CT111 Introduction to Communication Systems
Lecture 10: A Review of Probability Theory

Yash M. Vasavada

Associate Professor, DA-IICT, Gandhinagar

9th February 2018

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

1. Average Value
2. Probability
3. A Vectorial Point of View
4. Characterization of RVs
5. Examples of Discrete RVs
6. Examples of Continuous RVs

Overview of Today's Talk

Overview of Today's Talk

Taking an Average

- We have earlier seen this operation on a time domain signal $w(t)$:

$$\langle w(t) \rangle$$

Taking an Average

- We have earlier seen this operation on a time domain signal $w(t)$:

$$\langle w(t) \rangle \stackrel{\text{def}}{=} \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} w(t) \, dt$$

- Let us reflect on this a bit more

Taking an Average

- Let us say that in a quiz of 10 marks, the following is the observed performance of 30 groups:
  - $\rightarrow$ 8, 9, 1, 10, 6, 1, 3, 6, 10, 10, 1, 10, 10, 5, 8, 1, 4, 10, 8, 10, 7, 0, 9, 10, 7, 8, 8, 4, 7, 1
- Suppose we want to know the average score. One way is simple:
  - $\rightarrow$ Average score: $(8+9+1+10+6+1+3+6+10+10+1+10+10+5+8+1+4+10+8+10+7+0+9+10+7+8+8+4+7+1)/30 = 6.4$

Taking an Average

- Let us write this mathematically:
    - $\rightarrow$ Let k denote the index of the group, where $k = 1, 2, \ldots, N$
    - $\rightarrow$ $N = 30$ is the total number of groups
    - $\rightarrow$ Let $x(k)$ denote the mark obtained by k^{th} group. Thus, $x_1 = 8, x_2 = 9$ and so on
- With this, the average marks can be written as follows:

$$\bar{X} = \frac{1}{N} \sum_{k=1}^{N} x(k)$$

Here X is a notation for the Random Variable (or RV), and $\bar{X}$ denotes the average

Taking an Average

- Let us write the average operation mathematically:

$$\bar{X} = \frac{1}{N} \sum_{k=1}^{N} x(k)$$

- X is a notation for the Random Variable (or RV)
  - ▷ An RV is a number, it can be an integer (in which case the RV is called a Discrete RV) or a real number (in which case it is called a continuous-valued RV) or a complex number (Complex RV)
- In our example, X stands for the score obtained by different groups
  - ▷ Note that X is inherently random; we don't know what marks different groups are going to get in the next quiz
- $x(k)$ denotes k^{th} realization of X
- $\bar{X}$ denotes the average

Random Variables in Communication Systems

Random variables are extremely useful to model the communication systems. These let us talk about quantities and signals which are *not* known in advance:

- Data sent through the communication channel is best modeled as a sequence of random variables
- Message signal value at any given time is also an RV (e.g., consider the speech signal)
- Noise, interference and fading affecting this data transmission are all signals whose values at any given time is given by an RV
- Receiver performance is measured probabilistically and by means of an RV

Taking an Average

- Let us say that in a quiz of 10 marks, the following is the observed performance of 30 groups:
  - → 8, 9, 1, 10, 6, 1, 3, 6, 10, 10, 1, 10, 10, 5, 8, 1, 4, 10, 8, 10, 7, 0, 9, 10, 7, 8, 8, 4, 7, 1
- Suppose we want to know the average score. One way is simple:
  - → Average score: (8+9+1+10+6+1+3+6+10+10+1+10+10+5+8+1+ 4+10+8+10+7+0+9+10+7+8+8+4+7+1)/30 = 6.4
- There is another way possible:
  - → Average score:
    $((1 \times 0) + (5 \times 1) + (0 \times 2) + (1 \times 3) + (2 \times 4) + (1 \times 5)+$
    $(2 \times 6) + (3 \times 7) + (5 \times 8) + (2 \times 9) + (8 \times 10))/30 = 6.4$

Taking an Average

- Let us say that in a quiz of 10 marks, the following is the observed performance of 30 groups:
  - $\rightarrow$ 8, 9, 1, 10, 6, 1, 3, 6, 10, 10, 1, 10, 10, 5, 8, 1, 4, 10, 8, 10, 7, 0, 9, 10, 7, 8, 8, 4, 7, 1
- Suppose we want to know the average score. One way is simple:
  - $\rightarrow$ Average score: $(8+9+1+10+6+1+3+6+10+10+1+10+10+5+8+1+4+10+8+10+7+0+9+10+7+8+8+4+7+1)/30 = 6.4$
- There is another way possible:
  - $\rightarrow$ Average score:
    $((1 \times 0) + (5 \times 1) + (0 \times 2) + (1 \times 3) + (2 \times 4) + (1 \times 5) + (2 \times 6) + (3 \times 7) + (5 \times 8) + (2 \times 9) + (8 \times 10))/30 = 6.4$

Taking an Average

- Let us write the average operation mathematically *in an alternate way*:

$$\bar{X} = \frac{1}{N} \sum_{k=1}^{N} x(k) = \frac{1}{N} \sum_{m=0}^{M} n_{x_m} x_m$$

- Here, we have introduced two different notations:
    - ▷ In our case, X is a discrete RV, and its value can be any one integer from $\{0, 1, \ldots, 10\}$, where 10 is the maximum score possible in the quiz.
    - ▷ Integer x_m denotes the values that the RV can take and M is the maximum value.
    - ▷ Finally, n_{x_m} denotes the number of times x_m occurs in the sample set

Understanding the Probability

- Suppose someone says that the probability of the quiz score of 10 out of 10 is 5%, on what basis he/she made this statement?
- One answer is that it is his or her belief.
- How did he/she arrive at this belief?
- Because he/she saw that on average only 5% of the groups got full marks
- However, what does this mean mathematically?

$$p_{x_m} \stackrel{\text{def}}{=} \lim_{N \to \infty} \frac{n_{x_m}}{N}$$

Understanding the Probability

- Suppose someone says that the probability of the quiz score of 10 out of 10 is 5%, on what basis he/she made this statement?
- One answer is that it is his or her belief.
- How did he/she arrive at this belief?
- Because he/she saw that on average only 5% of the groups got full marks
- However, what does this mean mathematically?

$$p_{x_m} \overset{\mathrm{def}}{=} \lim_{N \to \infty} \frac{n_{x_m}}{N}$$

Understanding the Probability

- Suppose someone says that the probability of the quiz score of 10 out of 10 is 5%, on what basis he/she made this statement?
- One answer is that it is his or her belief.
- How did he/she arrive at this belief?
- Because he/she saw that on average only 5% of the groups got full marks
- However, what does this mean mathematically?

$$p_{x_m} \overset{\text{def}}{=} \lim_{N \to \infty} \frac{n_{x_m}}{N}$$

Understanding the Probability

- Suppose someone says that the probability of the quiz score of 10 out of 10 is 5%, on what basis he/she made this statement?
- One answer is that it is his or her belief.
- How did he/she arrive at this belief?
- Because he/she saw that on average only 5% of the groups got full marks
- However, what does this mean mathematically?

$$p_{x_m} \overset{\text{def}}{=} \lim_{N \to \infty} \frac{n_{x_m}}{N}$$

Understanding the Probability

- Suppose someone says that the probability of the quiz score of 10 out of 10 is 5%, on what basis he/she made this statement?
- One answer is that it is his or her belief.
- How did he/she arrive at this belief?
- Because he/she saw that on average only 5% of the groups got full marks
- However, what does this mean mathematically?

$$p_{x_m} \overset{\text{def}}{=} \lim_{N \to \infty} \frac{n_{x_m}}{N}$$

Understanding the Probability

- Suppose someone says that the probability of the quiz score of 10 out of 10 is 5%, on what basis he/she made this statement?
- One answer is that it is his or her belief.
- How did he/she arrive at this belief?
- Because he/she saw that on average only 5% of the groups got full marks
- However, what does this mean mathematically?

$$p_{x_m} \stackrel{\text{def}}{=} \lim_{N \to \infty} \frac{n_{x_m}}{N}$$

Probability
and Averages

- Probability that a random variable X takes a value of x_m can be written as:

$$p_{x_m} \stackrel{\text{def}}{=} \lim_{N \to \infty} \frac{n_{x_m}}{N}$$

Here, n_{x_m} is the number of times x_m occurs in a total of N experiements

- With this definition of the probability, the average can be written as

$$\bar{X} = \frac{1}{N} \sum_{k=1}^{N} x(k) = \frac{1}{N} \sum_{m=0}^{M} n_{x_m} x_m = \sum_{m=0}^{M} p_{x_m} x_m$$

Probability
Averages and Expected Value for Discrete RVs

- Expected value μ_X of an RV X is defined as

$$E(X) = \mu_X = \sum_X p_X(x)x$$

- Technically, there is a subtle difference between the expected value μ_X and the average value $\bar{X}$

  ▷ $\bar{X} = \dfrac{1}{N}\sum_{k=1}^{N} x(k)$ is the time domain average, does not require knowledge of the probabilities $p_X(x)$, but it requires statistical experiments

  ▷ $E(X) = \mu_X = \sum_X p_X(x)x$ is the statistical average, requires knowledge of the probabiilties, but does not require experimentation

  ▷ As $N \to \infty$, $\bar{X} \to E(X)$

Expected Values

- The definition on the prior slide is for the discrete RV. It generalizes to the continuous RV:

  $\rightarrow$ Mean: $E(X) = m_X = \int_{-\infty}^{\infty} x \, p_X(x) \, dx$

- Thus, there is yet another name; average or expected value of the RV is also called the *mean* of the RV

- Note when X stands for a random signal (also called as random process), m_X can be thought of as the DC component of the signal.

  $\rightarrow$ When the DC component is zero, the random signal is called a zero-mean random signal (or process)

- Can we know what the mean value is if we are given the Fourier Transform of a signal?

Probability
Averages and Expected Value for Continuous RVs

- Expected value μ_X of an RV X is defined as

$$E(X) = \mu_X = \int_X p_X(x) x \, dx$$

- Technically, there is a subtle difference between the expected value μ_X and the time-domain average value $\bar{X}$
  - $\triangleright$ $\bar{X} = \langle x(t) \rangle$ is the time domain average, does not require knowledge of the probabilities $p_X(x)$, but it requires statistical experiments
  - $\triangleright$ $E(X) = \mu_X = \sum_X p_X(x) x$ is the statistical average, requires knowledge of the probabiilties, but does not require experimentation
  - $\triangleright$ As $N \to \infty$, $\bar{X} \to E(X)$

Expected Values
Mean-Squared

- There is one type of *mean* which is quite useful:

  $\rightarrow$ Mean-Squared: $E(X^2) = m_{X^2} = \int_{-\infty}^{\infty} x^2 \, p_X(x) \, dx$

- Why is $E(X^2)$ a useful mean value?

  $\rightarrow$ $E(X^2)$ can be equated to the average power in the random variable X

  $\rightarrow$ Recall RMS (Root Mean Squared). We have defined it earlier for deterministic signals. For random signals and random variables, the RMS is defined as $\sqrt{E(X^2)}$

Expected Values
Mean-Squared

- There is one type of *mean* which is quite useful:

  $\rightarrow$ Mean-Squared: $E(X^2) = m_{X^2} = \displaystyle\int_{-\infty}^{\infty} x^2 \, p_X(x) \, dx$

- Why is $E(X^2)$ a useful mean value?

  $\rightarrow$ $E(X^2)$ can be equated to the average power in the random variable X

  $\rightarrow$ Recall RMS (Root Mean Squared). We have defined it earlier for deterministic signals. For random signals and random variables, the RMS is defined as $\sqrt{E(X^2)}$

Expected Values
Mean-Squared

- There is one type of *mean* which is quite useful:

  $\rightarrow$ Mean-Squared: $E(X^2) = m_{X^2} = \int_{-\infty}^{\infty} x^2 \, p_X(x) \, dx$

- Why is $E(X^2)$ a useful mean value?
  - $\rightarrow$ $E(X^2)$ can be equated to the average power in the random variable X
  - $\rightarrow$ Recall RMS (Root Mean Squared). We have defined it earlier for deterministic signals. For random signals and random variables, the RMS is defined as $\sqrt{E(X^2)}$

Expected Values
Variance

- There is another type of *mean* which is quite useful:
    - $\rightarrow$ Variance: $\sigma^2 \stackrel{\text{def}}{=} E\left((x - m_X)^2\right) = \int_{-\infty}^{\infty} (x - m_X)^2 \, p(x) \, dx$
    - $\rightarrow$ Variance measures the power of the truly random component, after removing the constant or the DC component of the signal given by m_X
   - $\rightarrow$ It can be proven that $E(X^2) = m_X^2 + \sigma_X^2$.

Expected Values
Correlation

- For random signals (or random processes), there is yet another type of mean value which is very important:
    - $\rightarrow$ Define a product signal $Y(t, \tau) = X(t)X^*(t + \tau)$
    - $\rightarrow$ Correlation: $R_X(\tau) = E[Y(t, \tau)]$
    - $\rightarrow$ Correlation: $\langle Y(t, \tau) \rangle \rightarrow R_X(\tau)$ as the length of the time domain averaging increases
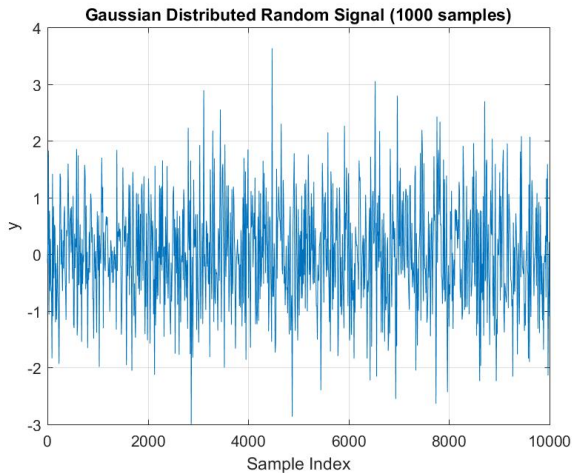   - $\rightarrow$ Correlation measures how fast the random process changes with time t.

Example Random Process: 1
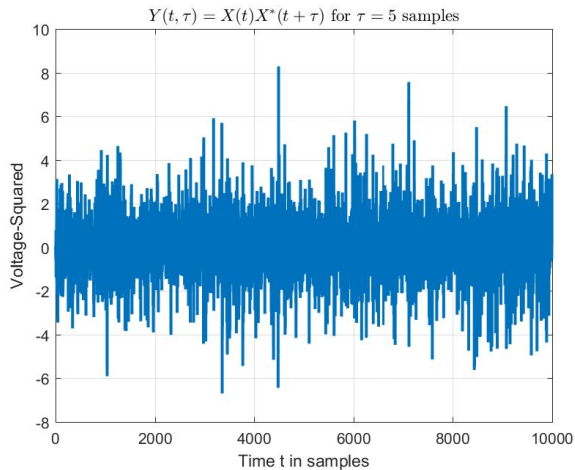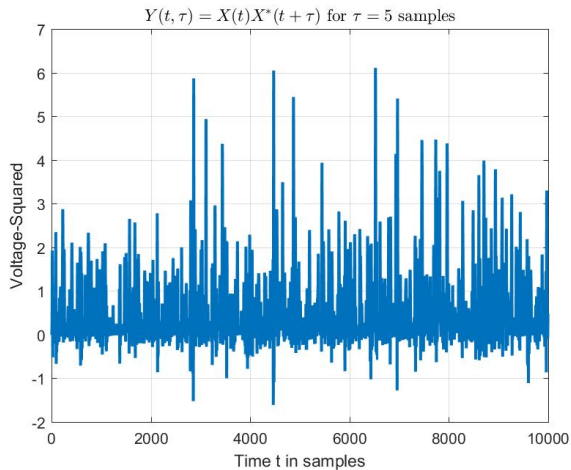$X_1(t)$
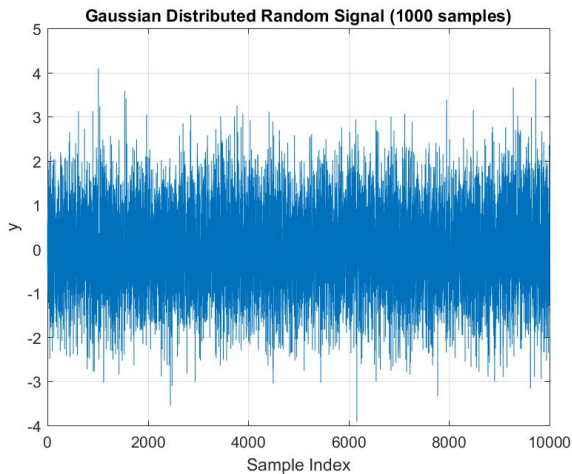
Example Random Process: 2
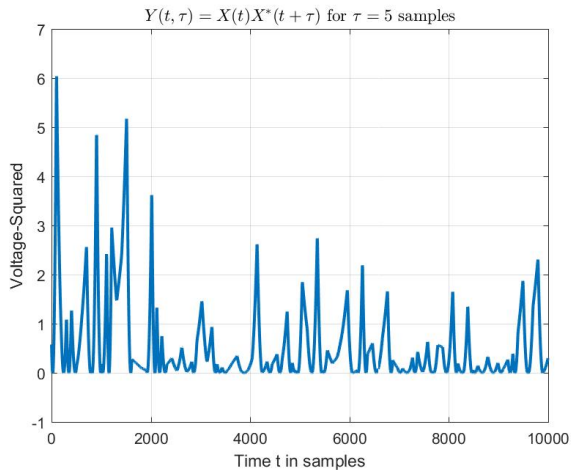$X_2(t)$

Example Random Process: 3
$X_3(t)$
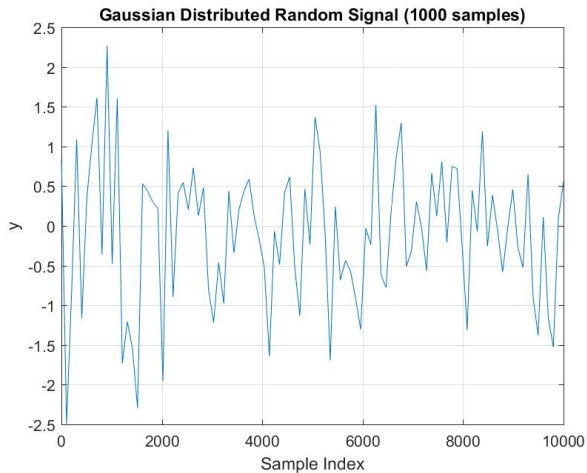
$Y_1(t,\tau) = X_1(t)X_1^*(t+\tau)$ for $\tau = 5$ samples for
Example Random Process: 1

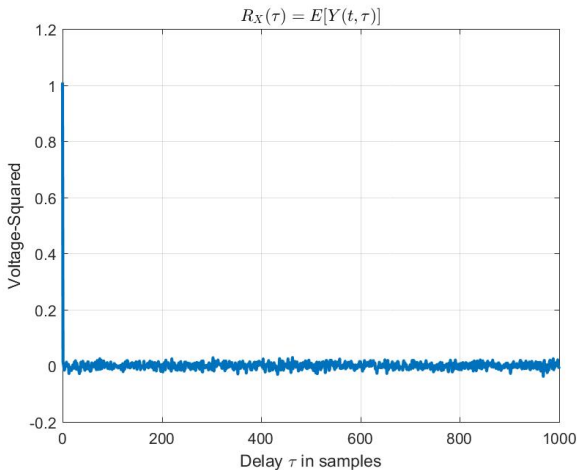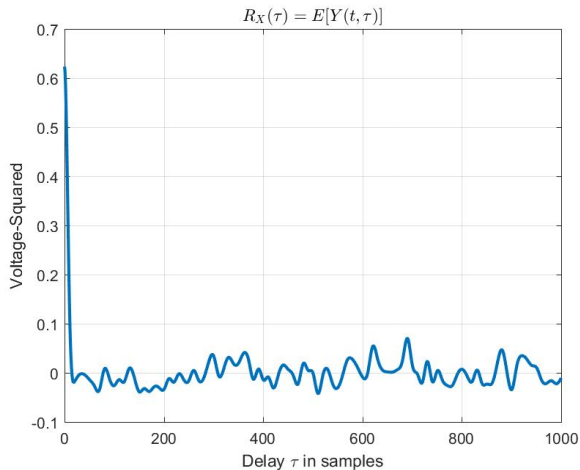# $Y_2(t,\tau) = X_2(t)X_2^*(t+\tau)$ for $\tau = 5$ samples for
Example Random Process: 2

$Y(t,\tau) = X(t)X^*(t+\tau)$ for $\tau = 5$ samples

$Y_3(t, \tau) = X_3(t)X_3^*(t + \tau)$ for $\tau = 5$ samples for
Example Random Process: 3

$R_{X_1}(\tau) = E\left[Y_1(t, \tau)\right]$ for all τ
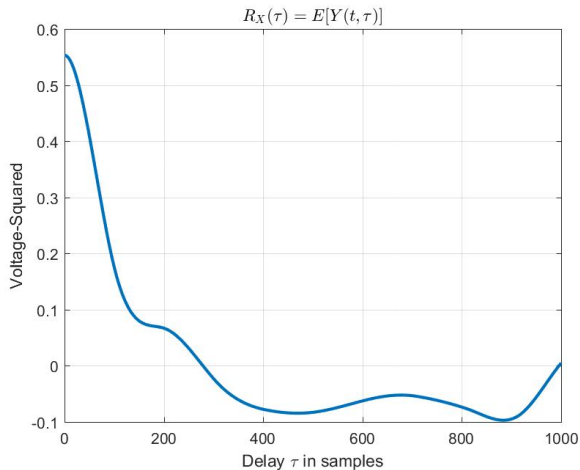Example Random Process: 2

$R_{X_2}(\tau) = E\left[Y_2(t, \tau)\right]$ for all τ
Example Random Process: 2

$R_{X_3}(\tau) = E\left[Y_3(t, \tau)\right]$ for all τ
Example Random Process: 2

Wiener Khinchine Theorem

- $P(f) = E\left[|X(f)|^2\right]$ is called the Power Spectral Density or PSD
- Weiner Khinchine Theorem: $R(\tau) \rightleftharpoons P(f)$
  - $\rightarrow$ Correlation and the PSD form a Fourier Transform Pair
  - $\rightarrow$ A question: what is $P(f)$ if $R(\tau)$ can be approximated by $\delta(\tau)$?
  - $\rightarrow$ Answer: spectrally flat PSD over $f \in [-\infty, +\infty]$

Wiener Khinchine Theorem

- $P(f) = E\left[|X(f)|^2\right]$ is called the Power Spectral Density or PSD
- Weiner Khinchine Theorem: $R(\tau) \rightleftharpoons P(f)$
    - $\rightarrow$ Correlation and the PSD form a Fourier Transform Pair
    - $\rightarrow$ A question: what is $P(f)$ if $R(\tau)$ can be approximated by $\delta(\tau)$?
    - $\rightarrow$ Answer: spectrally flat PSD over $f \in [-\infty, +\infty]$

Wiener Khinchine Theorem

- $P(f) = E\left[|X(f)|^2\right]$ is called the Power Spectral Density or PSD
- Weiner Khinchine Theorem: $R(\tau) \rightleftharpoons P(f)$
  - $\rightarrow$ Correlation and the PSD form a Fourier Transform Pair
  - $\rightarrow$ A question: what is $P(f)$ if $R(\tau)$ can be approximated by $\delta(\tau)$?
  - $\rightarrow$ Answer: spectrally flat PSD over $f \in [-\infty, +\infty]$

Dot Product

- Correlation operation can be thought of as taking a dot product
- We have seen the dot product operation earlier:
  1. Fourier Transformation
  2. Convolution Operation (in Filtering)

- Dot Product is an extremely powerful concept, which can be best understood by means of vector algebra

Dot Product

- Correlation operation can be thought of as taking a dot product
- We have seen the dot product operation earlier:
  1. Fourier Transformation
  2. Convolution Operation (in Filtering)
- Dot Product is an extremely powerful concept, which can be best understood by means of vector algebra

Vectorial Form

Let us define vectors as a collection of numbers:

$$\mathbf{h} = [h_1, h_2, \ldots, h_N]^T$$
$$\mathbf{r} = [r_1, r_2, \ldots, r_N]^T$$
$$\mathbf{h} \cdot \mathbf{r} = \mathbf{h}^T \mathbf{r} = h_1 r_1 + h_2 r_2 + \ldots + h_N r_N$$

Here, the last term $\mathbf{h} \cdot \mathbf{r} = \mathbf{h}^T \mathbf{r}$ is known as the *dot product* between the vectors $\mathbf{h}$ and $\mathbf{r}$.

Vectorial Form
Notations

- Transposition
  - ▷ $\mathbf{h} = [h_1, h_2, \ldots, h_N]$ denotes a *row* vector with 1 row and N columns (i.e., of size $1 \times N$)
  - ▷ $\mathbf{h}^T = [h_1, h_2, \ldots, h_N]^T$ denotes a *column* vector with N rows and 1 column (size $N \times 1$)
    - → Think of the row vector as a *horizontal* vector sleeping on a couch, and the column vector as a *vertical* vector standing *tall*
    - → Here $(\cdot)^T$ denotes the transposition operation: turns row vectors into column vectors and vice versa

Vectorial Form
Notations

- A rule of vector product:
    ▷ Cannot multiply a $1 \times N$ vector with $1 \times N$ vector. The internal dimensions don't match
    ▷ Can multiply a $1 \times N$ vector with $N \times 1$ vector. The internal dimensions do match.

    Therefore, in $\mathbf{h} \cdot \mathbf{r} = \mathbf{h}^T \mathbf{r}$, the first vector is a horizontal vector of size $1 \times N$ and the second vector is a tall vector of size $N \times 1$
    $\rightarrow$ If the vectors have complex-valued coordinates, the dot product requires Hermitian Transpose (transpose and conjugation): $\mathbf{h} \cdot \mathbf{r} = \mathbf{h}^H \mathbf{r}$

Vectors
Two Key Properties

Following are two key properties of the dot products.

1. Dot product of any vector **x** with itself is the square of the vector length.
2. Dot product of vectors **x** and **y** is directly related to the *cosine* of the angle between the vectors **x** and **y**.
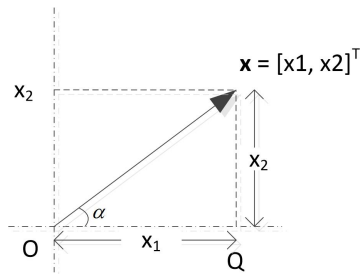
We examine each of these two properties next.

Dot Product between Two Vectors
First Key Property: Length of any Vector

- Length of a vector **x** is denoted as $\|\mathbf{x}\|$.
- Length of a vector $\mathbf{x} = [x_1, x_2]$ in 2D plane is the length of *hypotenuse* of a right angled triangle whose bases have lengths x_1 and x_2.
- The square of length of this hypotenuse is given by Pythagoras theorem:

$$\|\mathbf{x}\|^2 = x_1^2 + x_2^2 = \mathbf{x}^T\mathbf{x}$$

Dot Product between Two Vectors
First Key Property: Length of any Vector

- This can be extended to three dimensional space.

$$\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + x_3^2 = \mathbf{x}^T\mathbf{x}$$

- Therefore, we conclude, generalizing to N dimensional space, that the dot product of any N-dimensional vector $\mathbf{x}$ with itself gives the square of the length of the vector.

$$\mathbf{x}^T\mathbf{x} = x_1^2 + x_2^2 + x_3^2 + \ldots + x_N^2 = \|\mathbf{x}\|^2$$

First Key Property: Length of any Vector

- What does $\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + x_3^2 + \ldots + x_N^2 = \mathbf{x}^T\mathbf{x}$ represent physically?

  $\triangleright$ Approaches $N \times E\left[X^2\right]$ as $N \to \infty$

- Thus, the power of the signal $E\left[X^2\right]$ can be thought of as the normalized length $\|\mathbf{x}\|^2/N$ of the signal vector

First Key Property: Length of any Vector

- What does $\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + x_3^2 + \ldots + x_N^2 = \mathbf{x}^T\mathbf{x}$ represent physically?
  - ▷ Approaches $N \times E\left[X^2\right]$ as $N \to \infty$
- Thus, the power of the signal $E\left[X^2\right]$ can be thought of as the normalized length $\|\mathbf{x}\|^2 / N$ of the signal vector

Dot Product between Two Vectors
Second Key Property: Angle between Two Vectors

- Now, let us consider the dot product of vectors **x** and **y**.
- Length $\|x\|$ is the hypotenuse in the triangle OxQ, and the sine and consine of α are

$$\sin \alpha = \frac{x_2}{\|\mathbf{x}\|}, \quad \cos \alpha = \frac{x_1}{\|\mathbf{x}\|}$$
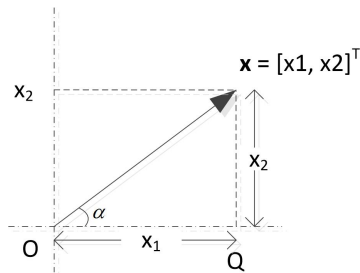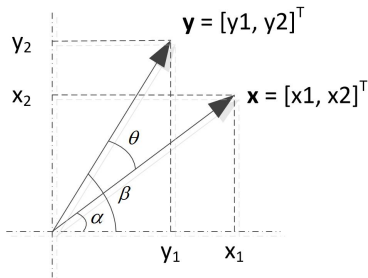
Dot Product between Two Vectors
Second Key Property: Angle between Two Vectors

- Similarly, for angle β for vector **y**,
  $\rightarrow \sin \beta = \dfrac{y_2}{\|\mathbf{y}\|}$, and $\cos \beta = \dfrac{y_1}{\|\mathbf{y}\|}$.

- Now,

$$
\begin{aligned}
\cos \theta &= \cos (\beta - \alpha) \\
&= \cos \beta \cos \alpha + \sin \beta \sin \alpha \\
&= \frac{x_1 y_1 + x_2 y_2}{\|\mathbf{x}\| \, \|\mathbf{y}\|} \\
&= \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}
\end{aligned}
$$

Dot Product between Two Vectors
Second Key Property: Angle between Two Vectors

- dot product is proportional to the cosine of the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$:

$$\mathbf{x}^T\mathbf{y} = \cos\theta \times \|\mathbf{x}\|\,\|\mathbf{y}\|$$

- If $\mathbf{x}$ and $\mathbf{y}$ are unit length vectors, dot product equals the cosine of the angle between the vectors.

$$\mathbf{x}^T\mathbf{y} = \cos\theta$$

- If $\mathbf{x}$ and $\mathbf{y}$ point in the same direction, $\theta = 0$ and the dot product is maximized.

$$\mathbf{x}^T\mathbf{y} = \cos(\theta = 0) = 1$$

- If $\mathbf{x}$ and $\mathbf{y}$ are perpendicular, $\theta = 90^o$ and the dot product becomes zero.

$$\mathbf{x}^T\mathbf{y} = \cos(\theta = 90^o) = 0$$

Second Key Property: Alignment between Vectors

- Therefore, whenever we are taking dot products between two vectors (in discrete time) or two signals (in continuous time), think (or visualize) this as if
  - ▷ we are measuring how well the two signals are aligned with each other in a multi-dimensional space
- In $\mathbf{h} \cdot \mathbf{r}$, one of the two vectors, let us say $\mathbf{h}$, is sometimes fixed (i.e., it makes up a reference), and $\mathbf{r}$ varies
  - ▷ The dot product $\mathbf{h} \cdot \mathbf{r}$ will be high when the variable vector $\mathbf{r}$ happens to be well-aligned to our reference vector $\mathbf{h}$
  - ▷ If the dot product turns out to be zero or near zero, think of this as $\mathbf{r}$ being *orthogonal* to $\mathbf{h}$ (i.e., pointing in a direction that is 90° offset to the direction of $\mathbf{h}$)

Probability Mass Functions or PMFs

- For discrete random variables, the PMF is the collection of all probabilities for different values of X, i.e., $p(x) \overset{\text{def}}{=} P_X(X = x)$
- Properties of the PMFs:

  1. $p(x) \geq 0$
  2. $\displaystyle\sum_X p(x) = 1$
  3. $P(a \leq X \leq b) = \displaystyle\sum_{x=a}^{b} p(x)$

Probability Density Function (PDF)

- PDF is the generalization of the PMF for continuous RVs.
- PDF measures how likely a random variable is to lie at a particular value
- Properties:
  1. $p(x) \geq 0$
  2. $\int_{-\infty}^{\infty} p(x)dx = 1$
  3. $P(a \leq X \leq b) = \int_{a}^{b} p(x)dx$

Cumulative Distribution Function (CDF)

- Definition: $F_X(x) = F(x) = P(X \leq x) = \int_{-\infty}^{x} p_X(y) dy$
- CDF is the integration of PDF; PDF is the derivative of the CDF
- Properties:
  1. $F(x)$ is monotonically nondecreasing
  2. $F(-\infty) = 0$
  3. $F(\infty) = 1$
  4. $P(a < X \leq b) = F(b) - F(a)$

Example PMFs
Binary Distribution

- Outcome of the toss of a fair coin: $p(x) = \begin{cases} 1/2, & x = 0(\text{head}), \\ 1/2, & x = 1(\text{tail}) \end{cases}$

  $\rightarrow$ Mean: $m_x = \sum_x x\, p(x) = 0 \times 1/2 + 1 \times 1/2 = 1/2$

  $\rightarrow$ Variance: $\sigma_x^2 = 1/4$

- If X_1 and X_2 are independent binary random variables,
  $P_{X_1 X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0)$

Example PMFs
Binomial Distribution

- Suppose a biased coin comes up heads with probability $p = 0.3$ when tossed. What is the probability of achieving $0, 1, \ldots, 6$ heads after six tosses?
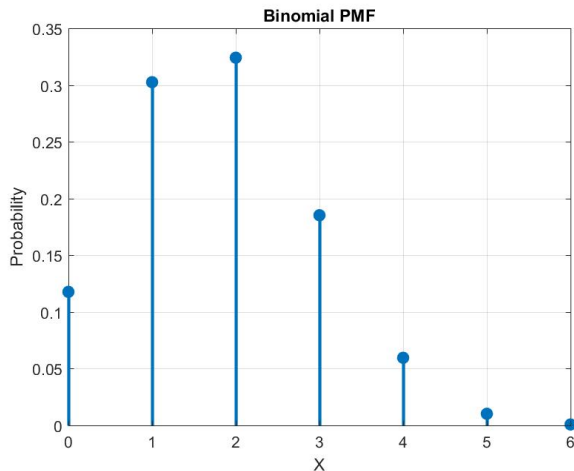- Let heads represent 0 and tails represent 1.

Example PMFs
Binomial Distribution

- Let $Y = \sum_{i=1}^{n} X_i$, where $\{X_i\}$; $i = 1, \ldots, N$ are independent binary

  RVs with $p(x) = \begin{cases} 1 - p, & x = 0(\text{head}), \\ p, & x = 1(\text{tail}) \end{cases}$

- In this case, RV Y follows the Binomial Distribution given as
  $P_Y(y) = \binom{N}{y} p^y (1 - p)^{N-y}$
- Mean $m_y = N \times p$
- Variance $\sigma_y^2 = N \times p \times (1 - p)$

Example PMFs
Binomial PMF

Example PMFs
Understanding Binomial Distribution

- Suppose I perform an experiment N times, where each outcome is a Binary RV, with probability of 1 equal to p and 0 equal to $1 - p$
- Questions:
  - ▷ What are all the possible outcomes?
  - ▷ What is the *sum* of N binary outcomes?
- Let us look at the cases $N = 2, 3$ and 4.

Example PMFs
Binomial PMF

Experiement ID		Binomial RV
1	2	Y
0	0	0
0	1	1
1	0	1
1	1	2

Example PMFs
Binomial PMF

Experiement ID			Binomial RV
1	**2**	**3**	**Y**
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	2
1	0	0	1
1	0	1	2
1	1	0	2
1	1	1	3

Example PMFs
Binomial PMF

Experiement ID				Binomial RV
1	2	3	4	Y
0	0	0	0	0
0	0	0	1	1
0	0	1	0	1
0	0	1	1	2
0	1	0	0	1
0	1	0	1	2
0	1	1	0	2
0	1	1	1	3
1	0	0	0	1
1	0	0	1	2
1	0	1	0	2
1	0	1	1	3
1	1	0	0	2
1	1	0	1	3
1	1	1	0	3
1	1	1	1	4

Example PMFs
Understanding Binomial Distribution

1. Outcomes of N experiments form a binary "string" of length N bits
2. When these N bits are *added up*, we get a variable Y, which is the Binomial RV
3. Value of Y can be any integer between 0 (all-zeros binary string) to N (all-ones string)
4. There is a total of 2^N possible strings (these form a set of all possible outcomes)
5. A total of $\binom{N}{y}$ strings have y ones and $N - y$ zeros. When any of these strings occur, the variable Y will take a value y.
6. Due to independence between N experiments, probability of occurrence of any one of $\binom{N}{y}$ string is the same, and it is given as $p^y(1-p)^{N-y}$.
7. Therefore, the total probability that $Y = y$ is given as $\binom{N}{y}p^y(1-p)^{N-y}$.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.

- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?
  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

  Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?
  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y (1 - p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y (1 - p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

  Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?
  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

  Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?

  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

  Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?

  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

  Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?
  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
    - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
    - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
    - ▷ $N = 10$, $p = 0.1$, $y = 3$.
    - ▷ $\binom{N=10}{y=3} = 120$
    - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
    - ▷ $\binom{N=10}{y=3} p^y (1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.

- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?

    - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
    - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
    - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
    - ▷ $N = 10$, $p = 0.1$, $y = 3$.
    - ▷ $\binom{N=10}{y=3} = 120$
    - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
    - ▷ $\binom{N=10}{y=3} p^y (1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?

    - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
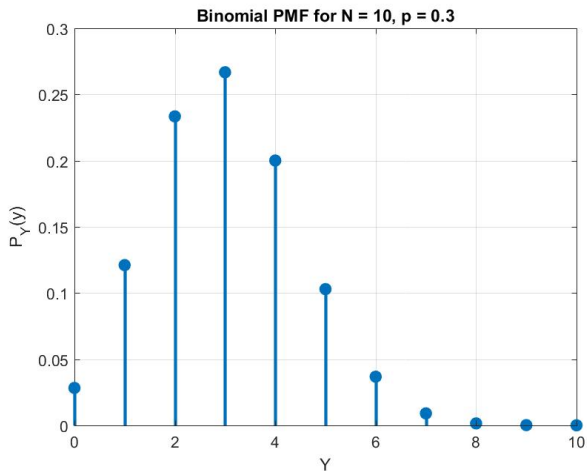Understanding Binomial Distribution

- *You are told:*
    - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
    - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
    - ▷ $N = 10$, $p = 0.1$, $y = 3$.
    - ▷ $\binom{N=10}{y=3} = 120$
    - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
    - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

    Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?

    - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

  Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?
  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Understanding Binomial Distribution

- *You are told:*
  - ▷ At a bicycle company, 70% of the newly made cycles pass the quality inspection and 30% fail.
- *Question:*
  - ▷ What are the chances that 3 of 10 cycles that you're inspecting will fail?
- Answer:
  - ▷ $N = 10$, $p = 0.1$, $y = 3$.
  - ▷ $\binom{N=10}{y=3} = 120$
  - ▷ $p^y(1-p)^{N-y} = 0.3^3 0.7^7 = 0.0022235661$
  - ▷ $\binom{N=10}{y=3} p^y(1-p)^{N-y} = 120 \times 0.0022235661 = 0.2668$

  Therefore, the chances that 30% of $N = 10$ cycles you're inspecting will fail are $\approx 27\%$.
- What are the chances that 2 cycles will fail, or 5 will fail, or all 10 will fail?
  - ▷ Binomial PMF gives the answers to all such questions.

Example PMFs
Binomial PMF

Example PMFs
Binomial PMF: An observation

- Although the failure rate is 30%, there is a good amount of chance that only one out of 10 cycles fails (10% failure rate) or even that none fail (0% failure rate)
- This happens because $N = 10$ is not large enough as the number of independent experiments
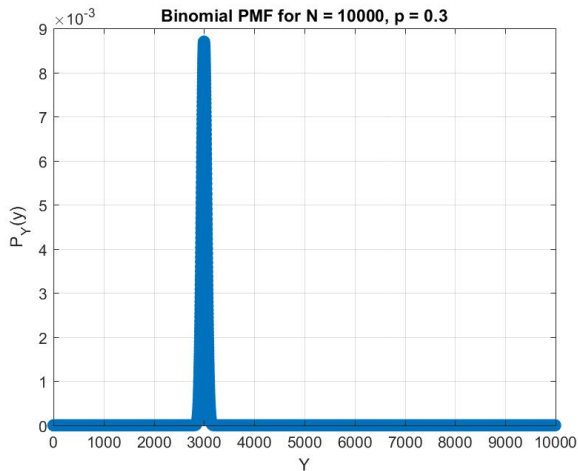 - ▷ Let us see what happens as we increase N

Example PMFs
Binomial PMF
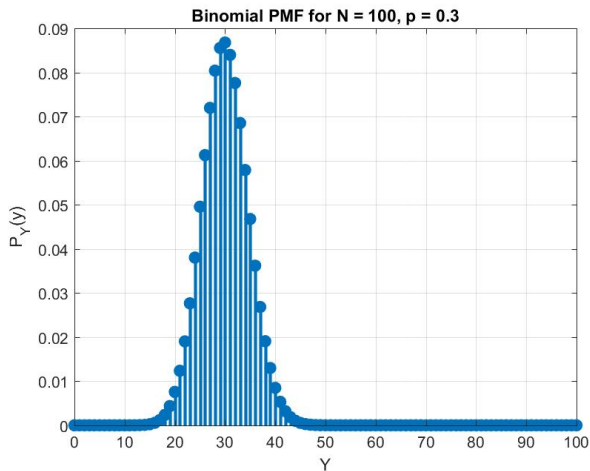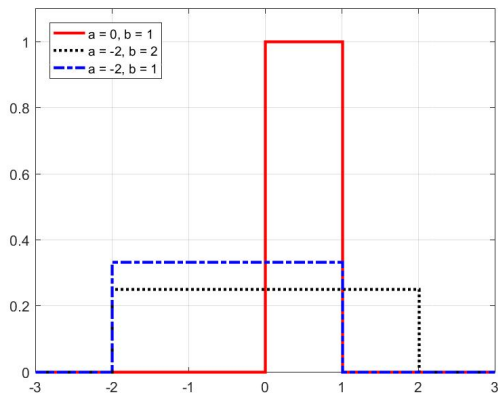
Example PMFs
Binomial PMF
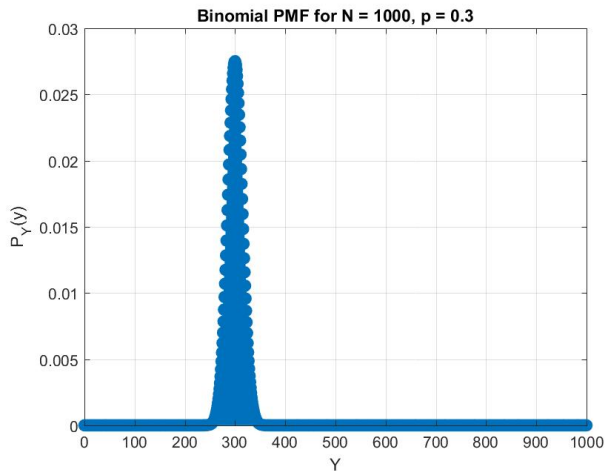
Example PMFs
Binomial PMF

Example PMFs
Binomial PMF: An observation

- As $N \to \infty$, we will indeed (as expected) see with a very high probability that $0.3 \times N$ cycles failing
- Alternatively, the chances that the number of cycles failing is different from $0.3 \times N$ will diminish and become near zero
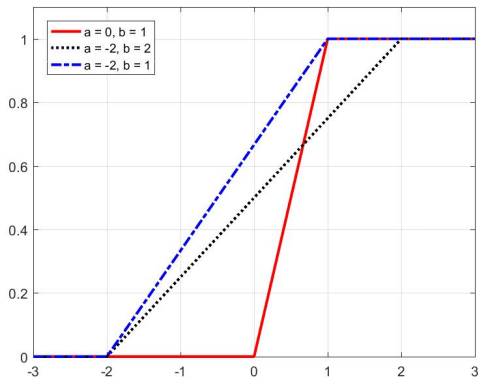
Example RVs
Uniform PDF

- $p(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{else} \end{cases}$

Example PDFs
Uniform CDF

- $F(x) = \begin{cases} 0, & x \leq 0 \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x \geq b \end{cases}$

Example PDFs
Uniform RV

- Mean: $m_x = \int_a^b x\, p(x)\, dx = \frac{1}{b-a} \int_a^b x\, dx = \frac{a+b}{2}$
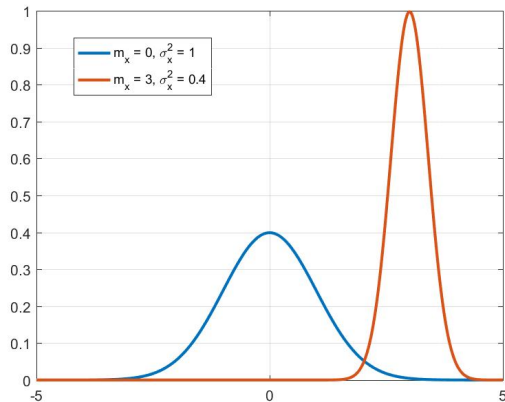- Variance: $\sigma_x^2 = \int_a^b (x - m_x)^2\, p(x)\, dx = \frac{(b-a)^2}{12}$
- Probability:

  $P(a_1 \leq x < b_1) = \int_{a_1}^{b_1} p(x)\, dx = \frac{b_1 - a_1}{b-a}, \quad a < a_1, b_1 < b$
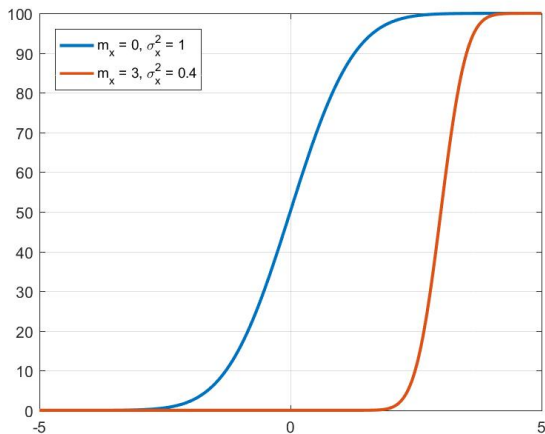
Example RVs
Gaussian PDF

- $p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x - m_x)^2}{2\sigma_x^2}\right)$

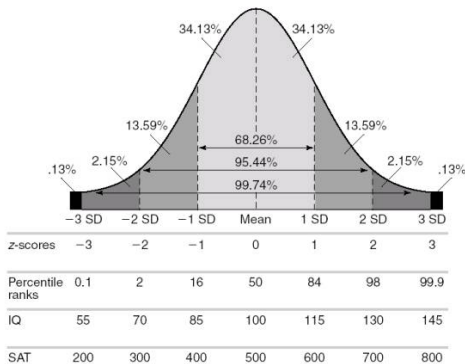# Example RVs
Gaussian CDF

Example RVs
Gaussian PDF

FIGURE 15.8 Percentile ranks and standard scores in relation to the normal curve.

http:
//turtleinvestor888.blogspot.in/2012_01_01_archive.html

Example RVs
Central Limit Theorem

- Let $X_1, X_2, \ldots, X_N$ be N independent RVs with identical PDFs
- Let $Y = \displaystyle\sum_{i=1}^{N} X_i$
- A theorem of probability theory called Central Limit Theorem or CLT: as $N \to \infty$, distribution of Y tends to a Gaussian distribution
  - $\to$ In practice, $N = 10$ is sufficient to see the tendency of Y to follow the Gaussian PDF
- Importance of CLT:
  - $\to$ Thermal noise results from random movements of many electrons, and it is well modeled by the Gaussian PDF
  - $\to$ Interference from many identically distributed interferers in a CDMA system tends toward the Gauassian PDF
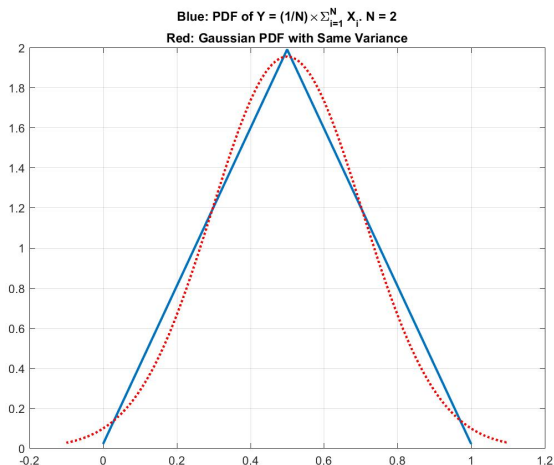
Example RVs
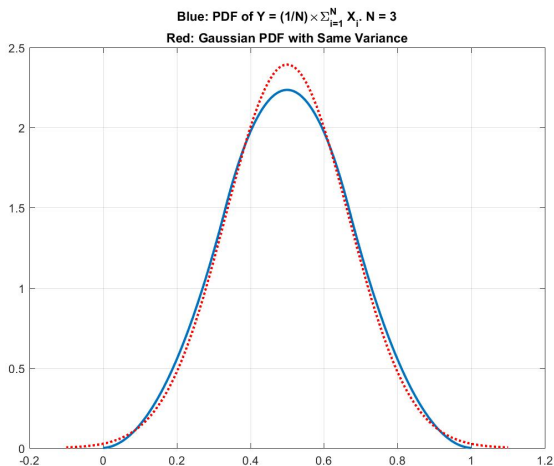## Central Limit Theorem: A Uniform Distribution
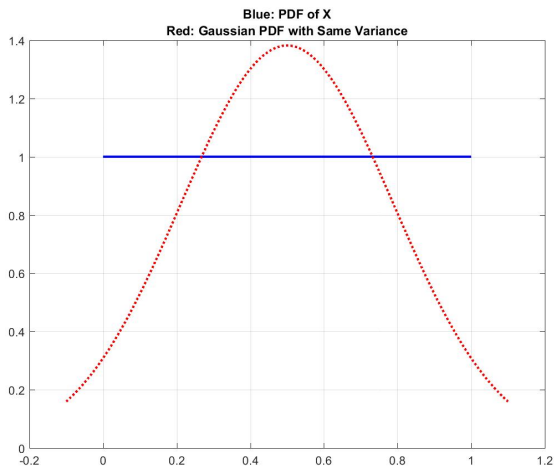
Example RVs
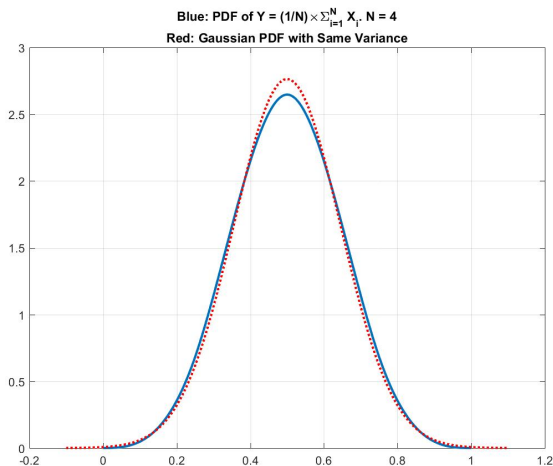Central Limit Theorem. Average of $N = 2$ identically distributed Uniform RVs

Blue: PDF of $Y = (1/N) \times \Sigma_{i=1}^{N} X_i$, $N = 2$
Red: Gaussian PDF with Same Variance

Example RVs
Central Limit Theorem. Average of $N = 3$ identically distributed Uniform RVs

Blue: PDF of Y = (1/N) $\times \Sigma_{i=1}^{N}$ X$_i$, N = 3
Red: Gaussian PDF with Same Variance

Example RVs
Central Limit Theorem. Average of $N = 4$ identically distributed Uniform RVs

Blue: PDF of $Y = (1/N) \times \Sigma_{i=1}^{N} X_i$, $N = 4$
Red: Gaussian PDF with Same Variance

Example RVs
Central Limit Theorem. Average of $N = 10$ identically distributed Uniform RVs

Blue: PDF of $Y = (1/N) \times \Sigma_{i=1}^{N} X_i$. $N = 10$
Red: Gaussian PDF with Same Variance
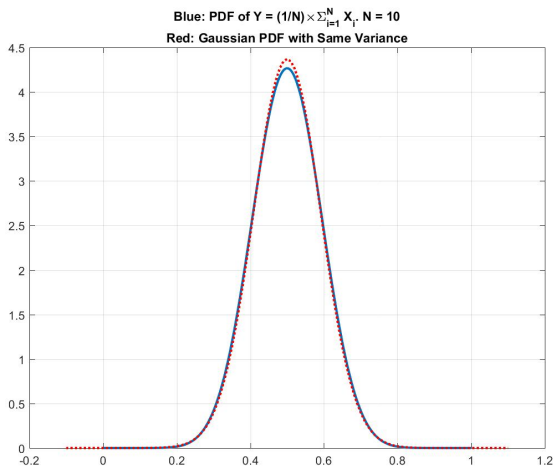
Example RVs
Central Limit Theorem. Average of $N = 19$ identically distributed Uniform RVs

Blue: PDF of Y = (1/N)$\times \Sigma_{i=1}^{N}$ X$_i$. N = 19
Red: Gaussian PDF with Same Variance
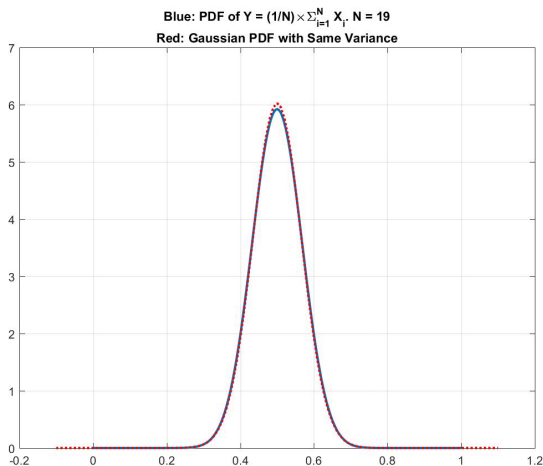
Example RVs
Gaussian PDF

- An application of Gaussian PDFs: signal level at the output of a digital communications receiver can often be given as $r = s + n$, where
  - $\rightarrow$ r is the received signal level,
  - $\rightarrow$ $s = -a$ is the transmitted signal level, and
  - $\rightarrow$ n is the Gaussian noise with mean 0 and variance σ_n^2
- Probability that the signal level $-a$ can be mistaken by the receiver as the signal level $+a$ is given as:

$$P(r > 0) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x+a)^2}{2\sigma_n^2}\right) dx = Q\left(\frac{a}{\sigma_n}\right)$$

- Definition: $Q(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_x^\infty \exp\left(-\frac{v^2}{2}\right) dv$

Example RVs
Rayleigh PDF

- Suppose $r = \sqrt{x_1^2 + x_2^2}$, where x_1 and x_2 are Gaussian with zero mean and variance σ^2
- PDF $p(r) = \frac{r}{\sigma^2} \exp{-\frac{r^2}{2\sigma^2}}$ is the Rayleigh PDF
- Used to model fading when no line of sight is present