Introduction to Probability

Chandrashekar L

Department of Computer Science and Automation
Indian Institute of Science
chandrul@csa.iisc.ernet.in

What is the Probability of obtaining

- Atleast 1 Tail in 2 Tosses of a coin.

What is the Probability of obtaining

- Atleast 1 Tail in 2 Tosses of a coin.

$$\frac{3}{4} \tag{1}$$

- An odd number in a single roll of a die.

What is the Probability of obtaining

- Atleast 1 Tail in 2 Tosses of a coin.

$$\frac{3}{4} \tag{1}$$

- An odd number in a single roll of a die.

$$\frac{3}{6} = \frac{1}{2} \tag{2}$$

- A 'Spade' in a pack of shuffled cards.

What is the Probability of obtaining

- Atleast 1 Tail in 2 Tosses of a coin.

$$\frac{3}{4} \tag{1}$$

- An odd number in a single roll of a die.

$$\frac{3}{6} = \frac{1}{2} \tag{2}$$

- A 'Spade' in a pack of shuffled cards.

$$\frac{13}{52} = \frac{1}{4} \tag{3}$$

2 Tosses of a coin

- What all can come??

2 Tosses of a coin

- What all can come??
  HH, **HT**, **TH**, **TT**

2 Tosses of a coin

- What all can come??
  HH, **HT**, **TH**, **TT**
- What is that we are interested in??

2 Tosses of a coin

- What all can come??
  HH, **HT**, **TH**, **TT**
- What is that we are interested in??
  HT, **TH**, **TT**

2 Tosses of a coin

- What all can come??
  HH, **HT**, **TH**, **TT**
- What is that we are interested in??
  HT, **TH**, **TT**
- Probability of *Event* of interest

2 Tosses of a coin

- What all can come??
  HH, **HT**, **TH**, **TT**
- What is that we are interested in??
  HT, **TH**, **TT**
- Probability of *Event* of interest

$$\frac{\text{number of favourable outcomes}}{\text{number of all possible outcomes}} \tag{4}$$

Some Terminology

- What all can come??

- What all can come??

Sample Space **S**={{**HH**},{**HT**},{**TH**},{**TT**}}

- What all can come??

  Sample Space $\mathbf{S}=\{\{\mathbf{HH}\},\{\mathbf{HT}\},\{\mathbf{TH}\},\{\mathbf{TT}\}\}$

- What is that we may be interested in??

- What all can come??

  Sample Space **S**=$\{\{$**HH**$\}$**,**$\{$**HT**$\}$**,**$\{$**TH**$\}$**,**$\{$**TT**$\}\}$

- What is that we may be interested in??

  Set of Events **E**

- What all can come??

Sample Space **S**=$\{\{$**HH**$\}$,$\{$**HT**$\}$,$\{$**TH**$\}$,$\{$**TT**$\}\}$

- What is that we may be interested in??

Set of Events **E**

$$\mathbf{E} = \Big\{\{\{\mathbf{HH}\}\}, \{\{\mathbf{HT}\}\}, \{\{\mathbf{TH}\}\}, \{\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{TH}\}\}, \{\{\mathbf{HH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HT}\},\{\mathbf{TH}\}\}, \{\{\mathbf{HT}\},\{\mathbf{TT}\}\}, \{\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\},\{\mathbf{TH}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HT}\},\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\},\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\Phi\}\Big\}$$

- What all can come??

Sample Space **S**={{**HH**},{**HT**},{**TH**},{**TT**}}

- What is that we may be interested in??

Set of Events **E**

E = {{{**HH**}}, {{**HT**}}, {{**TH**}},{{**TT**}}, {{**HH**},{**HT**}},
{{**HH**},{**TH**}},{{**HH**},{**TT**}},{{**HT**},{**TH**}}, {{**HT**},{**TT**}},
{{**TH**},{**TT**}},{{**HH**},{**HT**},{**TH**}},{{**HH**},{**HT**},{**TT**}},
{{**HH**},{**TH**},{**TT**}},{{**HT**},{**TH**},{**TT**}},
{{**HH**},{**HT**},{**TH**},{**TT**}},{Φ}}

- Probability of *Event* of interest

- What all can come??

Sample Space **S**=$\{$$\{$**HH**$\}$,$\{$**HT**$\}$,$\{$**TH**$\}$,$\{$**TT**$\}$$\}$

- What is that we may be interested in??

Set of Events **E**

$$\mathbf{E} = \Big\{ \{\{\mathbf{HH}\}\}, \{\{\mathbf{HT}\}\}, \{\{\mathbf{TH}\}\}, \{\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{TH}\}\}, \{\{\mathbf{HH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HT}\},\{\mathbf{TH}\}\}, \{\{\mathbf{HT}\},\{\mathbf{TT}\}\}, \{\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\},\{\mathbf{TH}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HT}\},\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\{\mathbf{HH}\},\{\mathbf{HT}\},\{\mathbf{TH}\},\{\mathbf{TT}\}\}, \{\Phi\} \Big\}$$

- Probability of *Event* of interest

The probability assignment $\mathbf{P} : \mathbf{E} \rightarrow [0, 1]$

How to put an elephant inside a refrigerator?

How to put an elephant inside a refrigerator?

- Open the refrigerator.

How to put an elephant inside a refrigerator?

- Open the refrigrator.
- Put the Elephant inside.

How to put an elephant inside a refrigerator?

- Open the refrigerator.
- Put the Elephant inside.
- Close the refrigerator.

How to put an elephant inside a refrigrator?

- Open the refrigrator.
- Put the Elephant inside.
- Close the refrigrator.

How to compute probabilities

- Identify **S**.

How to put an elephant inside a refrigrator?

- Open the refrigrator.
- Put the Elephant inside.
- Close the refrigrator.

How to compute probabilities

- Identify **S**.
- Identify **E**.

How to put an elephant inside a refrigrator?

- Open the refrigrator.
- Put the Elephant inside.
- Close the refrigrator.

How to compute probabilities

- Identify **S**.
- Identify **E**.
- Do the assignment **P**.

How to put an elephant inside a refrigrator?

- Open the refrigrator.
- Put the Elephant inside.
- Close the refrigrator.

How to compute probabilities

- Identify **S**.
- Identify **E**.
- Do the assignment **P**.

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

1. $0 \leq \mathbf{P}(E_1) \leq 1$, for any event $E_1 \in \mathbf{E}$.

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

1. $0 \leq \mathbf{P}(E_1) \leq 1$, for any event $E_1 \in \mathbf{E}$.
2. $\mathbf{P}(\mathbf{S}) = 1$.

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

1. $0 \leq \mathbf{P}(E_1) \leq 1$, for any event $E_1 \in \mathbf{E}$.
2. $\mathbf{P}(\mathbf{S}) = 1$.
3. For any sequence $E_1, E_2, \ldots$ of disjoint sets (*mutually exclusive*) we have

$$\mathbf{P}(\cup_i E_i) = \sum_i \mathbf{P}(E_i). \tag{5}$$

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

1. $0 \leq \mathbf{P}(E_1) \leq 1$, for any event $E_1 \in \mathbf{E}$.
2. $\mathbf{P}(\mathbf{S}) = 1$.
3. For any sequence $E_1, E_2, \ldots$ of disjoint sets (*mutually exclusive*) we have

$$\mathbf{P}(\bigcup_i E_i) = \sum_i \mathbf{P}(E_i). \tag{5}$$

4. The set **E** satisfies the following conditions
    1. $E_1 \in \mathbf{E} \Rightarrow E_1^c \in \mathbf{E}$.

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

1. $0 \leq \mathbf{P}(E_1) \leq 1$, for any event $E_1 \in \mathbf{E}$.
2. $\mathbf{P}(\mathbf{S}) = 1$.
3. For any sequence $E_1, E_2, \ldots$ of disjoint sets (*mutually exclusive*) we have

$$\mathbf{P}(\cup_i E_i) = \sum_i \mathbf{P}(E_i). \tag{5}$$

4. The set **E** satisfies the following conditions
    1. $E_1 \in \mathbf{E} \Rightarrow E_1^c \in \mathbf{E}$.
    2. $E_1, E_2 \in \mathbf{E} \Rightarrow (E_1 \cap E_2) \in \mathbf{E}$.

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

1. $0 \le \mathbf{P}(E_1) \le 1$, for any event $E_1 \in \mathbf{E}$.
2. $\mathbf{P}(\mathbf{S}) = 1$.
3. For any sequence $E_1, E_2, \ldots$ of disjoint sets (*mutually exclusive*) we have

$$\mathbf{P}(\underset{i}{\cup} E_i) = \sum_i \mathbf{P}(E_i). \quad (5)$$

4. The set **E** satisfies the following conditions
   1. $E_1 \in \mathbf{E} \Rightarrow E_1^c \in \mathbf{E}$.
   2. $E_1, E_2 \in \mathbf{E} \Rightarrow (E_1 \cap E_2) \in \mathbf{E}$.
   3. For any sequence $E_1, E_2, \ldots \in \mathbf{E} \Rightarrow \underset{i}{\cup} E_i \in \mathbf{E}$.

A probability space is a 3-tuple (**S**,**E**,**P**) with the following conditions

1. $0 \leq \mathbf{P}(E_1) \leq 1$, for any event $E_1 \in \mathbf{E}$.
2. $\mathbf{P}(\mathbf{S}) = 1$.
3. For any sequence $E_1, E_2, \ldots$ of disjoint sets (*mutually exclusive*) we have

$$\mathbf{P}(\underset{i}{\cup} E_i) = \sum_i \mathbf{P}(E_i). \tag{5}$$

4. The set **E** satisfies the following conditions
   1. $E_1 \in \mathbf{E} \Rightarrow E_1^c \in \mathbf{E}$.
   2. $E_1, E_2 \in \mathbf{E} \Rightarrow (E_1 \cap E_2) \in \mathbf{E}$.
   3. For any sequence $E_1, E_2, \ldots \in \mathbf{E} \Rightarrow \underset{i}{\cup} E_i \in \mathbf{E}$.

Certain Properties

- If $E_1, E_2 \in \mathbf{E}$, with $E_1 \subset E_2$, then $\mathbf{P}(E_1) \leq \mathbf{P}(E_2)$.

- If $E_1, E_2 \in \mathbf{E}$, with $E_1 \subset E_2$, then $\mathbf{P}(E_1) \leq \mathbf{P}(E_2)$.
- $\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2) - \mathbf{P}(E_1 \cap E_2)$.

- If $E_1, E_2 \in \mathbf{E}$, with $E_1 \subset E_2$, then $\mathbf{P}(E_1) \leq \mathbf{P}(E_2)$.
- $\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2) - \mathbf{P}(E_1 \cap E_2)$.
- $\mathbf{P}(\mathbf{S}) = 1$.

Certain Properties

- If $E_1, E_2 \in \mathbf{E}$, with $E_1 \subset E_2$, then $\mathbf{P}(E_1) \leq \mathbf{P}(E_2)$.
- $\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2) - \mathbf{P}(E_1 \cap E_2)$.
- $\mathbf{P}(\mathbf{S}) = 1$.
- $\mathbf{P}(\Phi) = 0$.

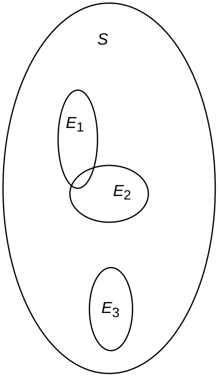


Figure: Abstract Probability Space

- World is not as simple as Coins, Dice and Cards.

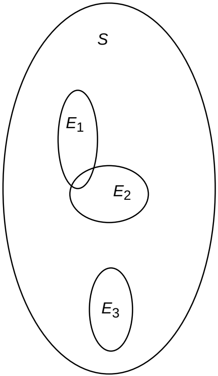
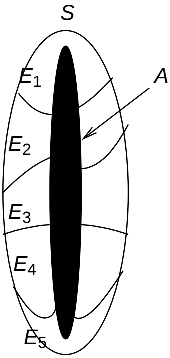
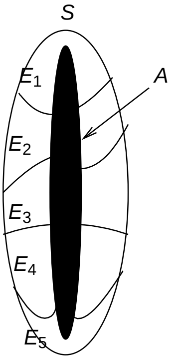


Figure: Abstract Probability Space

- World is not as simple as Coins, Dice and Cards.
- Any abstraction is just as useful as $(a + b)^2 = a^2 + b^2 + 2ab$.

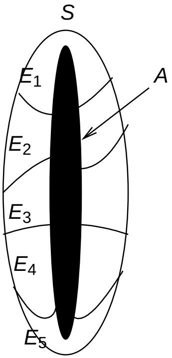


Here we are interested in probabilities of events given an event A has occured.



Here we are interested in probabilities of events given an event A has occured.

$$\mathbf{P}(E_1|A) = \frac{\mathbf{P}(E_1 \cap A)}{\mathbf{P}(A)} \tag{6}$$



Here we are interested in probabilities of events given an event A has occured.

$$\mathbf{P}(E_1|A) = \frac{\mathbf{P}(E_1 \cap A)}{\mathbf{P}(A)} \tag{6}$$

Events E_1 and E_2 are independent when

$$\mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1) \times \mathbf{P}(E_2). \tag{7}$$

Events E_1 and E_2 are independent when

$$\mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1) \times \mathbf{P}(E_2). \tag{7}$$

What is the probability of obtaining atleast one 'Tail' when a coin is Tossed twice. The probability of obtaining 'Tail' in a single Toss is p.

Events E_1 and E_2 are independent when

$$\mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1) \times \mathbf{P}(E_2). \tag{7}$$

What is the probability of obtaining atleast one ‘Tail’ when a coin is Tossed twice. The probability of obtaining ‘Tail’ in a single Toss is p.

$$\mathbf{P}(\{\{\mathbf{TT}\},\{\mathbf{HT}\},\{\mathbf{TH}\}\}) = p \times p + (1-p) \times p + p \times (1-p) \tag{8}$$

Events E_1 and E_2 are independent when

$$\mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1) \times \mathbf{P}(E_2). \tag{7}$$

What is the probability of obtaining atleast one ‘Tail’ when a coin is Tossed twice. The probability of obtaining ‘Tail’ in a single Toss is p.

$$\mathbf{P}(\{\{\mathbf{TT}\},\{\mathbf{HT}\},\{\mathbf{TH}\}\}) = p \times p + (1-p) \times p + p \times (1-p) \tag{8}$$

Examples of independent events

- Getting ‘Tail’ in the first toss is independent of getting ‘Head’ in the second toss.
- When I toss a coin and roll a die simultaneously the outcomes are independent of each other.
- How many glasses of water I drink is independent of whether it will rain Tomorrow.

The above examples are **Correct** but **Useless**

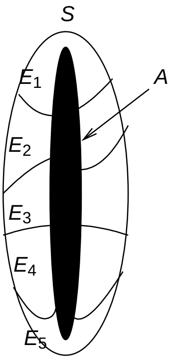
$E_1, E_2, \ldots, E_n$ be disjoint sets such that $\bigcup_i E_i = \mathbf{S}$, let A be any event

$$\mathbf{P}(A) = \mathbf{P}(A|E_1) \times \mathbf{P}(E_1) + \mathbf{P}(A|E_2) \times \mathbf{P}(E_2) + \ldots + \mathbf{P}(A|E_n) \times \mathbf{P}(E_n) \tag{9}$$

$E_1, E_2, \ldots, E_n$ be disjoint sets such that $\bigcup_i E_i = \mathbf{S}$, let A be any event

$$\mathbf{P}(A) = \mathbf{P}(A|E_1) \times \mathbf{P}(E_1) + \mathbf{P}(A|E_2) \times \mathbf{P}(E_2) + \ldots + \mathbf{P}(A|E_n) \times \mathbf{P}(E_n) \tag{9}$$

This is similar to the *cut-off* mark calulations.

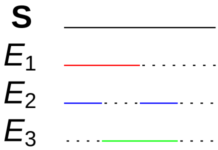


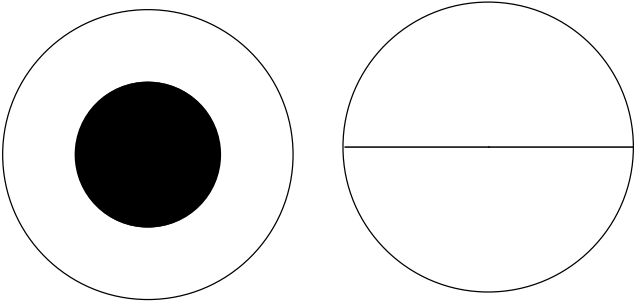
This leads to the following relation

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B|A) \times \mathbf{P}(A)}{\mathbf{P}(B|A) \times \mathbf{P}(A) + \mathbf{P}(B|A^c) \times \mathbf{P}(A^c)}. \tag{10}$$

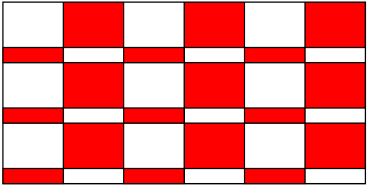
Independent Events Contd

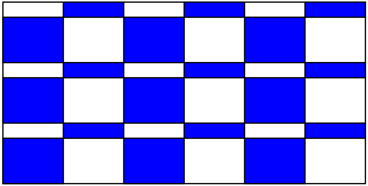
$$\mathbf{S} = [0,1], E_1 = [0, \frac{1}{2}], E_2 = [0, \frac{1}{4}] \cup [\frac{1}{2}, \frac{3}{4}], E_3 = [\frac{1}{4}, \frac{3}{4}]. \quad (11)$$





Independent Events Contd





A random variable is not

- Random.

A random variable is not

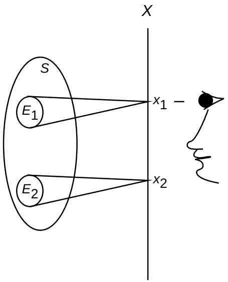
- Random.
- A Variable.

What is it?

A random variable is not

- Random.
- A Variable.

What is it?
It is a function $X : \mathbf{S} \to \mathbf{R}$.



Rs.1 on getting an odd number and 2 on getting a even number on roll of a die. We can define X as follows

Rs.1 on getting an odd number and 2 on getting a even number on roll of a die. We can define X as follows

$$\begin{aligned}
\mathbf{S} &= \{1, 2, 3, 4, 5, 6\} \\
X(1) &= 1 \\
X(2) &= 2 \\
X(3) &= 1 \\
X(4) &= 2 \\
X(5) &= 1 \\
X(6) &= 2
\end{aligned} \tag{12}$$

One can define another R.V. Y where Rs.$x/2$ when x shows up.

One can define another R.V. Y where Rs.$x/2$ when x shows up.

$$
\begin{aligned}
\mathbf{S} &= \{1, 2, 3, 4, 5, 6\} \\
Y(1) &= \frac{1}{2} \\
Y(2) &= 1 \\
Y(3) &= \frac{3}{2} \\
Y(4) &= 2 \\
Y(5) &= \frac{5}{2} \\
Y(6) &= 3
\end{aligned} \tag{13}
$$

$\mathbf{P}(X = 2)$, $\mathbf{P}(X = 1, Y = 1)$, $\mathbf{P}(X = 1, Y = -1)$.

Probability Mass Function

Associated with a R.V. X we are interested in its *probability mass function* (pmf)

Probability Mass Function

Associated with a R.V. X we are interested in its *probability mass function* (pmf)

$f_X(x) = \mathbf{P}(X = x)$

Associated with a R.V. X we are interested in its *probability mass function* (pmf)

$f_X(x) = \mathbf{P}(X = x)$

Let us compute the pmfs for X and Y.

$$f_X(1) = \mathbf{P}(X = 1) = \frac{1}{2}, f_X(2) = \mathbf{P}(X = 2) = \frac{1}{2} \tag{14}$$

$$\begin{aligned}
f_Y(1) = \mathbf{P}(Y = 1) &= \frac{1}{2}, & f_Y(2) = \mathbf{P}(Y = 2) = && \frac{1}{2} \\
f_Y(3) = \mathbf{P}(Y = 3) &= \frac{1}{2}, & f_Y(4) = \mathbf{P}(Y = 4) = && 0 \\
f_Y(5) = \mathbf{P}(Y = 5) &= 0, & f_Y(6) = \mathbf{P}(Y = 6) = && 0 \\
f_Y(\frac{1}{2}) = \mathbf{P}(Y = \frac{1}{2}) &= \frac{1}{2}, & f_Y(\frac{3}{2}) = \mathbf{P}(Y = \frac{3}{2}) = && \frac{1}{2} \\
f_Y(\frac{5}{2}) = \mathbf{P}(Y = \frac{5}{2}) &= \frac{1}{2} & &&
\end{aligned} \tag{15}$$

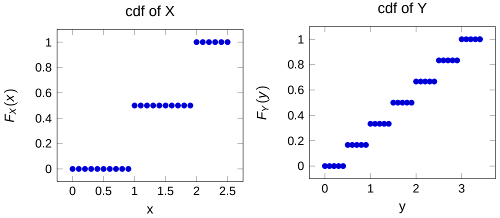
$f(x)$ can also be thought of as the frequency of occurence of x.

Cumulative Distribution Function

Sometimes we are interested in the quantity $F(x) = \mathbf{P}(X \leq x)$.

Cumulative Distribution Function

Sometimes we are interested in the quantity $F(x) = \mathbf{P}(X \leq x)$.
This is the *cumulative distribution function*



Bernoulli Random Variable X is either *Success*-1 or *Failure*-0. Thus X takes values 1 or 0.

$$\mathbf{P}(X = 0) = 1 - p,$$
$$\mathbf{P}(X = 1) = p, 0 \leq p \leq 1.$$

Bernoulli Random Variable X is either *Success*-1 or *Failure*-0. Thus X takes values 1 or 0.

$$\mathbf{P}(X = 0) = 1 - p,$$
$$\mathbf{P}(X = 1) = p, 0 \leq p \leq 1.$$

Binomial Random Variable X is the number *Successes* in n-independent trials. X takes values from 1 to n with

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}. \tag{16}$$

Bernoulli Random Variable X is either *Success*-1 or *Failure*-0. Thus X takes values 1 or 0.

$$\mathbf{P}(X = 0) = 1 - p,$$
$$\mathbf{P}(X = 1) = p, 0 \leq p \leq 1.$$

Binomial Random Variable X is the number *Successes* in n-independent trials. X takes values from 1 to n with

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}. \tag{16}$$

Geometric Random Variable X is the number of trials needed to obtain *Success*. X takes values $1, 2, 3, \ldots$ with

$$\mathbf{P}(x = k) = (1 - p)^{k-1} p. \tag{17}$$

Bernoulli Random Variable X is either *Success*-1 or *Failure*-0. Thus X takes values 1 or 0.

$$\mathbf{P}(X = 0) = 1 - p,$$
$$\mathbf{P}(X = 1) = p, 0 \leq p \leq 1.$$

Binomial Random Variable X is the number *Successes* in n-independent trials. X takes values from 1 to n with

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}. \tag{16}$$

Geometric Random Variable X is the number of trials needed to obtain *Success*. X takes values $1, 2, 3, \ldots$ with

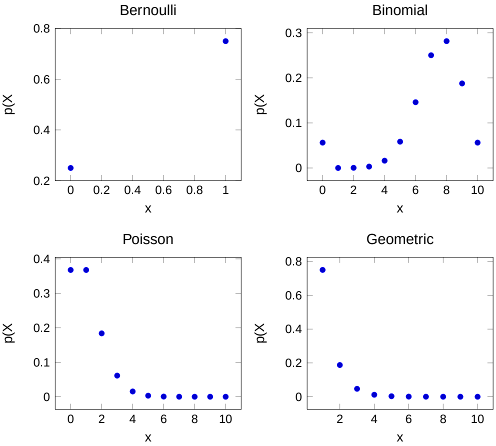
$$\mathbf{P}(x = k) = (1 - p)^{k-1} p. \tag{17}$$

Poisson Random Variable X takes values $0, 1, 2, 3, \ldots$ with

$$\mathbf{P}(X = k) = \exp(-\lambda) \frac{(\lambda)^k}{k!} \tag{18}$$

Useful to model number of arrivals in unit time.

Useful Random Variables



- Till now the Random Variables assumed only discrete values (finite or countable).

- Till now the Random Variables assumed only discrete values (finite or countable).
- Consider a Random Variable that takes a *continuum* of values.

Continuous Random Variables

- Till now the Random Variables assumed only discrete values (finite or countable).
- Consider a Random Variable that takes a *continuum* of values.
- Then $\mathbf{P}(X = x)$ is 0 for all x and pmf does not make sense.

- Till now the Random Variables assumed only discrete values (finite or countable).
- Consider a Random Variable that takes a *continuum* of values.
- Then $\mathbf{P}(X = x)$ is 0 for all x and pmf does not make sense.
- However the cdf still makes sense.

Continuous Random Variables

- Till now the Random Variables assumed only discrete values (finite or countable).
- Consider a Random Variable that takes a *continuum* of values.
- Then $\mathbf{P}(X = x)$ is 0 for all x and pmf does not make sense.
- However the cdf still makes sense.
- Equivalent to the pmf we define the *probability density function* (pdf).

- Till now the Random Variables assumed only discrete values (finite or countable).
- Consider a Random Variable that takes a *continuum* of values.
- Then $\mathbf{P}(X = x)$ is 0 for all x and pmf does not make sense.
- However the cdf still makes sense.
- Equivalent to the pmf we define the *probability density function* (pdf).
- The pdf is derivative of the cdf.

Uniform R.V X, where X takes values in (a, b).

Uniform R.V X, where X takes values in (a, b).
$f(x) = \frac{1}{b-a}, x \in (a, b)$ and $f_X(x) = 0, x \notin (a, b)$

Uniform R.V X, where X takes values in (a, b).
$f(x) = \frac{1}{b-a}, x \in (a, b)$ and $f_X(x) = 0, x \notin (a, b)$
Gaussian R.V, where X takes values from $-\infty$ to $+\infty$.

Uniform R.V X, where X takes values in (a, b).
$f(x) = \frac{1}{b-a}, x \in (a, b)$ and $f_X(x) = 0, x \notin (a, b)$
Gaussian R.V, where X takes values from $-\infty$ to $+\infty$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{\frac{x-\mu}{2\sigma^2}^2} \tag{19}$$

Exponential R.V, where X takes values in $[0, \infty)$,

Uniform R.V X, where X takes values in (a, b).
$f(x) = \frac{1}{b-a}, x \in (a, b)$ and $f_X(x) = 0, x \notin (a, b)$
Gaussian R.V, where X takes values from $-\infty$ to $+\infty$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{\frac{x-\mu^{2}}{2\sigma^2}} \tag{19}$$

Exponential R.V, where X takes values in $[0, \infty)$,

$$f(x) = \lambda exp^{-\lambda x} \forall x \geq 0. \tag{20}$$

Gamma R.V, where X takes values in $[0, \infty)$,

Uniform R.V X, where X takes values in (a, b).
$f(x) = \frac{1}{b-a}, x \in (a, b)$ and $f_X(x) = 0, x \notin (a, b)$
Gaussian R.V, where X takes values from $-\infty$ to $+\infty$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{\frac{x-\mu}{2\sigma^2}^2} \quad (19)$$

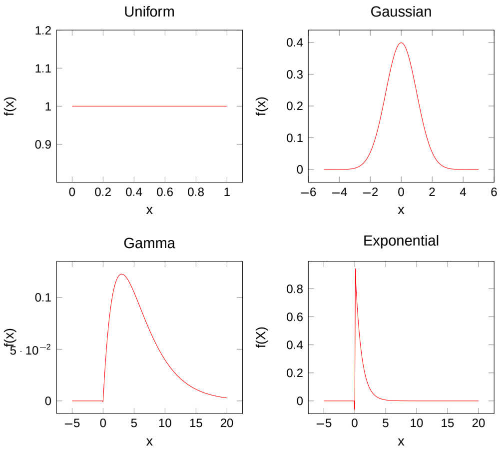
Exponential R.V, where X takes values in $[0, \infty)$,

$$f(x) = \lambda exp^{-\lambda x} \forall x \geq 0. \quad (20)$$

Gamma R.V, where X takes values in $[0, \infty)$,

$$f(x) = \frac{\lambda exp^{-\lambda x} (\lambda x)^{(n-1)}}{(n-1)!}. \quad (21)$$

Useful Continuous Random Variables



If X is a D.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \sum_x x\mathbf{P}(X = x). \tag{22}$$

If X is a D.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \sum_{x} x\mathbf{P}(X = x). \tag{22}$$

If X is a C.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} xf(x)dx. \tag{23}$$

If X is a D.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \sum_{x} x\mathbf{P}(X = x). \tag{22}$$

If X is a C.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} xf(x)dx. \tag{23}$$

Properties of Mean

- $\mathbf{E}(g(X)) = \sum_{x} g(x)\mathbf{P}(X = x)$ for D.R.V.,

  $\mathbf{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx.$

If X is a D.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \sum_{x} x\mathbf{P}(X = x). \tag{22}$$

If X is a C.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} xf(x)dx. \tag{23}$$

Properties of Mean

- $\mathbf{E}(g(X)) = \sum_{x} g(x)\mathbf{P}(X = x)$ for D.R.V.,
  $\mathbf{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx.$
- $\mathbf{E}(aX + bY) = a\mathbf{E}(X) + b\mathbf{E}(Y).$

If X is a D.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \sum_{x} x\mathbf{P}(X = x). \tag{22}$$

If X is a C.R.V. , the mean or *Expectation* of X is

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} xf(x)dx. \tag{23}$$

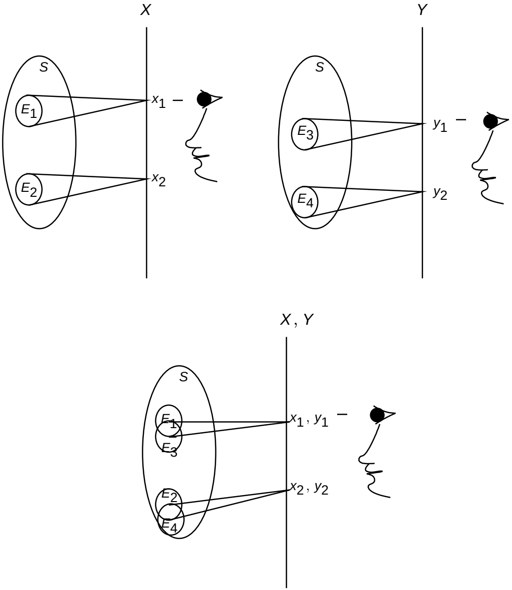
Properties of Mean

- $\mathbf{E}(g(X)) = \sum_{x} g(x)\mathbf{P}(X = x)$ for D.R.V.,
  $\mathbf{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx.$
- $\mathbf{E}(aX + bY) = a\mathbf{E}(X) + b\mathbf{E}(Y).$

- The Variance of X is given by $Var(X) = \mathbf{E}[(X - \mathbf{E}(X))^2]$

- The Variance of X is given by $Var(X) = \mathbf{E}[(X - \mathbf{E}(X))^2]$
- $Var(X)$ is always positive.
- Verify that $Var(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2$.

Two Random Variables



We have *joint* cdf, pmf, pdf defined as follows

- Joint cdf $F_{XY}(x, y) = \mathbf{P}(X \leq x, Y \leq y)$.

We have *joint* cdf, pmf, pdf defined as follows

- Joint cdf $F_{XY}(x, y) = \mathbf{P}(X \leq x, Y \leq y)$.
- Joint pmf meaning $f_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$.

We have *joint* cdf, pmf, pdf defined as follows

- Joint cdf $F_{XY}(x,y) = \mathbf{P}(X \leq x, Y \leq y)$.
- Joint pmf meaning $f_{XY}(x,y) = \mathbf{P}(X = x, Y = y)$.
- Joint pdf $f_{XY}(x,y)$, where $F_{XY}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(x,y)dxdy$.

We have *joint* cdf, pmf, pdf defined as follows

- Joint cdf $F_{XY}(x,y) = \mathbf{P}(X \le x, Y \le y)$.
- Joint pmf meaning $f_{XY}(x,y) = \mathbf{P}(X = x, Y = y)$.
- Joint pdf $f_{XY}(x,y)$, where $F_{XY}(x,y) = \int\limits_{-\infty}^{y} \int\limits_{-\infty}^{x} f_{XY}(x,y)dxdy$.

We have the *marginal* cdfs, pdfs and pmfs as follows.

- $F_X(x) = \mathbf{P}(X \le x) = \int\limits_{-\infty}^{\infty} f_{XY}(x,y)dy$ for a C.R.V.

We have *joint* cdf, pmf, pdf defined as follows

- Joint cdf $F_{XY}(x, y) = \mathbf{P}(X \le x, Y \le y)$.
- Joint pmf meaning $f_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$.
- Joint pdf $f_{XY}(x, y)$, where $F_{XY}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(x, y) dx dy$.

We have the *marginal* cdfs, pdfs and pmfs as follows.

- $F_X(x) = \mathbf{P}(X \le x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$ for a C.R.V.
- $f_X(x) = \mathbf{P}(X = x) = \sum_y \mathbf{P}(X = x, Y = y)$. for a D.R.V.

We have *joint* cdf, pmf, pdf defined as follows

- Joint cdf $F_{XY}(x, y) = \mathbf{P}(X \leq x, Y \leq y)$.
- Joint pmf meaning $f_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$.
- Joint pdf $f_{XY}(x, y)$, where $F_{XY}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(x, y) dx dy$.

We have the *marginal* cdfs, pdfs and pmfs as follows.

- $F_X(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$ for a C.R.V.
- $f_X(x) = \mathbf{P}(X = x) = \sum_y \mathbf{P}(X = x, Y = y)$. for a D.R.V.

Given the Marginals the joint distribution cannot be determined in general.

We have *joint* cdf, pmf, pdf defined as follows

- Joint cdf $F_{XY}(x, y) = \mathbf{P}(X \leq x, Y \leq y)$.
- Joint pmf meaning $f_{XY}(x, y) = \mathbf{P}(X = x, Y = y)$.
- Joint pdf $f_{XY}(x, y)$, where $F_{XY}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(x, y) dx dy$.

We have the *marginal* cdfs, pdfs and pmfs as follows.

- $F_X(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$ for a C.R.V.
- $f_X(x) = \mathbf{P}(X = x) = \sum_{y} \mathbf{P}(X = x, Y = y)$. for a D.R.V.

Given the Marginals the joint distribution cannot be determined in general.
We also can talk about the *conditional* $f_{X|Y=y}(x)$, distribution of X given that Y assumes a value y.

- Joint cdf

$$
\begin{aligned}
F_{XY}(x,y) &= \mathbf{P}(X \le x, Y \le y) \\
&= \mathbf{P}(X \le x)\mathbf{P}(Y \le y) \\
&= F_X(x)F_Y(y)
\end{aligned}
$$

- Joint cdf

$$
\begin{aligned}
F_{XY}(x,y) &= \mathbf{P}(X \le x, Y \le y) \\
&= \mathbf{P}(X \le x)\mathbf{P}(Y \le y) \\
&= F_X(x)F_Y(y)
\end{aligned}
$$

- The pdf (also the pmf) $f_{XY}(x,y) = f_X(x)f_Y(y)$.

In case of independent R.Vs

- We can determine the joints given the marginals.
- The conditional distribution is same as the marginal.

Two R.Vs X and Y are said to be identically distributed if

Two R.Vs X and Y are said to be identically distributed if $F_X(x) = F_Y(x)$.

Consider the following R.Vs defined on $\mathbf{S} = [0, 1]$.

Two R.Vs X and Y are said to be identically distributed if $F_X(x) = F_Y(x)$.
Consider the following R.Vs defined on $\mathbf{S} = [0, 1]$.
$X(s) = 1, s \in [0, 0.5), X(s) = 0, s \in [0.5, 1]$.

Two R.Vs X and Y are said to be identically distributed if $F_X(x) = F_Y(x)$.
Consider the following R.Vs defined on $\mathbf{S} = [0, 1]$.
$X(s) = 1, s \in [0, 0.5), X(s) = 0, s \in [0.5, 1]$.
$Y = 1 - X$

$$\begin{aligned}\mathbf{P}(X = 1) = 0.5, \mathbf{P}(Y = 1) = 0.5 \\ \mathbf{P}(X = 0) = 0.5, \mathbf{P}(Y = 0) = 0.5\end{aligned} \tag{24}$$

Illustration with Dice

X-outcome of first toss, Y-outcome of second toss.

Illustration with Dice

X-outcome of first toss, Y-outcome of second toss.

$X \backslash Y$					
(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Illustration with Dice

Z-sum of the two outcomes.

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

- Which R.Vs are independent?

- Which R.Vs are independent?
- Which R.Vs are identically distributed?

- Which R.Vs are independent?
- Which R.Vs are identically distributed?
- What is the marginal distribution of X, Y and Z?

- Which R.Vs are independent?
- Which R.Vs are identically distributed?
- What is the marginal distribution of X, Y and Z?
- Is $f_{XZ}(x, z) = f_X(x) f_Z(z)$?

- Which R.Vs are independent?
- Which R.Vs are identically distributed?
- What is the marginal distribution of X, Y and Z?
- Is $f_{XZ}(x,z) = f_X(x)f_Z(z)$?
- Is $f_{XY}(x,y) = f_X(x)f_Y(y)$?

- Which R.Vs are independent?
- Which R.Vs are identically distributed?
- What is the marginal distribution of X, Y and Z?
- Is $f_{XZ}(x, z) = f_X(x)f_Z(z)$?
- Is $f_{XY}(x, y) = f_X(x)f_Y(y)$?
- What is the conditional distribution $f_{Z|X=3}, f_{Z|X>3}, f_{Z|X>3,Y<5}, f_{Z|X=5,Y=5}$?

- Which R.Vs are independent?
- Which R.Vs are identically distributed?
- What is the marginal distribution of X, Y and Z?
- Is $f_{XZ}(x,z) = f_X(x)f_Z(z)$?
- Is $f_{XY}(x,y) = f_X(x)f_Y(y)$?
- What is the conditional distribution $f_{Z|X=3}, f_{Z|X>3}, f_{Z|X>3,Y<5}, f_{Z|X=5,Y=5}$?

Conditional Mean and Covariance

The Conditional Mean denoted as $\mathbf{E}(X|Y)$ is a function of Y, say $h(Y)$.

The Conditional Mean denoted as $\mathbf{E}(X|Y)$ is a function of Y, say $h(Y)$.

$$\begin{aligned} h(y) &= \mathbf{E}(X|y = y) \\ &= \sum x f_{X|Y=y}(x) \quad \text{for D.R.V} \\ &= \int x f_{X|Y=y}(x) dx \quad f\text{or C.R.V} \end{aligned}$$

The Conditional Mean denoted as $\mathbf{E}(X|Y)$ is a function of Y, say $h(Y)$.

$$\begin{aligned} h(y) &= \mathbf{E}(X|y = y) \\ &= \sum x f_{X|Y=y}(x) \quad \text{for D.R.V} \\ &= \int x f_{X|Y=y}(x) dx \quad f\text{or C.R.V} \end{aligned}$$

One can also check that $\mathbf{E}(h(Y)) = \mathbf{E}(X)$.

Conditional Mean and Covariance

The Conditional Mean denoted as $\mathbf{E}(X|Y)$ is a function of Y, say $h(Y)$.

$$
\begin{aligned}
h(y) &= \mathbf{E}(X|y = y) \\
&= \sum x f_{X|Y=y}(x) \quad \text{for D.R.V} \\
&= \int x f_{X|Y=y}(x) dx \quad f\text{or C.R.V}
\end{aligned}
$$

One can also check that $\mathbf{E}(h(Y)) = \mathbf{E}(X)$.
The covariance between R.Vs X and Y is defined as

The Conditional Mean denoted as $\mathbf{E}(X|Y)$ is a function of Y, say $h(Y)$.

$$h(y) = \mathbf{E}(X|y = y) = \sum x f_{X|Y=y}(x) \quad \text{for D.R.V} = \int x f_{X|Y=y}(x) dx \quad for \text{ C.R.V}$$

One can also check that $\mathbf{E}(h(Y)) = \mathbf{E}(X)$.
The covariance between R.Vs X and Y is defined as

$$\mathbf{E}\left[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))\right] \quad (25)$$

- What is best constant that approximates X, i.e., $\min_a \mathbf{E}\left[(X - a)^2\right]$

Approximation of a Random Variable X

- What is best constant that approximates X, i.e., $\min_a \mathbf{E}\left[(X - a)^2\right]$
  $a = \mathbf{E}(x)$.
- What is the best possible function of Y that approximates X, i.e., $\min_g \mathbf{E}\left[(X - g)^2\right]$

Approximation of a Random Variable X

- What is best constant that approximates X, i.e., $\min_a \mathbf{E}\left[(X - a)^2\right]$
  $a = \mathbf{E}(x)$.
- What is the best possible function of Y that approximates X, i.e., $\min_g \mathbf{E}\left[(X - g)^2\right]$
  $g = \mathbf{E}(X|Y)$.
- What is the best possible linear function of Y that approximates X, i.e., $\min_{a,b} \mathbf{E}\big(X - (bY + a)\big)^2$

- What is best constant that approximates X, i.e., $\min_a \mathbf{E}\left[(X - a)^2\right]$
  $a = \mathbf{E}(x)$.
- What is the best possible function of Y that approximates X, i.e., $\min_g \mathbf{E}\left[(X - g)^2\right]$
  $g = \mathbf{E}(X|Y)$.
- What is the best possible linear function of Y that approximates X, i.e., $\min_{a,b} \mathbf{E}\left(X - (bY + a)\right)^2$
  $b = \frac{CoVar(X,Y)}{Var(Y)}$
  $a = \mathbf{E}(X - bY)$

Sum of Two Independent Random Variables

Let X and Y be independent R.Vs, Let $Z = X + Y$.

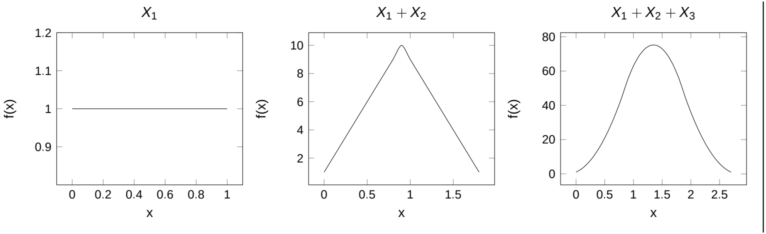
Let X and Y be independent R.Vs, Let $Z = X + Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx$$

Sum of Two Independent Random Variables

Let X and Y be independent R.Vs, Let $Z = X + Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx$$



Sum of infinitely many random Variables

If $X_1, X_2, \ldots, X_n \ldots$ are independent identically distributed random variables

- $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i \to \mathbf{E}(X_1)$.

Sum of infinitely many random Variables

If $X_1, X_2, \ldots, X_n \ldots$ are independent identically distributed random variables

- $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i \to \mathbf{E}(X_1)$.
- If $Y = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$. Then $f_Y is \mathcal{N}(0,1)$, i.e., Gaussian R.V with mean 0 and variance 1.