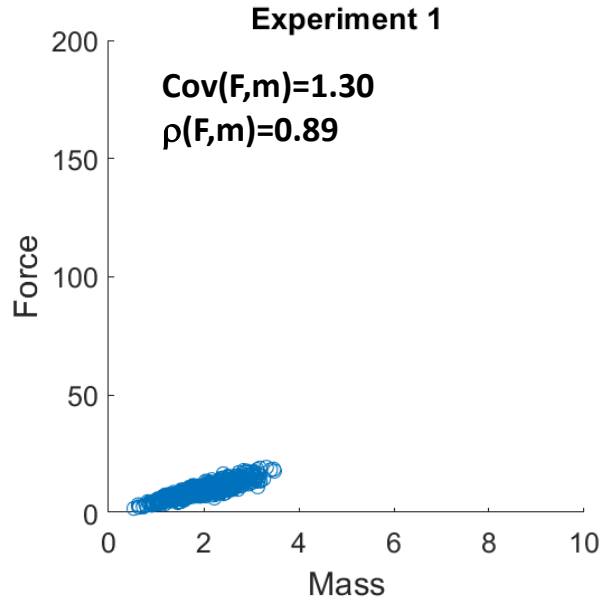
To express the relationship between two random variables independently from their variance, we use the **normalized correlation coefficient**

$$\rho_{X,Y} = \frac{\text{Cov}[X,Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}} = \frac{\text{Cov}[X,Y]}{\sigma_X \sigma_Y}.$$

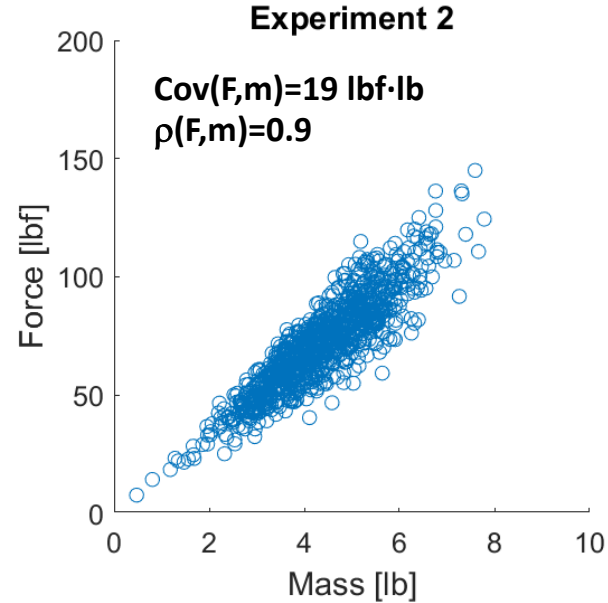
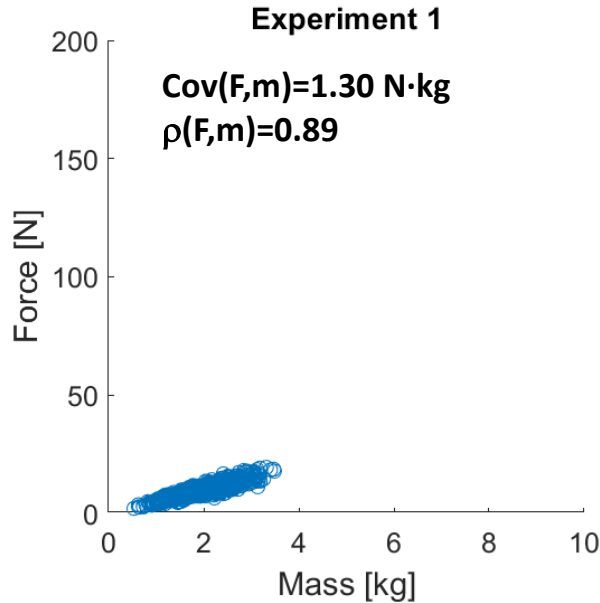
The correlation coefficient is bounded between -1 and 1

$$-1 \leq \rho_{X,Y} \leq 1.$$

Example: force measurement

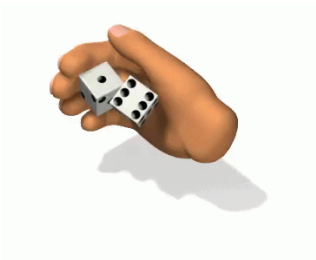


Example: force measurement



Joint probability distribution: example

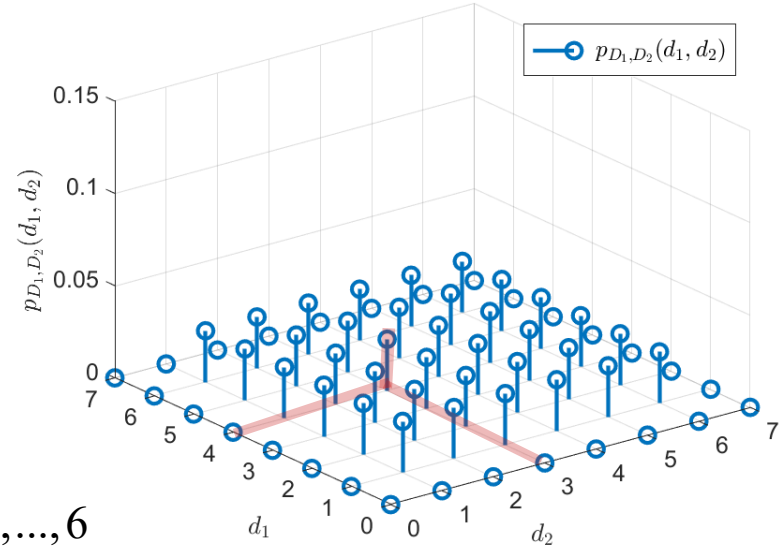
Random process: rolling two dice



D_1 = number on first die

D_2 = number on second die

$$p_{D_1, D_2}(d_1, d_2) = \begin{cases} 1/36 & \text{for } d_1 = 1, \dots, 6; d_2 = 1, \dots, 6 \\ 0 & \text{elsewhere} \end{cases}$$





Conditional probability distributions

Given the a priori knowledge that an event set B , with probability $\Pr[B]$, has occurred we can define

Conditional probability function:
$$p_{X,Y}(x, y|B) = \begin{cases} \frac{p_{X,Y}(x, y)}{\Pr[B]} & (x, y) \in B \\ 0 & \text{elsewhere} \end{cases}$$

The event set B becomes the sample space!

Conditional probability distributions: example

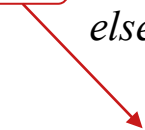
Random process: rolling two dice

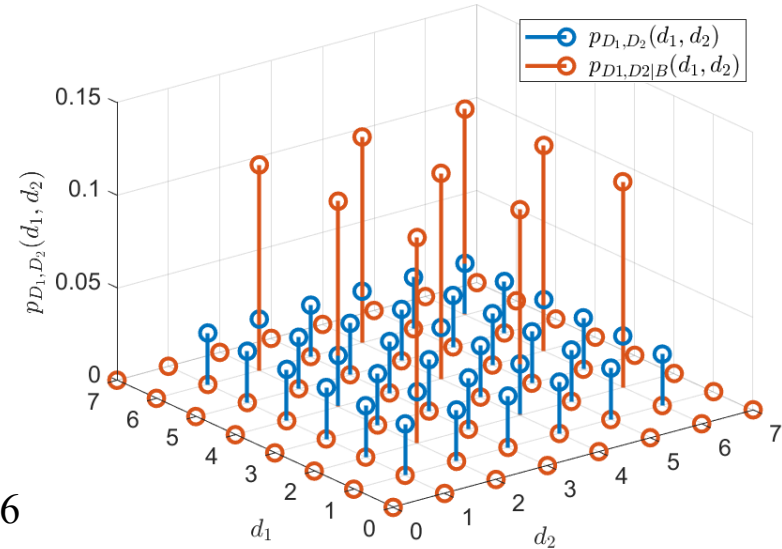
D_1 = number on first die

D_2 = number on second die

$$p_{D_1, D_2}(d_1, d_2) = 1/36 \text{ for } d_1 = 1, \dots, 6; d_2 = 1, \dots, 6$$

$$B = \{d_1 \text{ even}, d_2 \text{ even}\}$$

$$p_{D_1, D_2|B}(d_1, d_2) = \begin{cases} 1/36 & \text{for } d_1 = 2, 4, 6; d_2 = 2, 4, 6 \\ \boxed{1/4} & \\ 0 & \text{elsewhere} \end{cases}$$




d_1 and d_2 are independent: $\Pr[B] = \Pr[\{d_1 \text{ even}\}] \Pr[\{d_2 \text{ even}\}] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$



Conditional probability distributions

Given that Y is observed, we can define the following conditional probability distributions

Conditional PMF:
$$p_{X|Y}(x | y) = \Pr[X = x | Y = y] = \frac{\Pr[X = x, Y = y]}{\Pr[Y = y]} = \frac{p_{X,Y}(x, y)}{p_Y(y)}.$$

Conditional PDF:
$$p_{X|Y}(x | y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

Marginalization

To obtain the probability function of a random variable of interest, we “marginalize” over all other random variables

Discrete RVs

$$p_X(x) = \sum_{y \in \mathcal{S}_Y} p_{X,Y}(x, y)$$

Continuous RVs

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y) dy.$$

Joint random variables: example

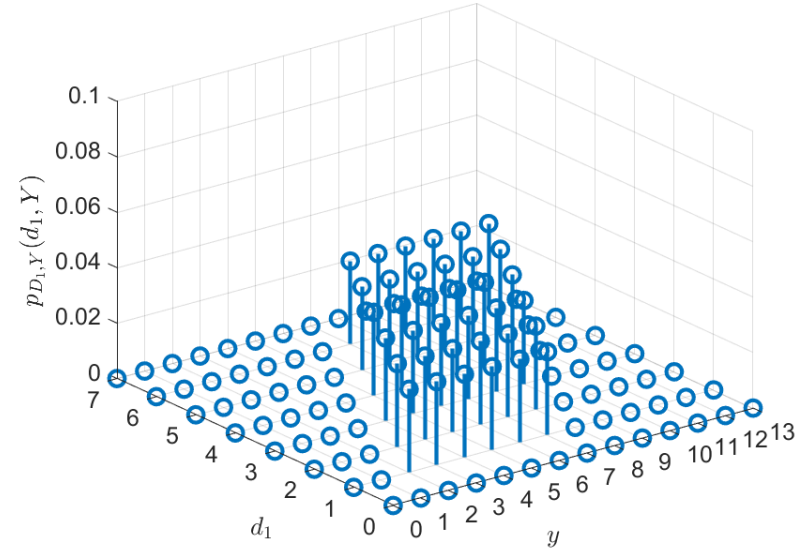
Random process: rolling two dice

D_1 = number on first die

D_2 = number on second die

$$Y = D_1 + D_2$$

$$p_{D_1, Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; y = 1 + d_1, \dots, 6 + d_1$$



Joint random variables: example

Random process: rolling two dice

D_1 = number on first die

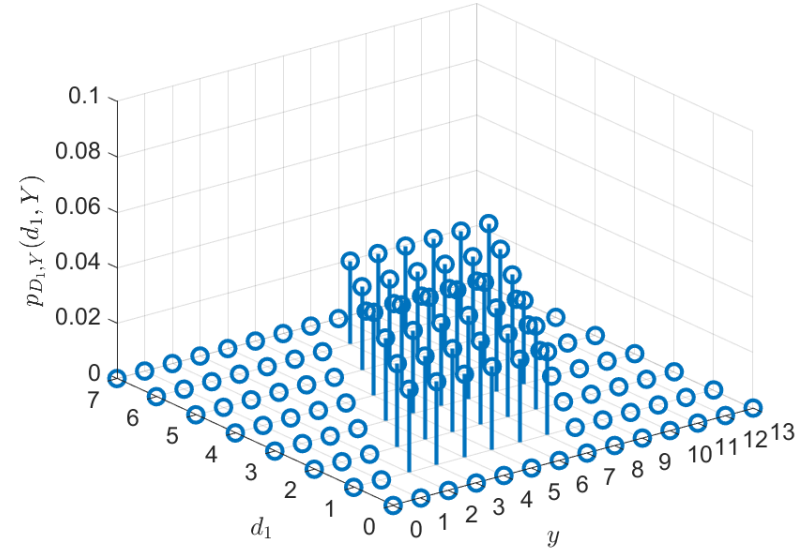
D_2 = number on second die

$$Y = D_1 + D_2$$

$$p_{D_1,Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; y = 1 + d_1, \dots, 6 + d_1$$

$$D_1 = d_1$$

$$p_{Y|D_1}(y | d_1) = \frac{p_{D_1,Y}(d_1, y)}{p_{D_1}(d_1)}$$



Marginalization: example

Random process: rolling two dice

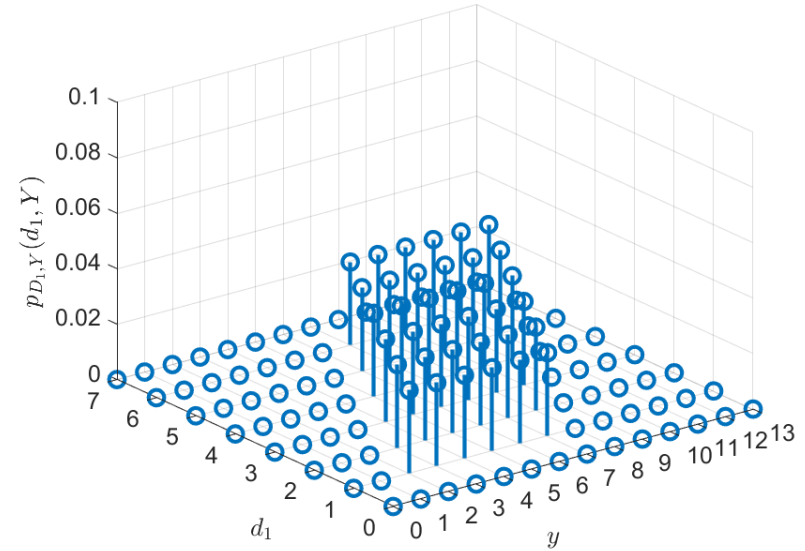
D_1 = number on first die

D_2 = number on second die

$$Y = D_1 + D_2$$

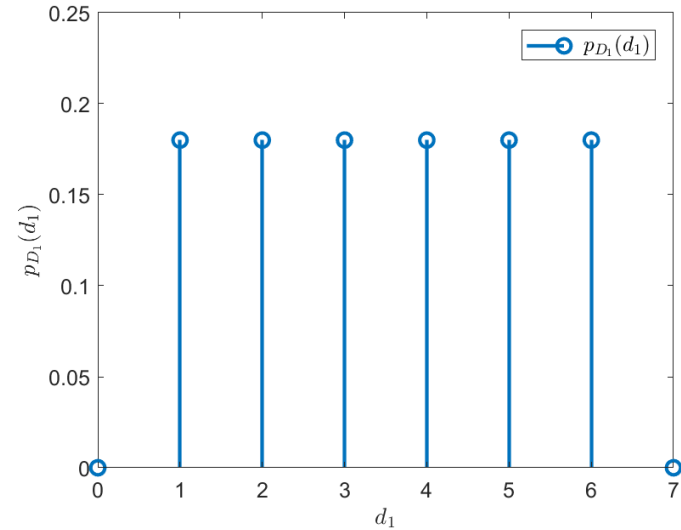
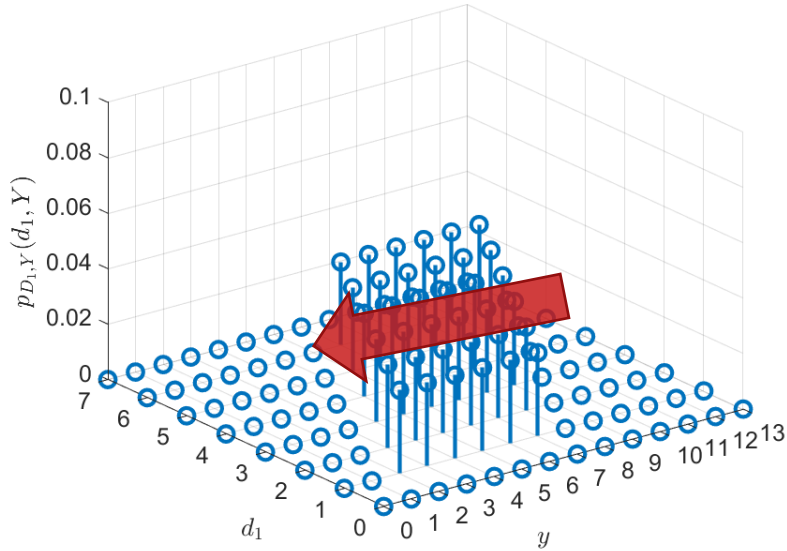
$$p_{D_1, Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; y = 1 + d_1, \dots, 6 + d_1$$

$$p_{D_1}(d_1) = \sum_{y \in S_Y} p_{D_1, Y}(d_1, y)$$



Marginalization: example

$$p_{D_1}(d_1) = \sum_{y \in S_Y} p_{D_1, Y}(d_1, y)$$



Marginalization: example

Random process: rolling two dice

D_1 = number on first die

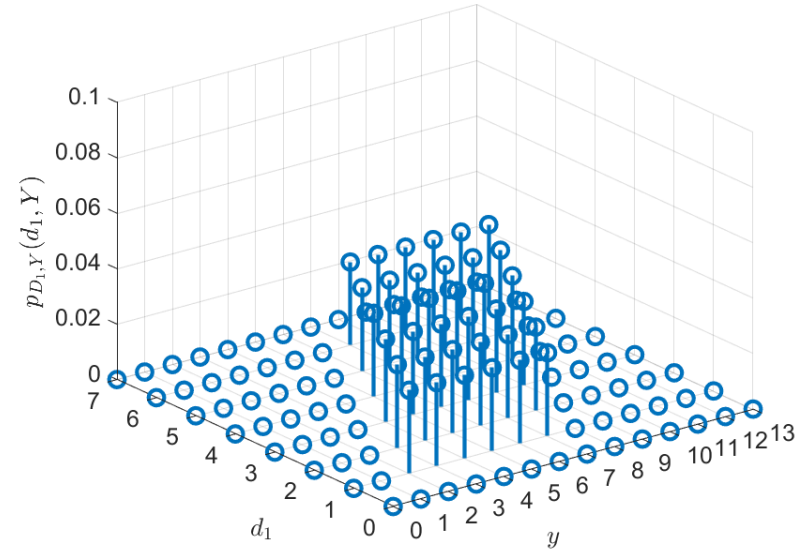
D_2 = number on second die

$$Y = D_1 + D_2$$

$$p_{D_1,Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; y = 1 + d_1, \dots, 6 + d_1$$

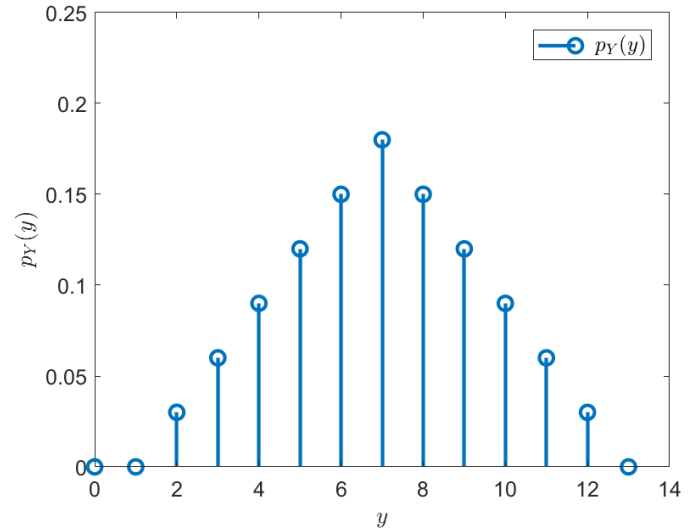
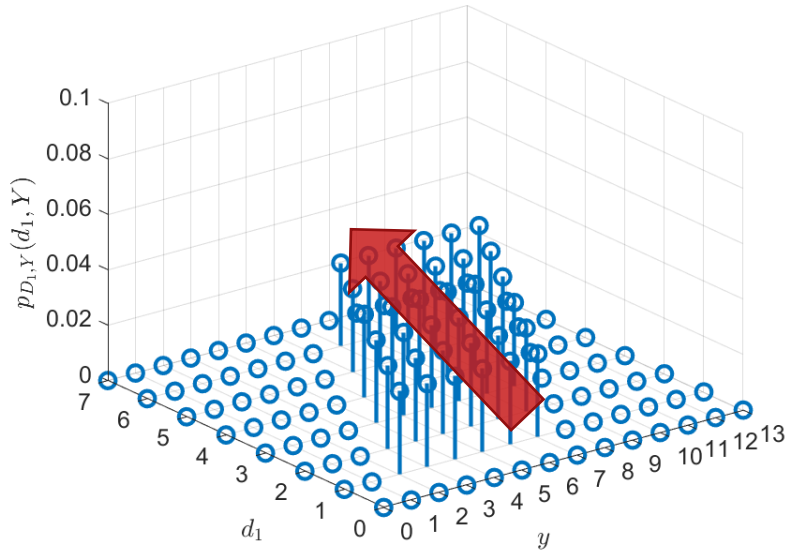
$$p_{D_1}(d_1) = \sum_{y \in S_Y} p_{D_1,Y}(d_1, y)$$

$$p_Y(y) = \sum_{d_1 \in S_{D_1}} p_{D_1,Y}(d_1, y)$$



Marginalization: example

$$p_Y(y) = \sum_{d_1 \in S_{D_1}} p_{D_1, Y}(d_1, y)$$



Marginalization: example

Random process: rolling two dice

D_1 = number on first die

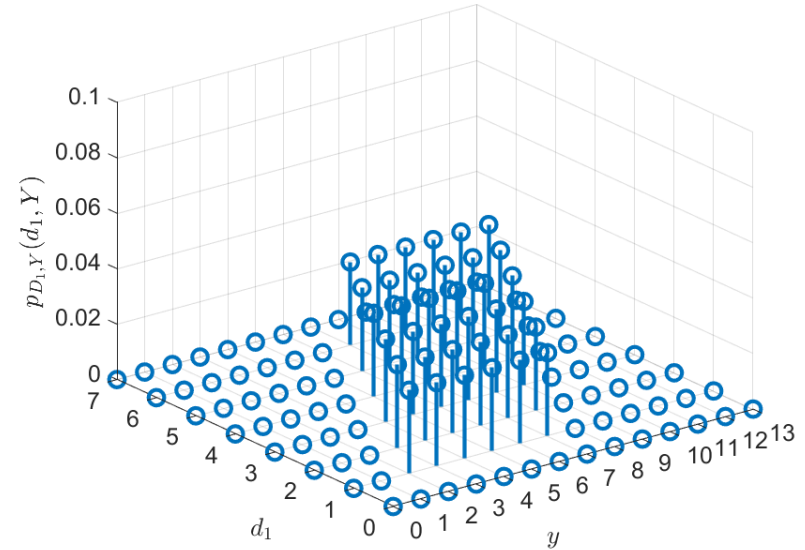
D_2 = number on second die

$$Y = D_1 + D_2$$

$$p_{D_1, Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; y = 1 + d_1, \dots, 6 + d_1$$

$$p_{D_1}(d_1) = \sum_{y \in S_Y} p_{D_1, Y}(d_1, y)$$

$$p_Y(y) = \sum_{d_1 \in S_{D_1}} p_{D_1, Y}(d_1, y)$$



Conditional PMF: example

Random process: rolling two dice

D_1 = number on first die

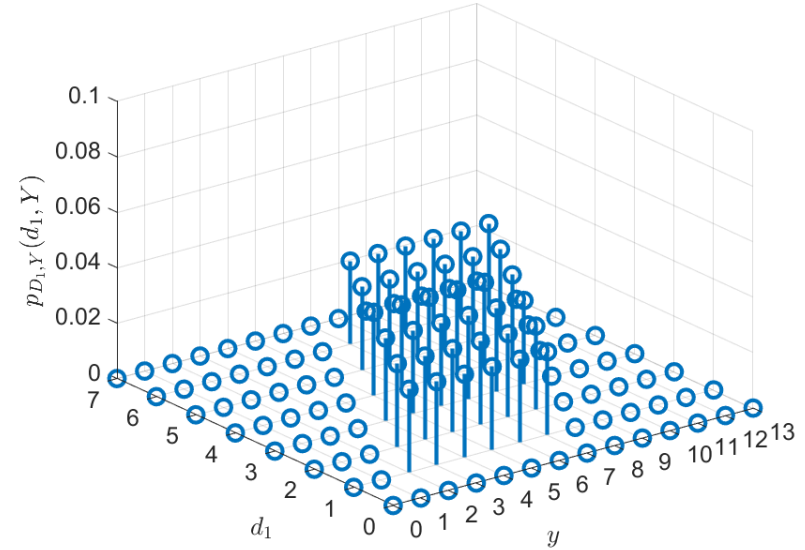
D_2 = number on second die

$$Y = D_1 + D_2$$

$$p_{D_1,Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; y = 1 + d_1, \dots, 6 + d_1$$

$$D_1 = d_1$$

$$p_{Y|D_1}(y | d_1) = \frac{p_{D_1,Y}(d_1, y)}{p_{D_1}(d_1)}$$

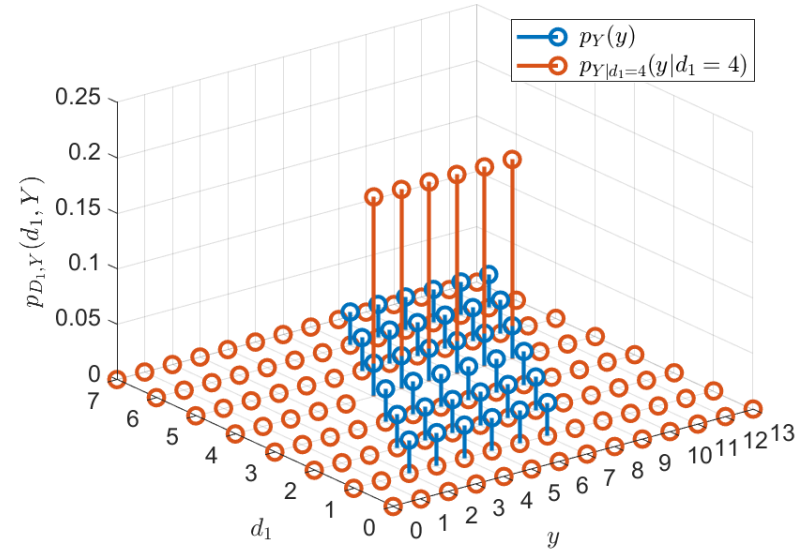


Conditional PMF: example

$$p_{D_1,Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; y = 1 + d_1, \dots, 6 + d_1$$

$$D_1 = d_1$$

$$p_{Y|D_1}(y | d_1 = 4) = \frac{p_{D_1,Y}(d_1, y)}{p_{D_1}(d_1 = 4)}$$



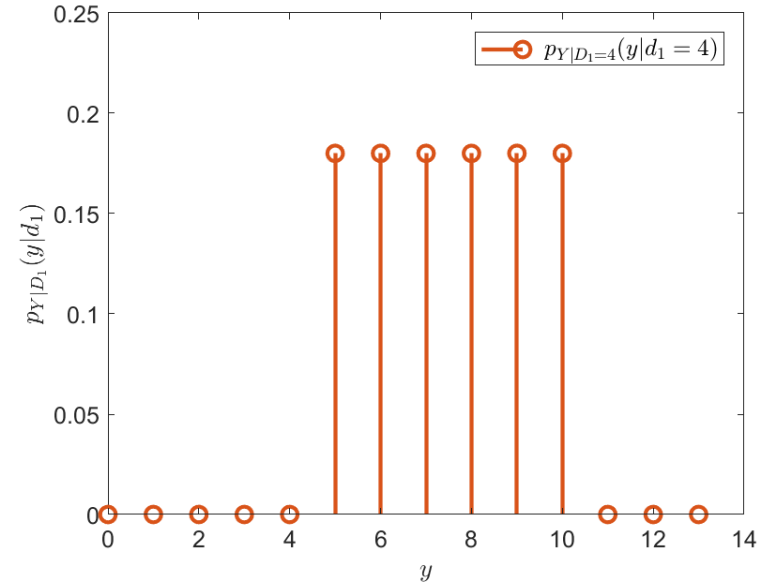
Conditional PMF : example

$$p_{D_1,Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; \text{ } y = 1 + d_1, \dots, 6 + d_1$$

$$D_1 = d_1$$

$$p_{Y|D_1}(y | d_1 = 4) = \frac{p_{D_1,Y}(d_1, y)}{p_{D_1}(d_1 = 4)}$$

$$p_{Y|D_1}(y | d_1) = \begin{cases} \frac{p_{D_1,Y}(d_1, y)}{p_{D_1}(d_1)} = \frac{1/36}{1/6} \text{ for } y = 1 + d_1, \dots, 6 + d_1 \\ 0 & \text{elsewhere} \end{cases}$$



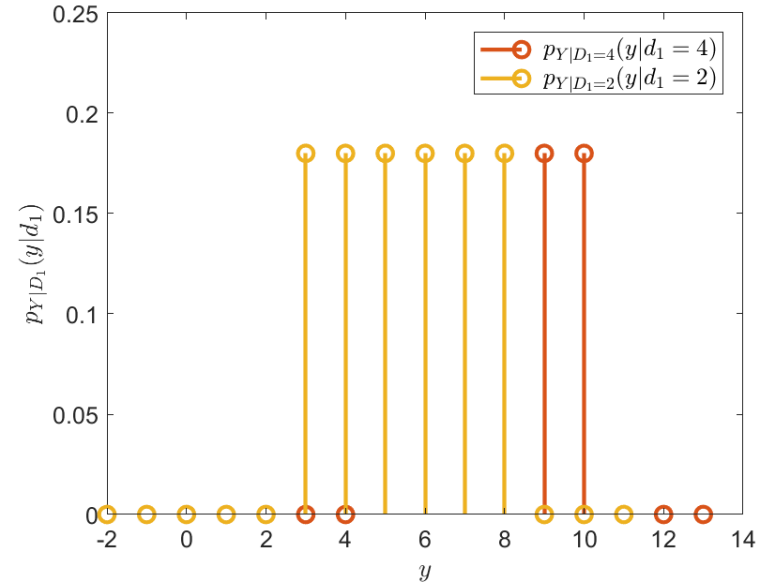
Conditional PMF: example

$$p_{D_1,Y}(d_1, y) = 1/36 \text{ for } d_1 = 1, \dots, 6; \text{ } y = 1 + d_1, \dots, 6 + d_1$$

$$D_1 = d_1$$

$$p_{Y|D_1}(y | d_1 = 4) = \frac{p_{D_1,Y}(d_1, y)}{p_{D_1}(d_1 = 4)}$$

$$p_{Y|D_1}(y | d_1) = \begin{cases} \frac{p_{D_1,Y}(d_1, y)}{p_{D_1}(d_1)} = \frac{1/36}{1/6} \text{ for } y = 1 + d_1, \dots, 6 + d_1 \\ 0 & \text{elsewhere} \end{cases}$$



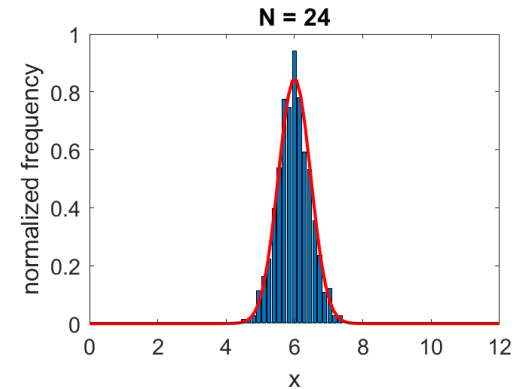
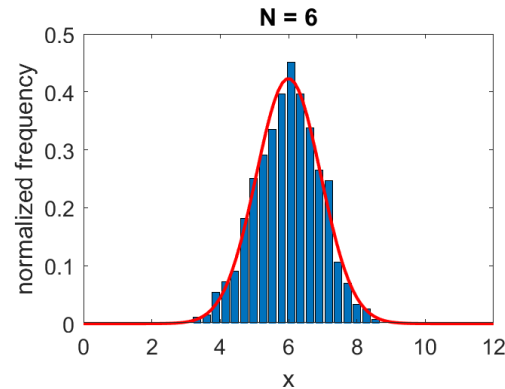
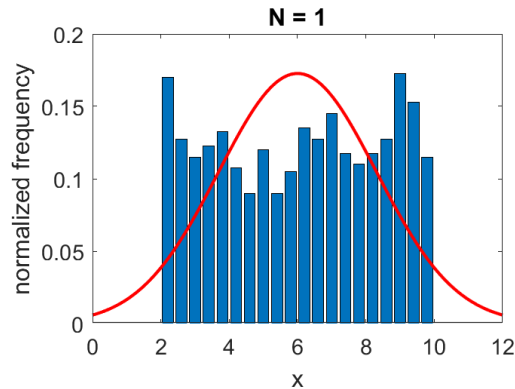
Central limit theorem (classic)

Let X_1, X_2, \dots, X_N be a set of N **independent identically-distributed (i.i.d)** random variables and each X_i has an arbitrary probability distribution $p(x_1, x_2, \dots, x_n)$ with finite mean $\mu_i = \mu$ and finite standard deviation $\sigma_i = \sigma$.

If the sample size N is “**sufficiently large**”, then the CDF of the sum converges to a Gaussian CDF

Examples

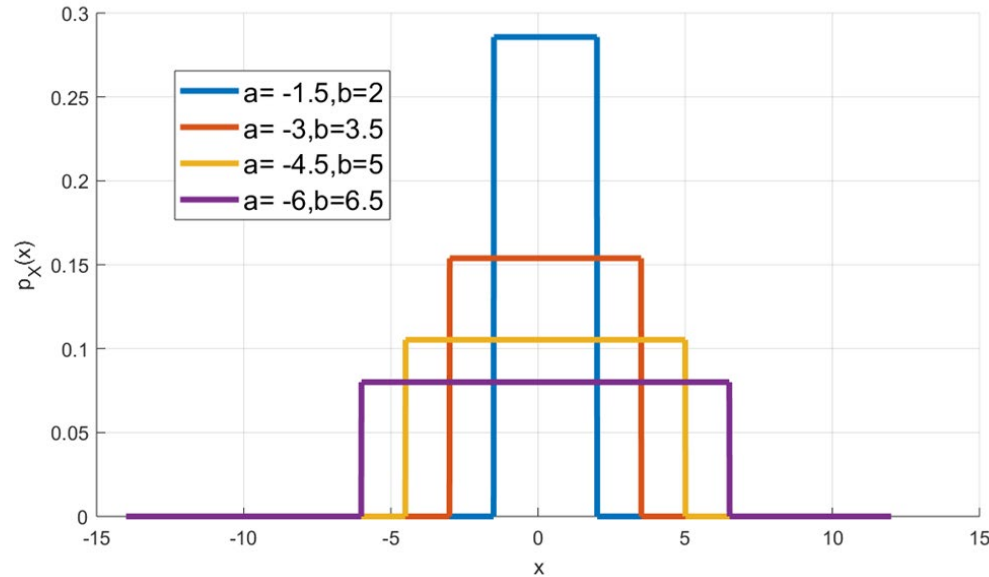
The normalized sum of a **sufficient** number of *i.i.d* random variables tends to a Gaussian distribution.



$X \sim \text{Uniform}(2, 10)$

Uniform distribution

- Continuous or discrete **uniform distribution**: used to model equally likely events

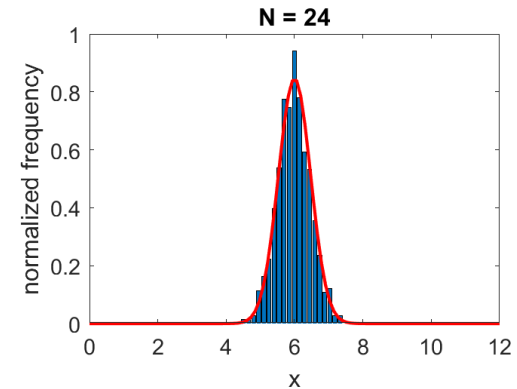
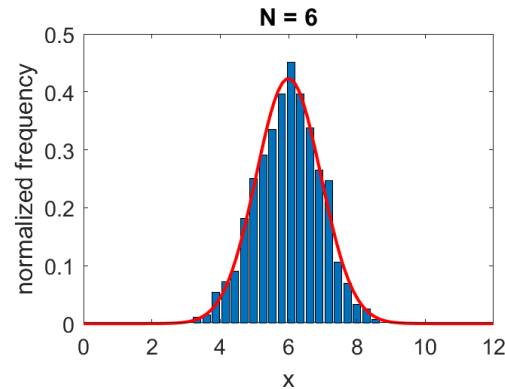
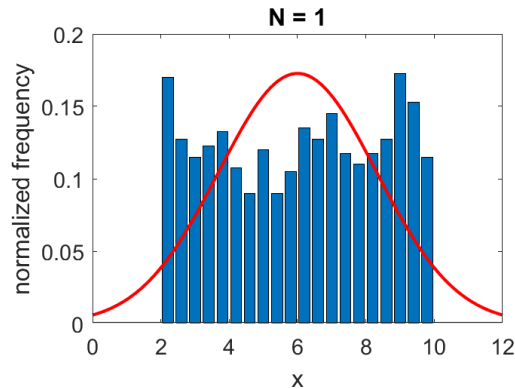


$$X \sim \text{Uniform}(a, b)$$

$$p_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Examples

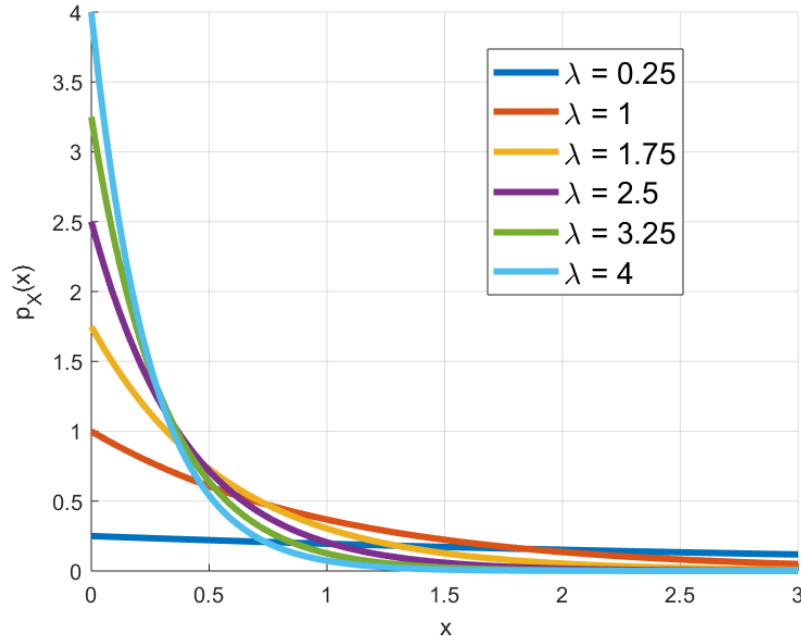
The normalized sum of a **sufficient** number of *i.i.d* random variables tends to a Gaussian distribution.



$X \sim \text{Uniform}(2,10)$

Exponential distribution

- Continuous **exponential distribution**: used to model the amount of time until some specific event occurs

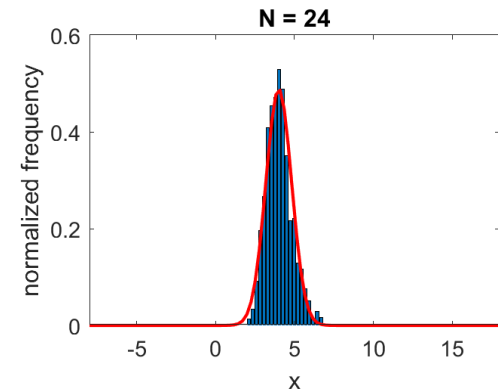
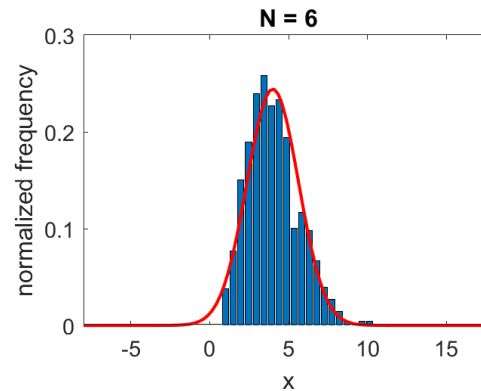
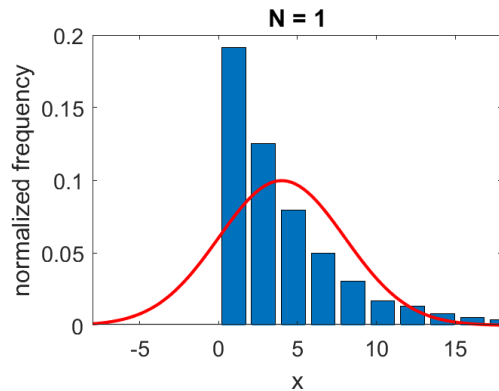


$$X \sim \text{Exponential}(\lambda)$$

$$p_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad \lambda > 0$$

Example

The normalized sum of a **sufficient** number of i.i.d random variables tends to a Gaussian distribution.



$X \sim \text{Exponential}(4)$

Central limit theorem (Lyapunov)

Let X_1, X_2, \dots, X_N be a set of N **independent** and each X_i has an **arbitrary probability distribution** $p(x_1, x_2, \dots, x_N)$ with finite mean μ_i and finite standard deviation σ_i .

If the sample size N is **sufficiently large**, and the **Lyapunov condition is satisfied**, then the CDF of the sum converges to a Gaussian CDF

$$X = \frac{1}{N} \sum_{i=1}^N X_i \quad P_X \xrightarrow{N} \mathcal{N}(\mu_X, \sigma_X) \quad \text{with:} \quad \begin{aligned} \mu_X &= \frac{1}{N} \sum_{i=1}^N \mu_i \\ \sigma_X^2 &= \frac{1}{N^2} \sum_{i=1}^N \sigma_i^2 \end{aligned}$$

Lyapunov condition

Given X_1, X_2, \dots, X_N a set of N **independent** with finite mean μ_i and finite standard deviation σ_i , and defining:

$$s_n^2 = \sum_{i=1}^n \sigma_i^2$$

For some $\delta > 0$, the **Lyapunov condition** is satisfied when

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n E \left[|X_i - \mu_i|^{2+\delta} \right] = 0$$

Lyapunov condition

Given X_1, X_2, \dots, X_N a set of N **independent** with finite mean μ_i and finite standard deviation σ_i , and defining:

$$S_n^2 = \sum_{i=1}^n$$

The Lyapunov condition puts a limit on the rate of growth of the moments

For some $\delta > 0$

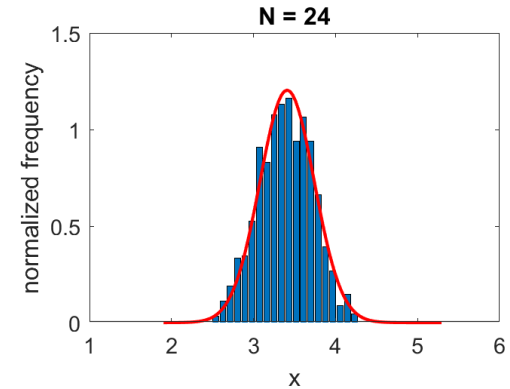
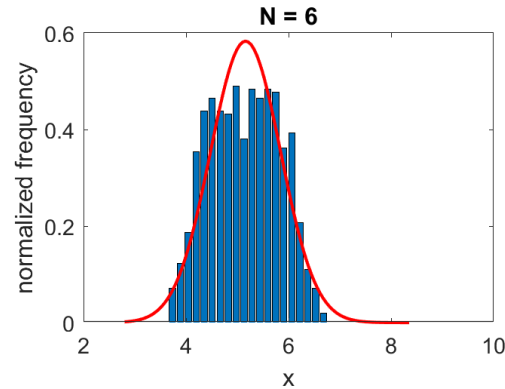
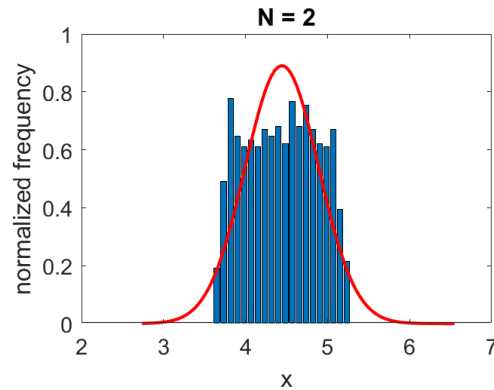
$$\lim_{n \rightarrow \infty} \frac{1}{S_n^{2+\delta}} \sum_{i=1}^n E[|X_i - \mu_i|^{2+\delta}] = 0$$

In simple words..

- All CLT formulations essentially state that the sum of multiple random variables converges to a gaussian distribution, provided that:
 - The number of random variables is “sufficiently large”
 - They are not too correlated
 - The variables are not too large (Lyapunov quantifies this by imposing a condition on the moments)

“Sufficiently large” number condition

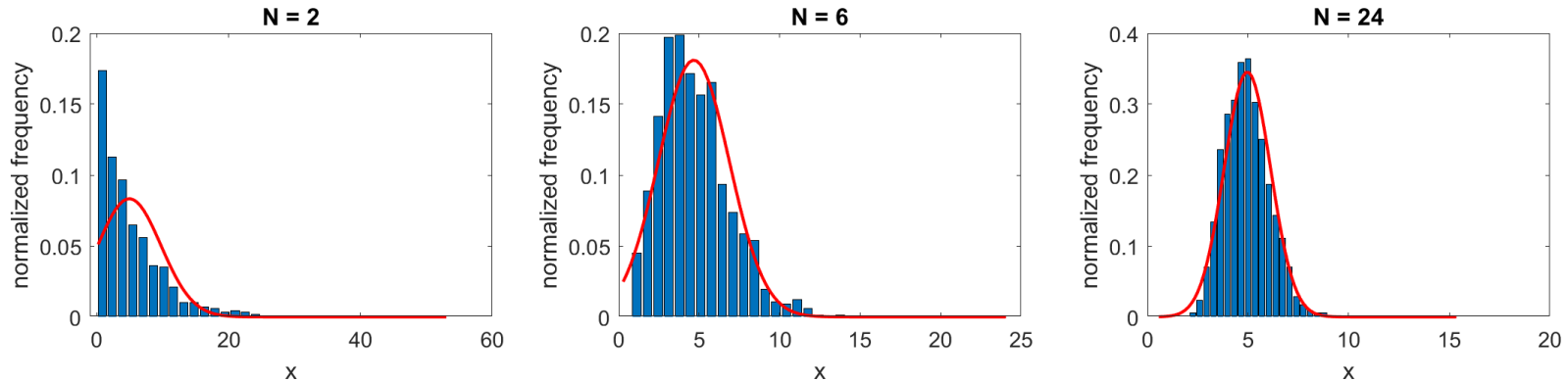
The normalized sum of a **sufficient** number of **independent** random variables tends to a Gaussian distribution.



$$X \sim \text{Uniform}(a_i, b_i)$$

“Sufficiently large” number condition

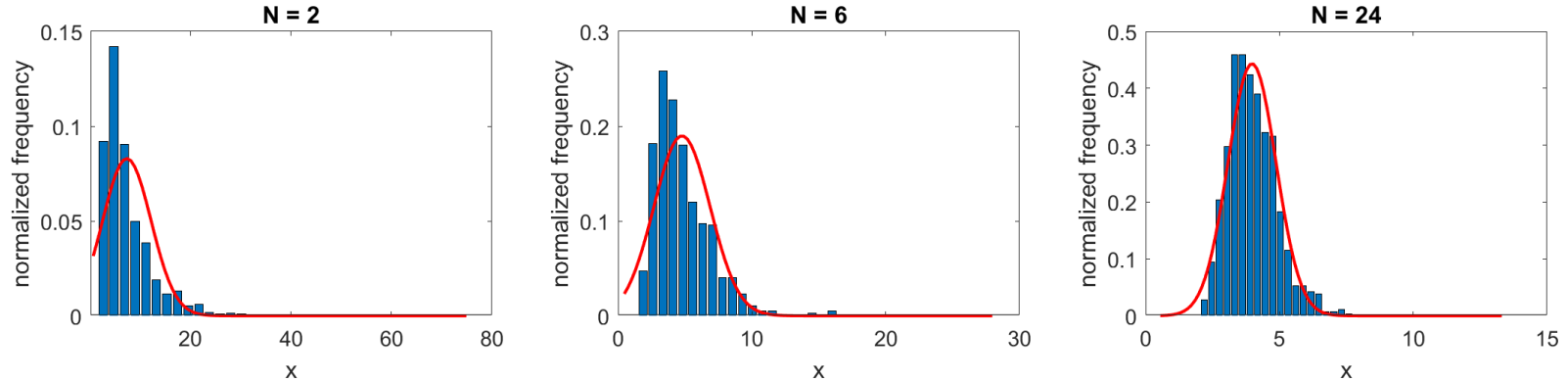
The normalized sum of a **sufficient** number of **independent** random variables tends to a Gaussian distribution.



$X \sim \text{Exponential}(\lambda_i)$

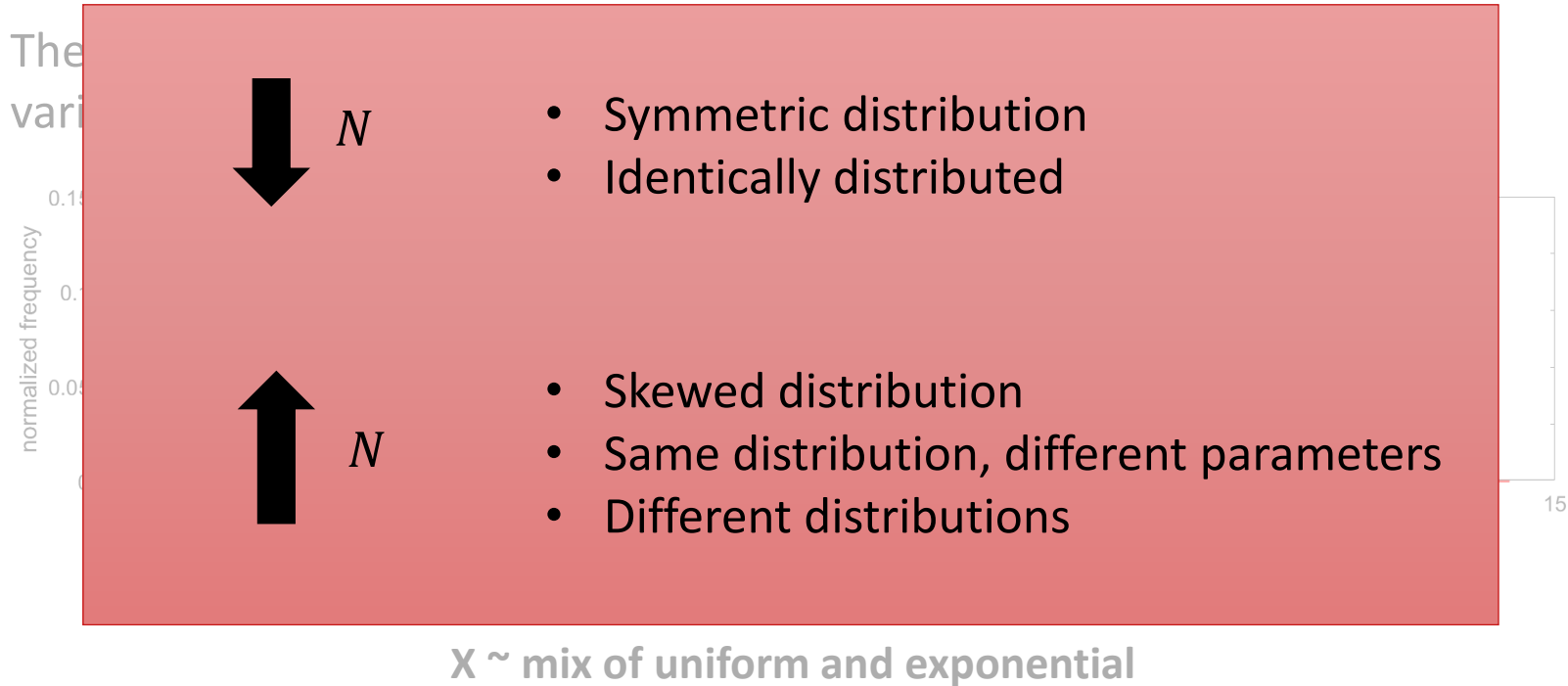
“Sufficiently large” number condition

The normalized sum of a **sufficient** number of **independent** random variables tends to a Gaussian distribution.



$X \sim \text{mix of uniform and exponential}$

“Sufficiently large” number condition



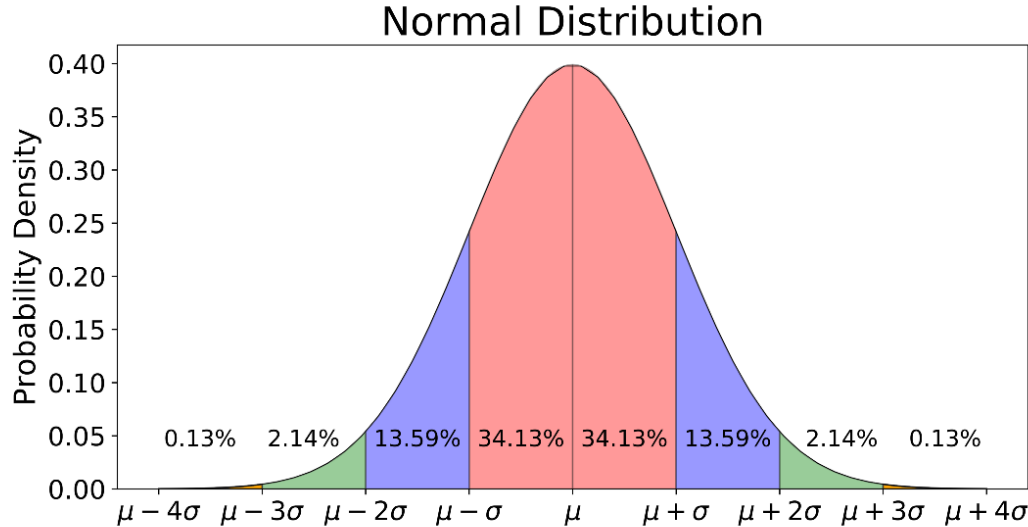
Application CLT

- If we sum multiple random variables, however distributed, we can apply the central limit theorem to obtain the full probability model (Gaussian with known expected value and variance)
- In practice: any random variable resulting from the sum of multiple random variables (e.g., multiple noise sources affecting a measurement) can be considered gaussian distributed
- From the obtained Gaussian probability model, we can perform easy and accurate calculations

Gaussian (normal) distribution

$$N(\mu, \sigma^2): \quad p_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- μ : mean
- σ : standard deviation



- $\Pr[-\sigma < T < \sigma] \sim 68.2\%$
- $\Pr[-2\sigma < T < 2\sigma] \sim 95.4\%$
- $\Pr[-3\sigma < T < 3\sigma] \sim 99.7\%$

Gaussian (normal) distribution: Q-function

- The **standard** normal distribution, also called **z-distribution** is a normal distribution with *zero-mean* and *unit variance* (standard deviation)

$$Z = \frac{X - \mu_X}{\sigma_X}$$

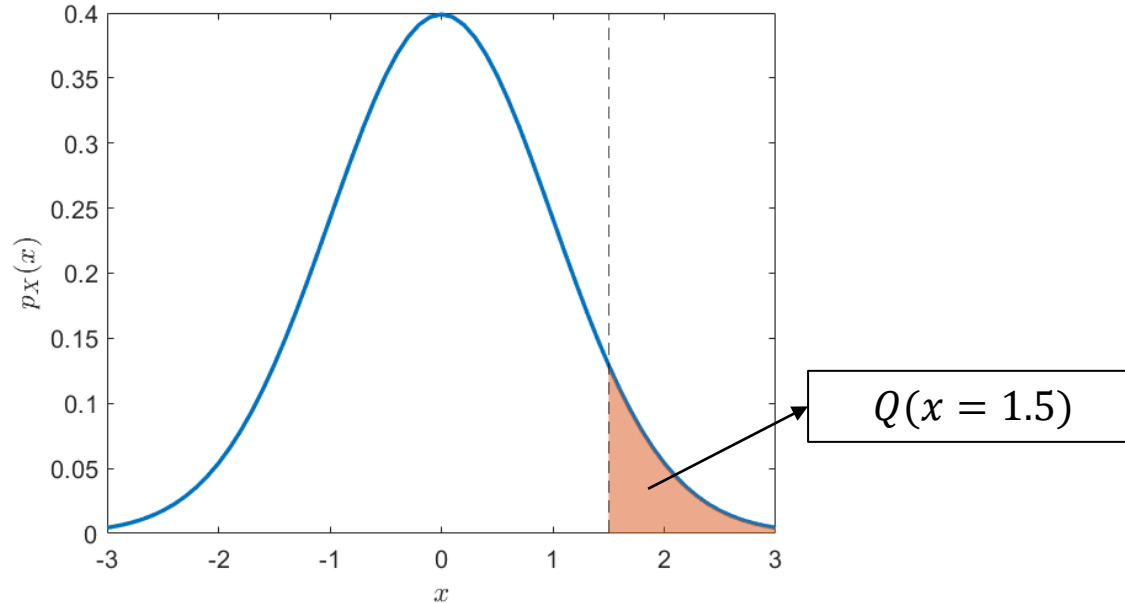
- X , gaussian random variable
- μ_X , mean of X
- σ_X , standard deviation

- The **Q-function** is the tail distribution function of the standard normal distribution

$$Q(x) = \Pr[X > x] = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}} du$$

Gaussian (normal) distribution: Q-function

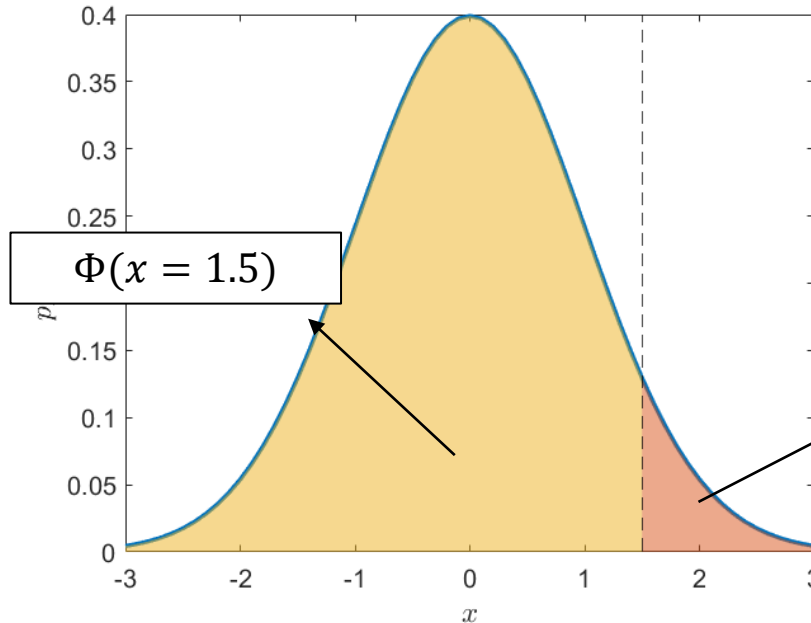
$$Q(x) = \Pr[X > x] = \frac{1}{\sqrt{2\pi}1^2} \int_x^\infty e^{-\frac{(u-0)^2}{2 \cdot 1^2}} du$$



Gaussian (normal) distribution: Q-function

$$Q(x) = \Pr[X > x] = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}} du = 1 - Q(-x) = 1 - \Phi(x)$$

CDF of standard
normal distribution



MATLAB functions

qfunc(x)
qfuncinv(x)
normcdf(x)
normcdf(x, mu, sigma)

Application CLT: example

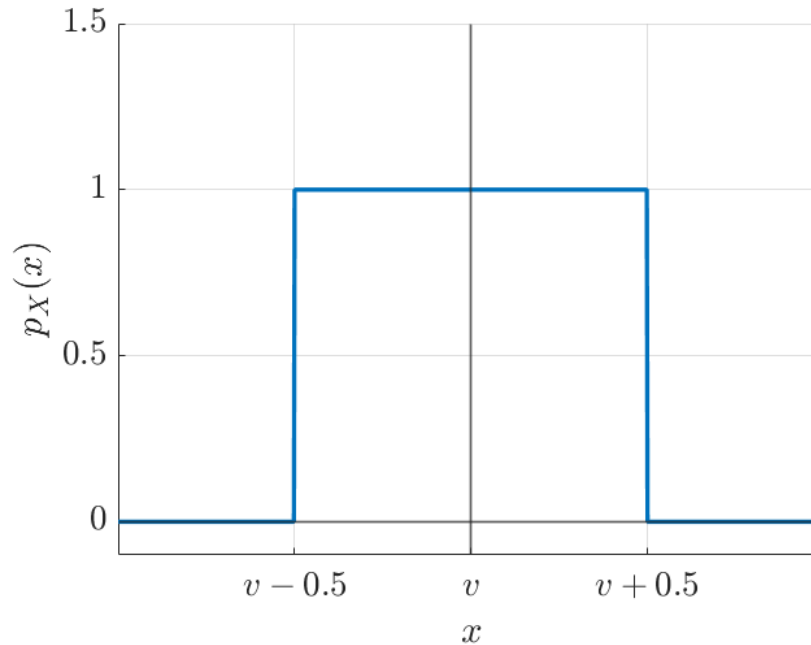


- A compact disc (CD) contains digitized samples of an acoustic waveform.
- In a CD player with a “one-bit digital to analog converter,” each digital sample is represented to an **accuracy of ± 0.5 mV**.
- The CD player “oversamples” the waveform by making **eight independent measurements** corresponding to each sample.
- The CD player obtains a waveform sample by calculating the average (sample mean) of the eight measurements.

What is the probability that the error in the waveform sample is greater than 0.1 mV?

Application CLT: example

- Each digital sample is represented to an **accuracy of ± 0.5 mV**
- With can model each measurement with a uniform distribution

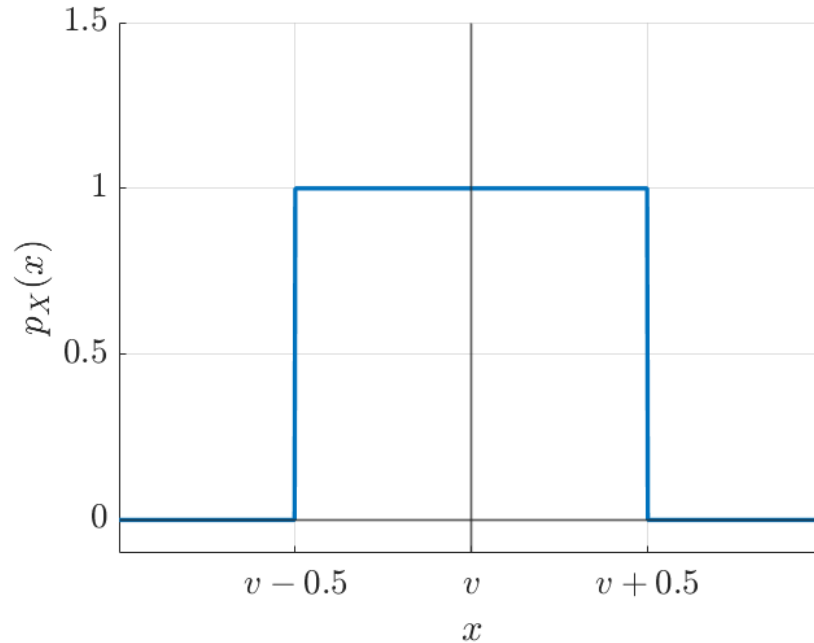


Application CLT: example

- **Approach 1: we don't know the CLT...**
 - Calculate exact probability model
 - Eightfold convolution of the uniform PDF [out of scope]
 - Moment generating function method [out of scope]
- **Approach 2: use the CLT!**
 - Calculate the expected mean and variance of the sum of random variables
 - Model the sum as Gaussian probability function
 - Use Gaussian CDF (or Q function) to calculate probability

Solution

Measurements X_i have all uniform distribution between $v-0.5$ mV and $v + 0.5$ mV



$$E[X_i] = \frac{a+b}{2}$$

$$\text{Var}[X_i] = \frac{1}{12}(b-a)^2$$

Solution

Measurements X_i have all uniform distribution between $v-0.5$ mV and $v + 0.5$ mV

$$E[X_i] = \frac{a+b}{2} = \frac{v-0.5+v+0.5}{2} = \frac{2v}{2} = v$$

$$\text{Var}[X_i] = \frac{1}{12}(b-a)^2 = \frac{1}{12}(v-0.5-(v-0.5))^2 = \frac{1}{12}$$

Solution

- Measurements X_i have all uniform distribution between $v - 0.5$ mV and $v + 0.5$ mV
- Output U of CD player:

$$U = \frac{1}{8} \sum_{i=1}^8 X_i$$

- Using central limit theorem:
 $E[U] = 8E[X_i] / 8 = v$
 $\text{var}[U] = 8 \text{var}[X_i] / 64 = 1 / 8 \cdot 1 / 12 = 1 / 96$

Output U is approximately Gaussian with mean v and variance $1/96$

Solution

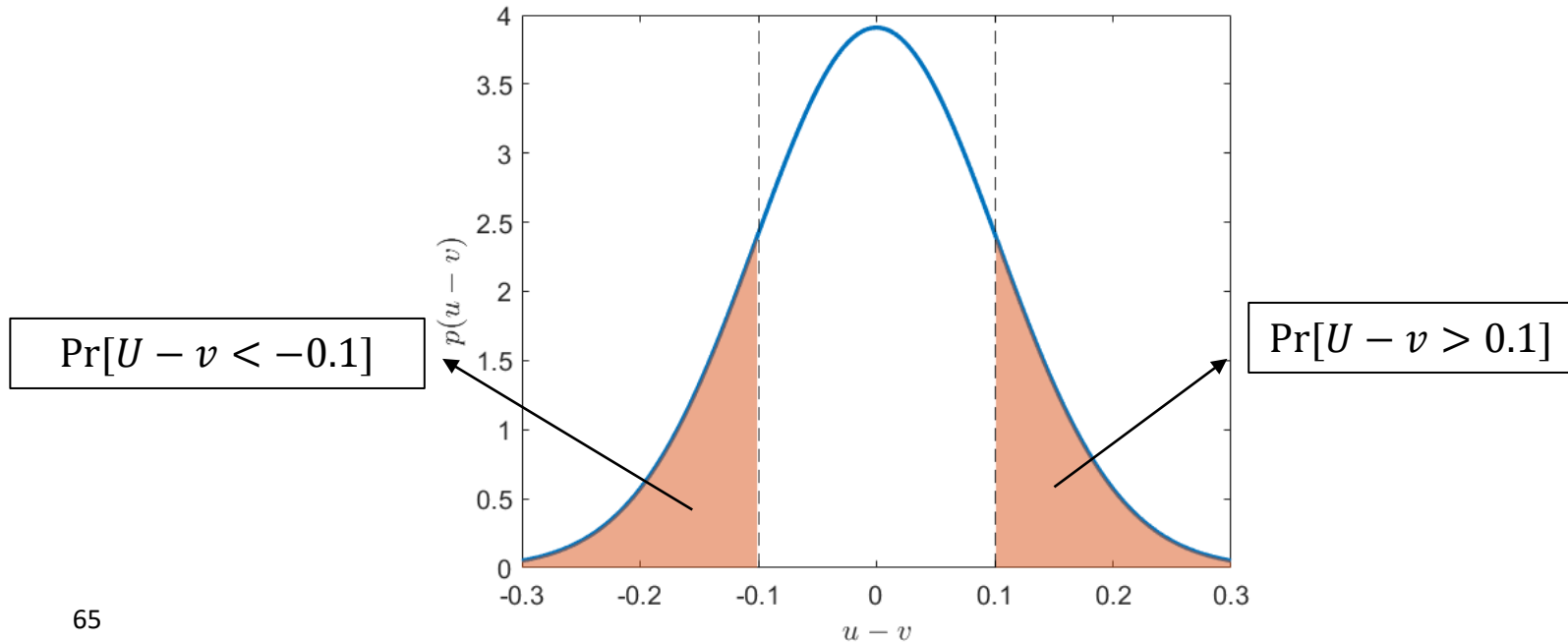
- We want to know $\Pr[|U - v| > 0.1]$
- Output U is approximately Gaussian with mean v and variance $1/96$



Error random variable $U - v$ is approximately Gaussian
with mean 0 and variance $1/96$

Solution

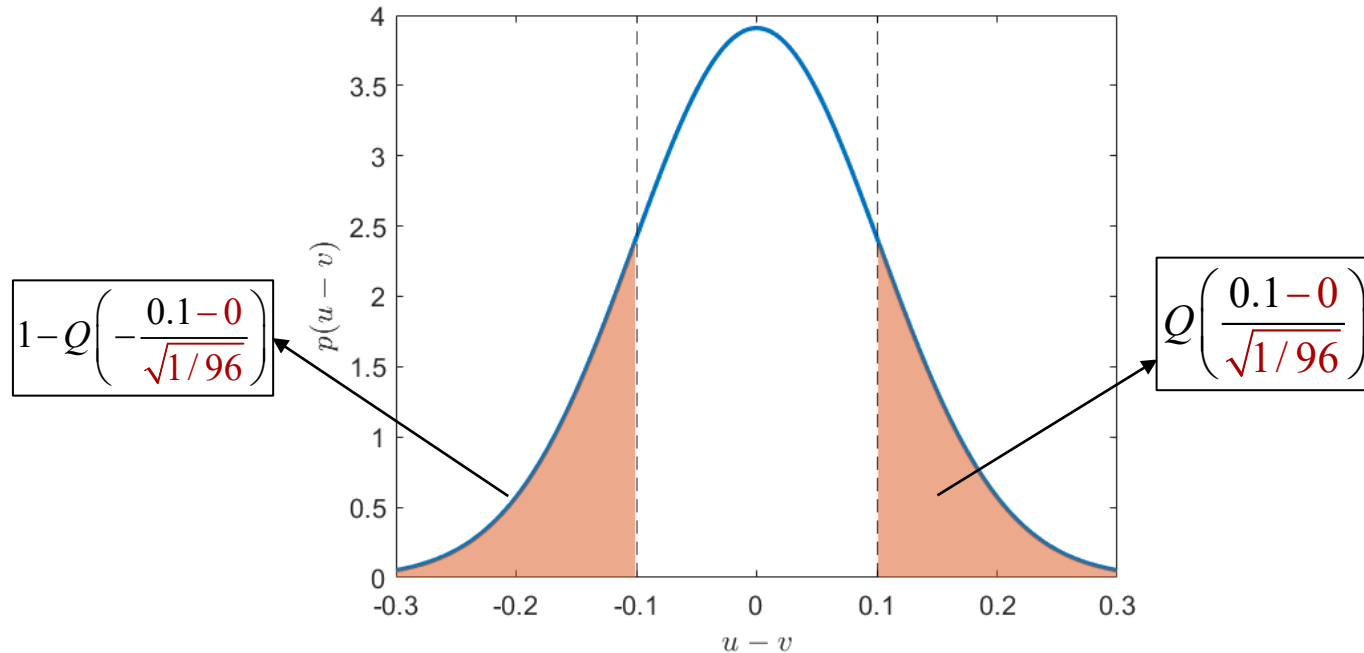
$$\Pr[|U - v| > 0.1]$$



Solution

$$\Pr[|U - v| > 0.1] = 2 \cdot Q\left(\frac{0.1 - 0}{\sqrt{1/96}}\right)$$

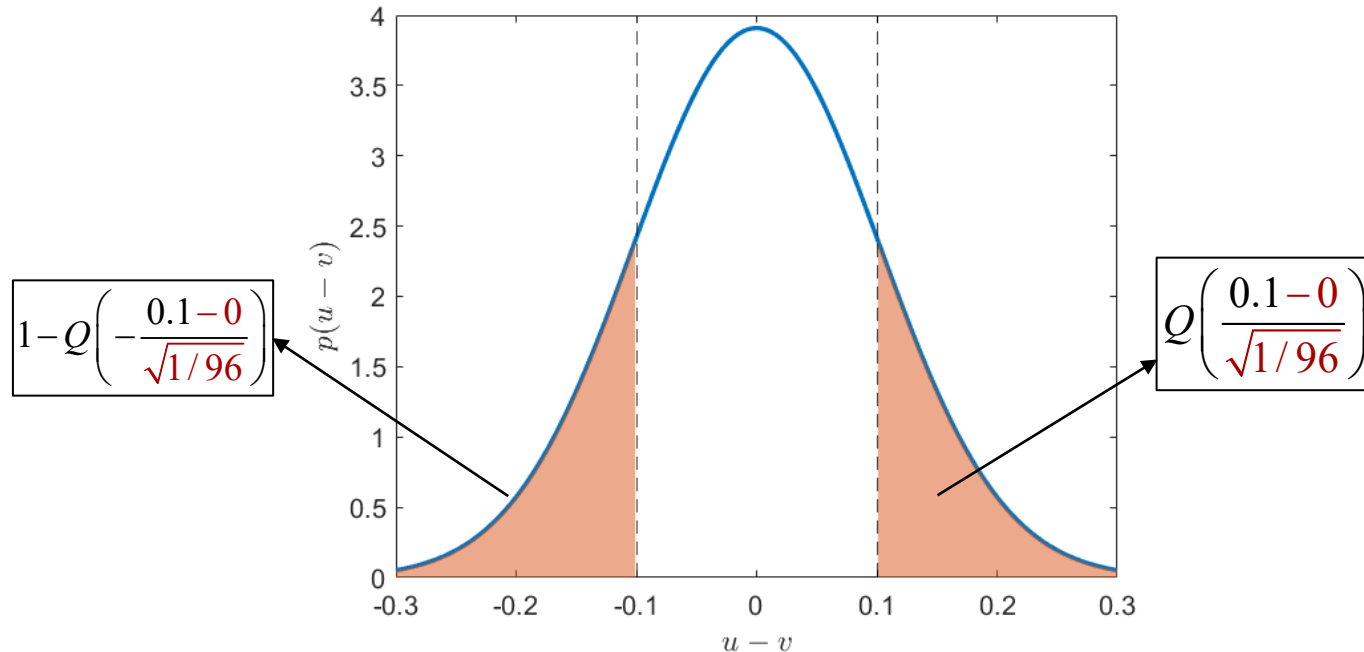
standardization



Solution

$$\Pr[|U - v| > 0.1] = 2 \cdot Q\left(\frac{0.1 - 0}{\sqrt{1/96}}\right) = 2 \cdot \left(1 - \Phi\left(\frac{0.1}{\sqrt{1/96}}\right)\right) = 0.3272$$

standardization

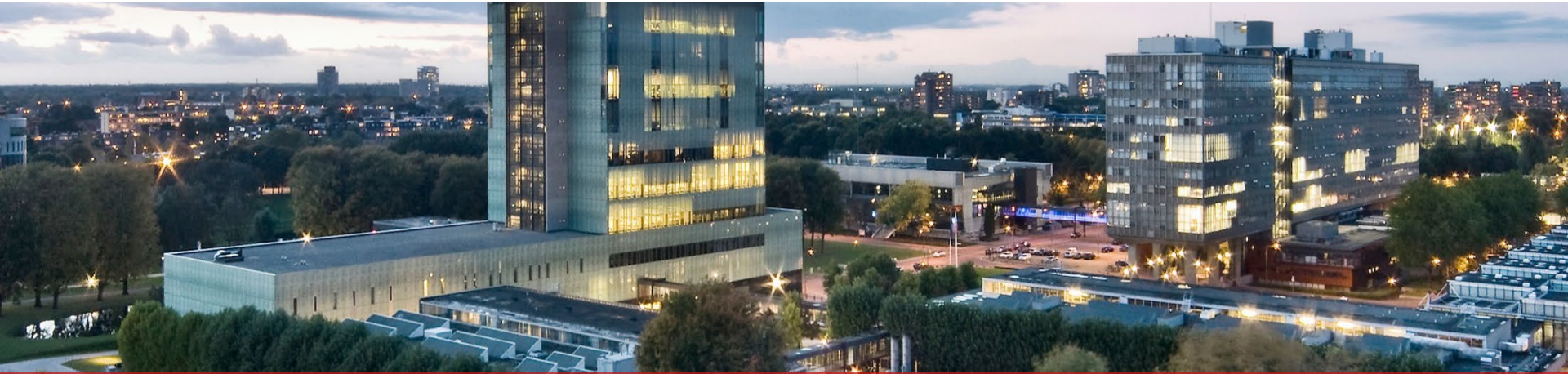


Wrap up (I)

- If we apply a **function** to a **random variable**, the probability model as well as the moments may change
- Stochastic process involving **pairs of random variables** can be described by joint probability models
- Given a joint probability distribution, and an observed event or random variable, we can define **conditional probability distributions**
- The operation of **marginalization** permits calculating the probability distribution of an individual random variable, given the joint probability distribution

Wrap up (II)

- The **central limit theorem** allows to assume a gaussian probability distribution as the result of summing “many”, “not-to-correlated”, “not-too-large” random variables
- The **standard normal random variable** is defined as a normal random variable with zero mean and unit variance
- The **Q-function** is the tail distribution function of the standard normal distribution



Statistical signal processing (5CTA0)

Lecture 2, part A

Lecturer: Simona Turco

Electrical Engineering, Signal Processing Systems group