ME 233 Advance Control II

Lecture 3
Introduction to Probability Theory

(ME233 Class Notes pp. PR1-PR3)

Outline

- Sample Space and Events
- Probability function
- Random Variable
- PDF, expectation and variance

Sample Space and Events

Assume:

- We do an experiment many times.
  - Each time we do an experiment we call that a ***trial***
- The outcome of the experiment may be different at each trial.

ω_i : The i-th possible outcome of the experiment

Sample Space and Events

Sample Space Ω **:**

The space which contains all possible outcomes of an experiment.

$$\Omega = \{\omega_1, \omega_2, \cdots, w_n\}$$

ω_i **:** The i-th possible outcome of the experiment

Each outcome is an element of Ω

Example: Dice

Experiment:

A situation whose ***outcome*** depends on chance

- throwing of a fair dice once

Sample Space Ω

The set of <u>all possible</u> ***outcomes*** of an experiment

Ω = 

Events

Event S_j :

Is a subset of the union of the sample space Ω and the empty set ϕ

Sample space with n outcomes:

$$\Omega = \{\omega_1, \omega_2, \cdots, w_n\}$$

There are 2^n possible events:

$$\mathcal{S} = \{S_1, \cdots, S_{2^n}\}$$

Probability - events

Experiment: throwing of a dice once

Ω = 

Outcomes: *elements of the sample space* S

Events: *Are <u>subsets</u> of the sample space* S

An event occurs if any of the outcomes in that event occurs.

Empty subsets are ***null*** *or* ***impossible events***

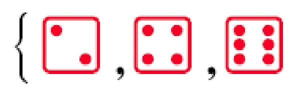
Probability - events

Experiment: throwing of a dice once

Ω = 

<u>Some events:</u>

- The event E of observing an even number of dots:

$E =$ 

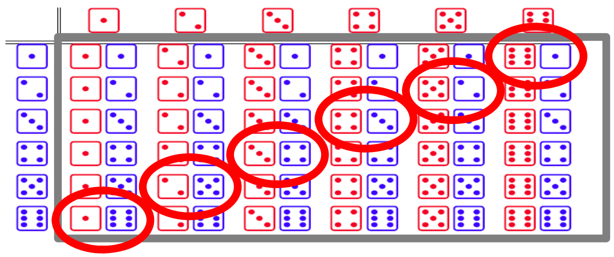
- The event O of observing an odd number of dots:

$O =$ 

Example: throwing a pair of dice
(one red and one blue)

– the sample space has ***36*** outcomes:

$\Omega =$



- *The event L of obtaining the number **7** is*

$$L = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

L occurs if any of the outcomes in L occurs.

Probability function

We now consider the probability that a certain event occurs.

An event occurs if any of the outcomes in that event occurs.

The probability of event A will be define by

$$P(A)$$

Probability

A number between 0 and 1, inclusive, that indicates **<u>how likely an event is to occur</u>**.

- An event with probability of 0 is a **<u>null event.</u>**
- An event with probability of 1 is a **<u>certain event.</u>**

- Probability of event A is denoted as $P(A)$.
- The closer $P(A)$ to 1, the more likely is A to happen.

Probability

A number between 0 and 1, inclusive, that indicates **<u>how likely an event is to occur</u>**.

- An event with probability of 0 is a **<u>null event.</u>**
  - a man gets pregnant
  - a woman dies of prostate cancer
- An event with probability of 1 is a **<u>certain event.</u>**
  - the sun will set tonight
  - a person eventually dies

Intuitive Notion of Probability Frequentist approach

The probability of event A is

$$P(A) = \frac{\text{Possible outcomes associated with } A}{\text{Total possible outcomes}}$$

$$0 \leq P(A) \leq 1$$

Assigning Probability - Frequentist approach

- An experiment is repeated n times under essentially <u>identical</u> conditions
- if the event A occurs m times, then as n grows <u>large</u>

$$P(A) \approx \frac{m}{n}$$

Dice example

Experiment: throwing of a **<u>fair</u>** dice once

$\Omega = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$ $\quad \Omega = \{1,2,3,4,5,6\}$

- ***P(Ω) = 1***
- ***P(1) = 1/6, P(3) =1/6, P(6) = 1/6***
- ***P(even number) = 3/6 = 1/2***
- ***P(odd number) = 3/6 = 1/2***

Example: poker

Example: In poker you are dealt 5 cards from a deck of 52

- What is the probability of being dealt four of a kind?
  - e.g. 4 aces or four kings, and so fourth?

$$P(\text{four of a kind}) = ?$$

Example: poker

Solution:

1. There are only 48 possible hands containing 4 aces, another 48 containing 4 kings, etc.

2. Thus, there are **13 x 48** possible "four of a kind" hands.

3. The possible number of hands is obtained from the combination formula for "52 things taken 5 at a time":

total possible outcomes: $\binom{52}{5} = \frac{52!}{5!(52-5)!} = 2,598,960$

4. Thus, $P(\text{four of a kind}) = \frac{13 \times 48}{2,598,960} = 0.00024$

Union, Complement and Intersection

For a sample space $\Omega = \{\omega_1, \omega_2, \cdots, w_n\}$
And the set of all events $\mathcal{S} = \{S_1, \cdots, S_{2^n}\}$

- Union of two events (or):

$$S_i \cup S_j = \{\omega_m \mid \omega_m \in S_i \text{ or } \omega_m \in S_j\}$$

- Intersection of two events (and):

$$S_i \cap S_j = \{\omega_m \mid \omega_m \in S_i \text{ and } \omega_m \in S_j\}$$

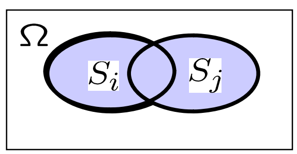
- Complement of an event (not):

$$\backslash S_i = \{\omega_m \mid \omega_m \in \Omega \cup \phi \text{ and } \omega_m \notin S_i\}$$

Union, Complement and Intersection

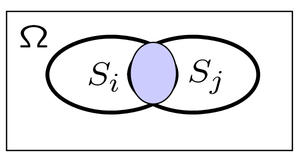
- Union of two events:

$$S_i \cup S_j$$



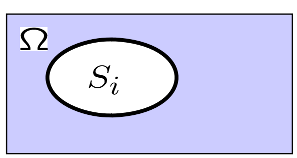
- Intersection of two events:

$$S_i \cap S_j$$



- Complement of an event:

$$\backslash S = S^c$$



Probability Space

The probability space to be the triple:

$$(\Omega, \mathcal{S}, P)$$

Where

- Ω is the sample space
- $\mathcal{S}$ the set of all possible events
- $P : \mathcal{S} \to [0, 1]$ is the probability function

Probability function

Probability function: $P : \mathcal{S} \rightarrow [0,1]$

Satisfies 3 axioms:

1. $P(S_i) \geq 0$

2. $P(\Omega) = 1$

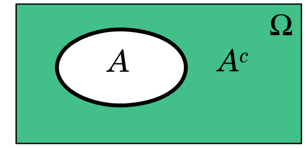
3. $P(S_i \cup S_j) = P(S_i) + Pr(S_j)$ if $S_i \cap S_j = \phi$

Complement

- The **complement** of an event $\boldsymbol{A}$, denoted by $\boldsymbol{A^c}$, is the set of outcomes that are not in $\boldsymbol{A}$

- $\boldsymbol{A^c}$ occurring *means* that $\boldsymbol{A}$ *does not occur*

$$A^c = \{\omega \mid \omega \in \Omega \cup \phi \text{ and } \omega \notin A\}$$

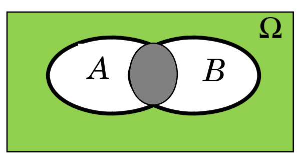


$$\boxed{P(A^c) = 1 - P(A)}$$

Intersection of two events

- The **intersection** of two events $\boldsymbol{A}$ and $\boldsymbol{B}$, denoted by $\boldsymbol{A} \bigcap \boldsymbol{B}$, is the set of outcomes that are in $\boldsymbol{A}$, **and** $\boldsymbol{B}$.

- *If the event $\boldsymbol{A} \cap \boldsymbol{B}$ occurs, then **both** $\boldsymbol{A}$ and $\boldsymbol{B}$ occur*



- *Events $\boldsymbol{A}$ and $\boldsymbol{B}$ are mutually exclusive, i.e. they cannot happen at the same time if*

$$A \bigcap B = \varnothing$$

Example of Intersection of two events

Experiment: throwing of a dice once

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}.$$

- Events E and O are mutually exclusive

$$E = \{⚁, ⚃, ⚅\} \qquad O = \{⚀, ⚂, ⚄\}.$$

$$E \bigcap O = \varnothing$$

Independent Events

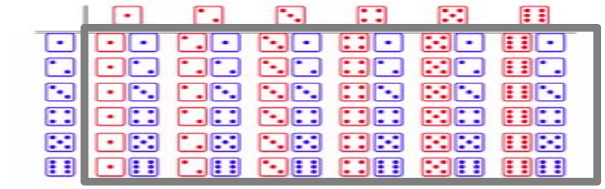
- $A \cap B$ means both A and B occur
- Two events are **independent** if the events do not influence each other.
  - That is, if event A occurs, it does not affect chances of B occurring, and vice versa.
- If two events are independent, then

$$\boxed{P(A \cap B) = P(A) \times P(B)}$$

Example of independence

Experiment: throwing a pair of dice (one red and one blue)

$\Omega =$



36 possible outcomes

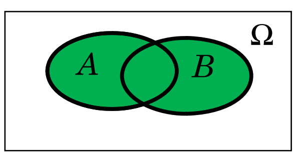
The probability of throwing a red 1 and a blue 5 is

$$P(1 \cap 5) = P(1) \times P(5)$$

$$= 1/6 \times 1/6 = 1/36$$

Union of two events

- The **union** of two events A and B, denoted by $A \cup B$, is the set of outcomes that are in A, or B, or both
- *If* the event $A \cup B$ *occurs, then either* A *or* B *or both occur*



Law of Union

- $A \cup B$ means both A or B or both occur
- If A and B are mutually exclusive, i.e. $A \cap B = \varnothing$

$$\boxed{P(A \cup B) = P(A) + P(B)}$$

- If A and B are not mutually exclusive

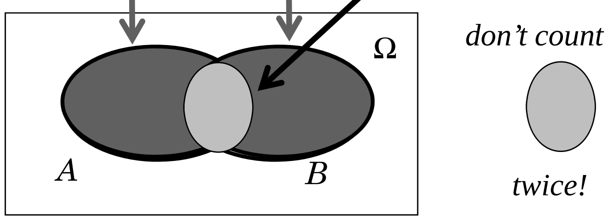
$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

Law of Union

- $A \cup B$ means both A or B or both occur

- If A and B are <u>not</u> mutually exclusive

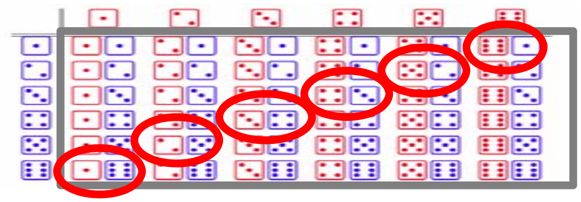
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Example

Experiment: throwing a pair of dice (one red and one blue)

$\Omega =$



- $P(L)$ = *the probability of obtaining a **7***

$$L = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

$$P(L) = P(1,6) + P(2,5) + P(3,4) + P(4,3) + P(5,2) + P(6,1)$$

$$P(L) = 6/36 = 1/6$$

Join Probability

Let A and B be two events

$$P(A \cap B)$$

is often called the ***join probability*** of A and B

$$P(A) \qquad P(B)$$

are often called the ***marginal probabilities*** of A and B

Conditional Probability

Let A and B be two events and $P(B) \neq 0$

The conditional probability of event A given that event B has occurred is

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

Bayes' Rule

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two events

$$\boxed{\begin{aligned} P(A|B)P(B) &= P(B|A)P(A) \\ &= P(A \cap B) \end{aligned}}$$

Independence

$\boldsymbol{A}$ and $\boldsymbol{B}$ *are **independent*** if

$$P(A \cap B) = P(A)\, P(B)$$

Or equivalently

$$\begin{aligned} P(A|B) &= P(A) \\ P(B|A) &= P(B) \end{aligned}$$

Array of Probabilities

Let C and D be two chance experiments.

Set of disjoint events associate with C

$$\mathcal{C} = \{C_1,\, C_2,\, \cdots C_m\}$$

Set of disjoint events associate with D

$$\mathcal{D} = \{D_1,\, D_2,\, \cdots D_n\}$$

Array of Probabilities

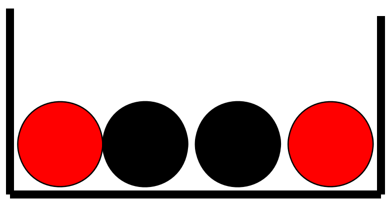
We can construct:

C \ D	Event D_1	Event D_2	$\cdots$	Event D_n	Marginal Probabilities
Event C_1	$P(C_1 \cap D_1)$	$P(C_1 \cap D_2)$	$\cdots$	$P(C_1 \cap D_n)$	$P(C_1) = \sum_{i=1}^{n} P(C_1 \cap D_i)$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
Event C_m	$P(C_m \cap D_1)$	$P(C_m \cap D_1)$	$\cdots$	$P(C_m \cap D_n)$	$P(C_m) = \sum_{i=1}^{n} P(C_m \cap D_i)$
Marginal Probabilities	$P(D_1) = \sum_{i=1}^{m} P(C_i \cap D_1)$	$P(D_2) = \sum_{i=1}^{m} P(C_i \cap D_2)$	$\cdots$	$P(D_n) = \sum_{i=1}^{m} P(C_i \cap D_n)$	$Sum = 1$

Example:

There are 4 balls in one jar, 2 balls are red and two balls are black.

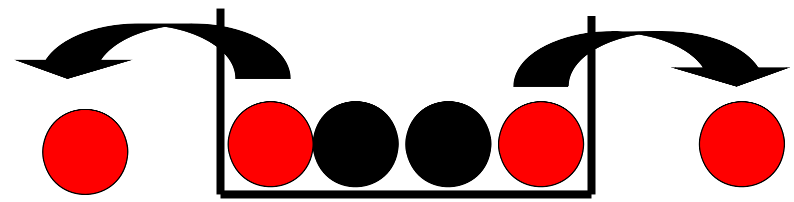
- A person can pull a ball from the jar two times, without seeing the balls inside the jar.



Example:

What is the probability of picking a red ball after having picked a red ball the first time?

To answer this question, lets built the table of probabilities.

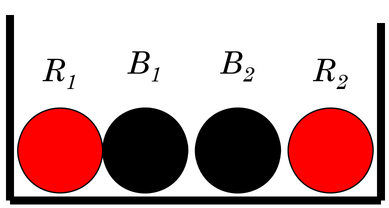


Example:

What is the probability of picking a red ball after having picked a red ball the first time?

To answer this question, lets built the table of probabilities.

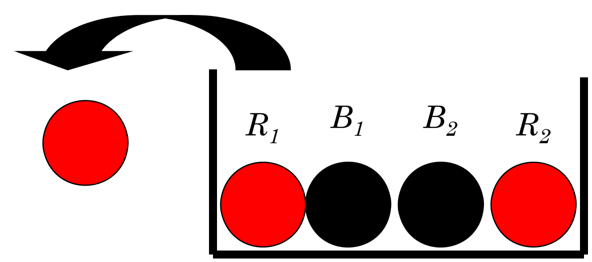
Labels:



Example:

Probability of picking R_1 the first time?

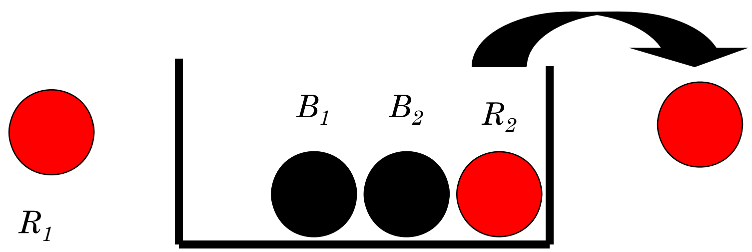
$$P(R_1) = 1/4$$



Example:

Probability of picking R_2 with only 3 balls left?

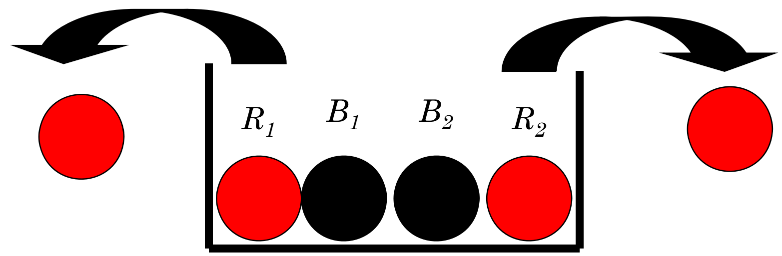
$$P(R_2) = 1/3 \qquad \text{(second time)}$$



Example:

Probability of picking R_1 the first time and R_2 the second time?

$$P(R_1 \cap R_2) = 1/4 \times 1/3 = 1/12$$



Example: Array of Probabilities

2 pick / 1 pick	R_1	R_2	B_1	B_2	Marginal Probabilities
R_1	o $P(R_1 \cap R_1)$	1/12 $P(R_1 \cap R_2)$	1/12 $P(R_1 \cap B_1)$	1/12 $P(R_1 \cap B_2)$	$P(R_1)$ 1/4
R_2	1/12	o	1/12	1/12	1/4
B_1	1/12	1/12	o	1/12	1/4
B_2	1/12	1/12	1/12	o	1/4
Marginal Probabilities	1/4	1/4	1/4	1/4	$Sum = 1$

Probability of pick red balls consecutively

Probability of picking <u>a red ball</u> the first time and <u>a red ball</u> the second time?

$$\left.\begin{array}{l}\text{• Picking } R_1 \text{ first and } R_2 \text{ second} \\ \text{Picking } R_2 \text{ first and } R_1 \text{ second}\end{array}\right\} \quad \textit{are } \underline{\textbf{\textit{mutually exclusive}}} \textbf{\textit{ events}}$$

$$P(\text{red} \cap \text{red}) = P\left[(R_2 \cap R_1) \cup (R_1 \cap R_2)\right]$$

$$P(\text{red} \cap \text{red}) = P(R_2 \cap R_1) + P(R_1 \cap R_2)$$

$$= \frac{1}{12} + \frac{1}{12} \quad = \frac{1}{6}$$

Example: Array of Probabilities

2 pick / 1 pick	*Red*	*Black*	Marginal Probabilities
Red	$1/6$ $P(red \cap red)$	$1/3$ $P(red \cap black)$	1/2
Black	$1/3$ $P(black \cap red)$	$1/6$ $P(black \cap black)$	1/2
Marginal Probabilities	1/2	1/2	$Sum = 1$

Example:

What is the probability of picking a red ball the second time after having picked a red ball the first time?

$$P(Red|Red) = \frac{P(Red \cap Red)}{P(Red)}$$

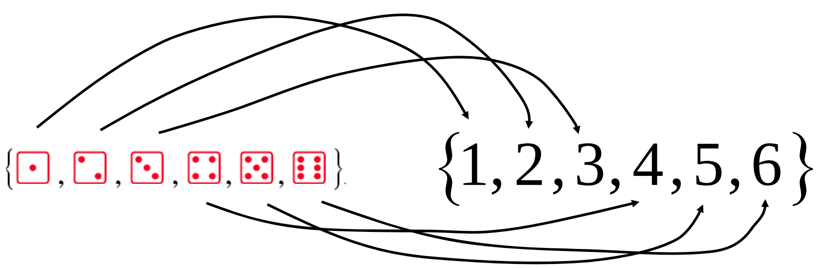
$$\boxed{P(Red|Red) = \frac{1/6}{1/2} = \frac{1}{3}}$$

Discrete random variable

Given a sample Space Ω, a random variable X is a function that assigns to each outcome a unique numerical value.

- Example: throwing of a dice once 

$\Omega =$  $= \{1,2,3,4,5,6\}$



Discrete random variable

- Example: throwing of a dice once 

$\Omega =$  $= \{1,2,3,4,5,6\}$

- In this case, the random variable X only takes discrete values

$$x_i \in \{1,2,3,4,5,6\}$$

- The random variable X is defined by the **probability mass function**

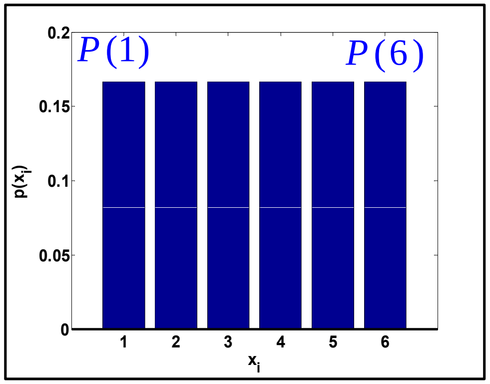
$$P(x_i) = P(X = x_i)$$

the probability that, after throwing a dice, X will be equal to x_i

Discrete random variable

- For a **fair dice**, the probability mass function of the random variable X is

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$$



the probability mass function satisfies:

$$\sum_{i=1}^{6} P(x_i) = 1$$

Expected value

- For a discrete random variable X taking on the N possible values

$$x_1, x_2, x_3, \ldots, x_k, \ldots, x_N$$

the **expected value** or **mean** of X is defined by

$$\boxed{E[X] = m_x = \hat{x} = \sum_{k=1}^{N} x_k P(x_k)}$$

$$E[x] = x_1 P(x_1) + x_2 P(x_2) + \ldots + x_N P(x_N)$$

Expected value

Example: For a **fair dice,**

$\Omega =$  $= \{1, 2, 3, 4, 5, 6\}$

- X takes 6 possible values $x_i = 1, 2, 3, 4, 5, 6$
- $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$

the **expected value** or **mean** of X

$$E(X) = m_x = \sum_{k=1}^{6} x_k P(x_k) = \frac{1}{6}\sum_{k=1}^{6} k = \frac{1}{6} 21 = 3.5$$

Variance and standard deviation

- For a discrete random variable X taking on the N possible values

$x_1, x_2, x_3, \ldots, x_k, \ldots, x_N$ and a mean $m_X = \hat{x}$

the **variance** of X is defined by

$$\boxed{E[(X - m_X)^2] = \sigma_X^{\ 2} = \sum_{k=1}^{N} (x_k - m_X)^2 P(x_k)}$$

where $\boldsymbol{\sigma_X}$ is the standard deviation of X

Variance and standard deviation

Example: For a **fair dice**, where $x_i = 1,2,3,4,5,6$

has mean $m_x = 3.5$ and $P(x_i) = 1/6$

the variance and standard deviation of X are

$$E[(x-m_x)^2] = \sum_{k=1}^{6}(x_k-3.5)^2 P(x_k) = \frac{1}{6}\sum_{k=1}^{6}(k-3.5)^2$$

$$= \frac{1}{6}\left[(1-3.5)^2+(2-3.5)^2+\ldots+(6-3.5)^2\right] = 2.9167$$

$$\sigma_x = \sqrt{E[(X-m_x)^2]} = \sqrt{2.9167} = 1.7078$$

Cumulative Distribution Function

- The cumulative distribution function (CDF) for a random variable X is

$$F_X(x) = P(X \le x)$$

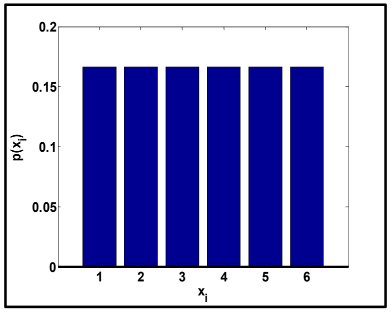
Find index $\boldsymbol{k}$ *such that* $x_k \le x < x_{k+1}$

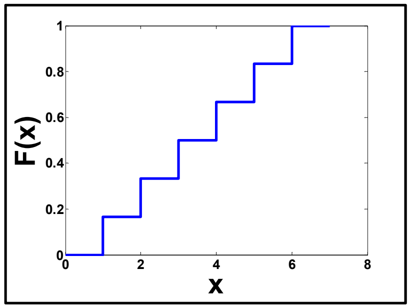
$$F_X(x) = \sum_{j=1}^{k} P(x_j)$$

Cumulative Distribution Function

- The cumulative distribution function (CDF) for a random variable X is

$$F_X(x) = \sum_{j=1}^{k} P(x_j) \qquad x_k \le x < x_{k+1}$$



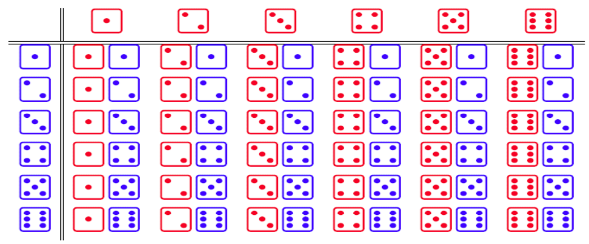


Sum of two uniform random variables

- Let X and Y be 2 random variables with constant probability mass function
- Let $Z = X + Y$
- The probability mass function of Z will not be constant

Throwing two fair dice

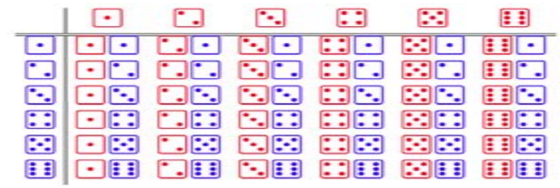
Experiment: throwing a pair of **fair** dice (red and blue)

$\Omega =$ 

– the sample space has ***36*** outcomes:

– each outcome has a ***1/36*** probability of occurring

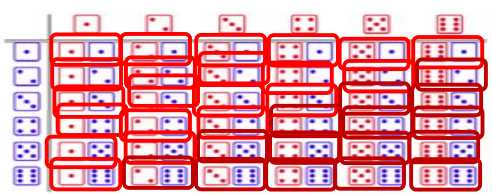
Throwing two fair dice

$\Omega =$ 

- Define the random variable Z associated with the **event** of observing the total number of dots on both dice after each throw

$Z = k$ when the throw results in the number k

Throwing two fair dice

$\Omega =$ 

number of outcomes ***36***

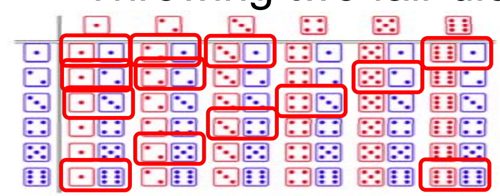
when the throw results in the number

only takes discrete values

$$z_i \in \{\boxed{2},\boxed{3},\boxed{4},\boxed{5},\boxed{6},\boxed{7},\boxed{8},\boxed{9},\boxed{10},\boxed{11},\boxed{12}\}$$

- *However, the union of events of observing the total number of dots on both dice after each throw contains the sample space*

Throwing two fair dice

$\Omega =$ 

probability of each outcome ***1/36***

when the throw results in the number

we now estimate:

$Z = 2 \rightarrow \; P(2) = 1/36$

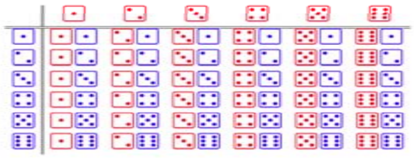
$Z = 3 \rightarrow \; P(3) = 2/36$

$Z = 4 \rightarrow \; P(4) = 3/36$

$Z = 7 \rightarrow \; P(7) = 6/36$

$Z = 12 \rightarrow \; P(12) = 1/36$

Throwing two fair dice

$\Omega =$ 

takes discrete values

$$z_i \in \{2,3,4,5,6,7,8,9,10,11,12\}$$

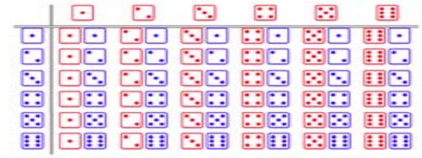
is defined by the **probability mass function**

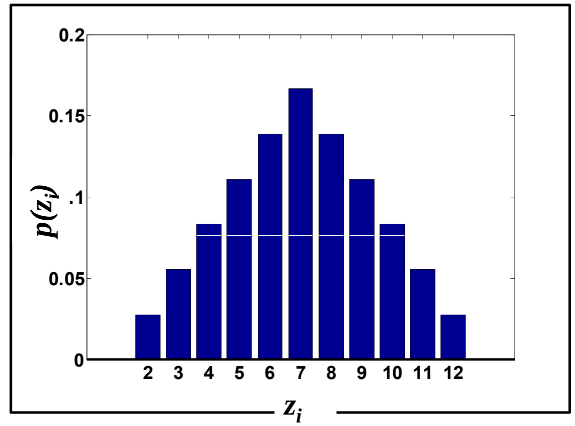
$P(2) = 1/36$ $\quad P(5) = 4/36$ $\quad P(8) = 5/36$ $\quad P(11) = 2/36$

$P(3) = 2/36$ $\quad P(6) = 5/36$ $\quad P(9) = 4/36$ $\quad P(12) = 1/36$

$P(4) = 3/36$ $\quad P(7) = 6/36$ $\quad P(10) = 3/36$

Throwing two fair dice

$\Omega =$ 



the probability mass function satisfies:

$$\sum_{k=2}^{12} P(k) = 1$$

Continuous random variable

A continuous-valued random X variable takes on a range of **real** values

- For the probability space, $(\Omega, \mathcal{S}, P)$
- A random variable X is a mapping $X : \Omega \to \mathcal{R}$

Example:

- An experiment whose outcome is a real number, e.g. measurement of a noisy voltage.

$$X \in [V_{\min}, V_{\max}]$$

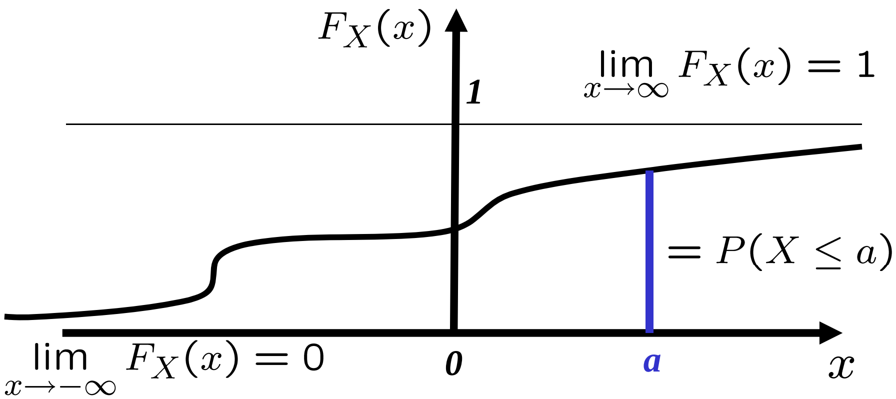
Probability Distribution Function

Probability distribution function associate with the random variable $\boldsymbol{X}$:

$$\boxed{F_X(x) = P(X \leq x)}$$

The probability that the random variable $\boldsymbol{X}$ will be less than or equal to the value $\boldsymbol{x}$

Properties of the probability distribution



Properties of the probability distribution

$$F_X(x) = P(X \le x)$$

1. $\lim_{x\to-\infty} F_X(x) = 0$

2. $\lim_{x\to\infty} F_X(x) = 1$

3. $F_X(x)$ is a monotone non decreasing

4. $F_X(x)$ is left-continuous

Probability Density Function

For a ***differentiable*** probability distribution function,

$$F_X(x) = P(X \le x)$$

Define the **probability density function (PDF)**,

$$\boxed{p_X(x) = \frac{dF_X(x)}{dx}}$$

Probability Density Function

$$p_X(x) = \frac{dF_X(x)}{dx}$$

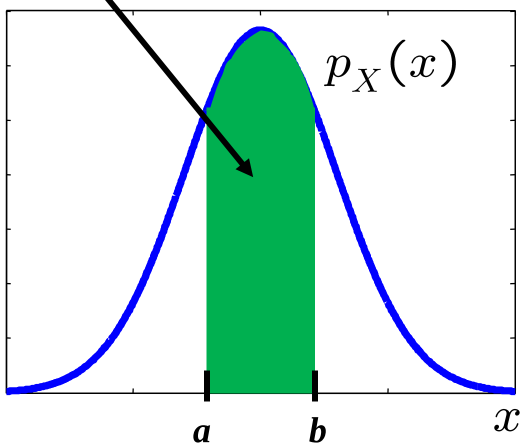
Interpretation:

$$p_X(x)\,\Delta x \approx P(x \le X \le x + \Delta x)$$

$$\int_a^b p_X(x)dx = P(a \le X \le b)$$

Probability Density Function

$$\int_a^b p_X(x)dx = P(a \le X \le b)$$



Expectation

The ***expected value*** of random variable X is:

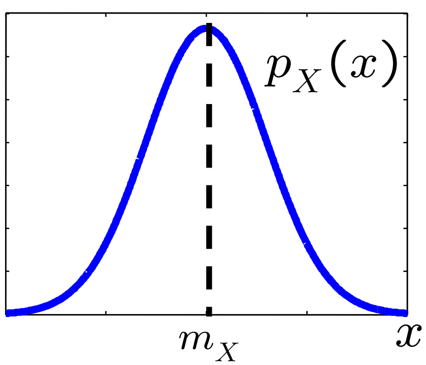
$$\boxed{E[X] = \int_{-\infty}^{\infty} x\, p_X(x)\, dx}$$

This is the average value of X.

It is also called the ***mean*** or first moment of X

Expected value - notation

$$\boxed{m_X = \hat{x} = E[X]}$$



Expected value of a function

f : real valued function of random variable X

$$Y = f(X)$$

The expected value of Y is

$$\boxed{E[Y] = \int_{-\infty}^{\infty} f(x)\, p_X(x)\, dx}$$

Variance

The ***variance*** of random variable X is:

$$\boxed{\begin{aligned}\sigma_X^2 &= E[(X-m_X)^2] \\ &= \int_{-\infty}^{\infty} (x-m_X)^2 \, p_X(x)\, dx\end{aligned}}$$

where $m_X = E[X]$

σ_X Is called the standard deviation of X

Variance

$$\begin{aligned}\sigma_X^2 &= E[(X-m_X)^2] \\ &= E[X^2] - m_X^2\end{aligned}$$

where

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \, p_X(x)\, dx$$

Proof

$$\begin{aligned}\sigma_X^2 &= \int_{-\infty}^{\infty} (x-m_X)^2 \, p_X(x)\, dx \\ &= \int_{-\infty}^{\infty} \left(x^2 - 2xm_X + m_X^2\right) p_X(x)\, dx\end{aligned}$$

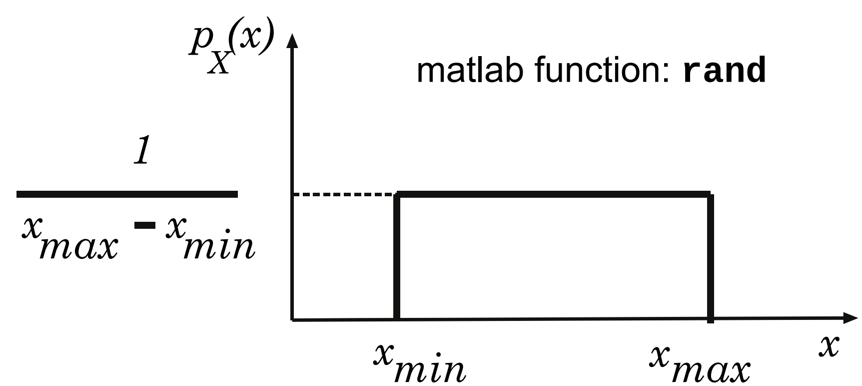
$$\left(\int_{-\infty}^{\infty} p_X(x)\, dx = 1\right)$$

$$= E[X^2] - 2\, m_X \underbrace{\int_{-\infty}^{\infty} x p_X(x)\, dx}_{m_X} + m_X^2$$

$$= E[X^2] - 2\, m_X^2 + m_X^2 \quad = \quad E[X^2] - m_X^2$$

Uniform Distribution

A random variable X which is uniformly distributed between x_{min} and x_{max} has the PDF:



Summing uniformly distributed random variables

- Let X and Y be 2 uniformly distributed variables between [0,1]
- The random variable

$$Z = X + Y$$

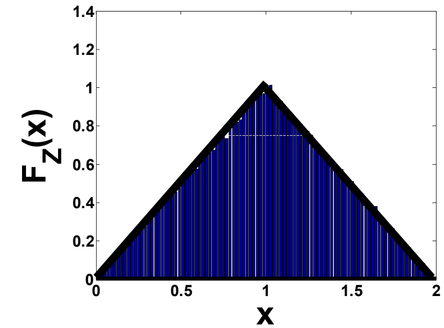
- is **not uniformly distributed**

Summing uniformly distributed random variables

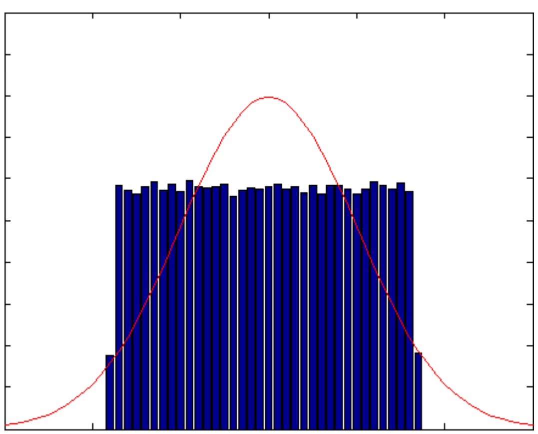
- Let X and Y be 2 uniformly distributed variables between [0,1]

$$Z = X + Y$$

```
X=rand(1,1e5);
Y=rand(1,1e5);
Z=X+Y;
[fz,x]=hist(Z,100);
w_fz=x(end)/length(fz);
fz=fz/sum(fz)/w_fz;
bar(x,fz)
xlabel('x')
ylabel('F_Z(x)')
```



Summing a very large number of random variables

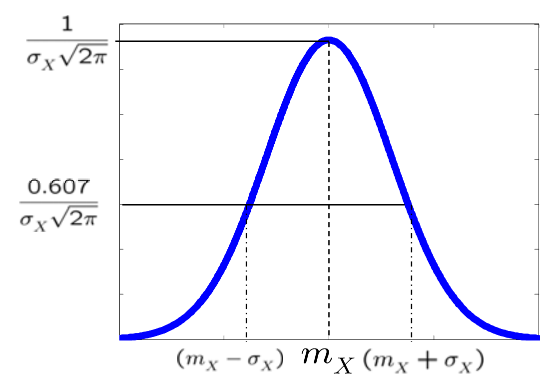


Gaussian (Normal) Distribution

$$\boxed{p_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}}$$

Normal distribution

$$X \sim N(m_X, \sigma_X^2)$$



History of the Normal Distribution

From Wikipedia, the free encyclopedia

- The normal distribution was first introduced by de **Moivre** in an article in **1733** in the context of approximating certain binomial distributions for large n.
- His result was extended by **Laplace** in his book Analytical Theory of Probabilities (1812), and is now called the theorem of de Moivre-Laplace.
- **Laplace** used the normal distribution in the analysis of errors of experiments.

History of the Normal Distribution

From Wikipedia, the free encyclopedia

- The important method of **least squares** was introduced by **Legendre** in 1805.
- **Gauss**, who claimed to have used the method since 1794, justified it rigorously in 1809 by assuming a normal distribution of the errors.
- That the distribution is called the normal or Gaussian distribution is an instance of Stigler's law of eponymy: "No scientific discovery is named after its original discoverer."

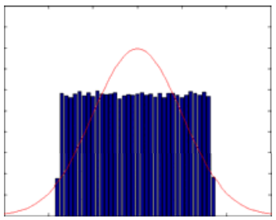
Central limit theorem

Let $X_1, X_2, \ldots$ be independent random variables each with mean m_x and variance σ_x^2. Then the sequence

$$Z_n = \frac{\sum_{k=1}^{n}(X_k - m_X)}{\sqrt{n}\sigma_X^2}$$

converges in distribution to a normal random variable with distribution

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



Properties of normal distributions

- **Consequence of the central limit theorem:**
  - In most engineering applications, noise is frequently due to the superposition of many small contributions.
  - Using a Gaussian distribution to model noise is often a good assumption.

Properties of normal distributions

The sum of two Gaussian random variables is also a Gaussian.

Assume X and Y are ***independent and Gaussian***

$$X \sim N(m_X, \sigma_X^2) \qquad Y \sim N(m_y, \sigma_Y^2)$$

$$Z = X + Y$$

$$Z \sim N(m_X + m_Y, \sigma_X^2 + \sigma_Y^2)$$

Bilateral Laplace and Fourier Transforms

Given $f : \mathcal{R} \to \mathcal{R}$

- Laplace transform: $F(s) = \int_{-\infty}^{\infty} e^{-st} f(t) dt$
  $s \in \mathcal{C}$
- Fourier transform: $F(j\omega) = \int_{-\infty}^{\infty} e^{-j\omega t} f(t) dt$
  $\omega \in \mathcal{R}$
- Inverse F. T. $f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} F(\omega) d\omega$

Properties of Normal distributions

The Fourier transform of a zero-mean Gaussian distribution is also Gaussian.

$$p_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_X^2}}$$

$$P_X(j\omega) = \mathcal{F}\{p_X(\cdot)\} = \int_{-\infty}^{\infty} e^{-j\omega x} p_X(x)\, dx$$

$$= e^{\frac{-\sigma_X^2 \omega^2}{2}}$$

Laplace transform of normal PDF

$$p_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$$

$$P_X(s) = \int_{-\infty}^{\infty} e^{-sx} p_X(x)\, dx = \frac{1}{\sigma_X \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-sx} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx$$

$$= \frac{1}{\sigma_X \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-A(x)} dx$$

where, after “completing the squares”,

$$A(x) = sx + \frac{x^2}{2\sigma_X^2} + \frac{m_X^2}{2\sigma_X^2} - \frac{2m_X x}{2\sigma_X^2}$$

$$= \frac{1}{2\sigma_X^2} \left\{ \left[x + (s\sigma_X^2 - m_X)\right]^2 - s^2\sigma_X^4 + 2m_X s \sigma_X^2 \right\}$$

Laplace transform of normal PDF

substituting,

$$P_X(s) = e^{(s^2\sigma_X^2/2)-sm_X} \underbrace{\int_{-\infty}^{\infty} \left\{ \frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x+s\sigma_X^2-m_X)^2/2\sigma_X^2} \right\} dx}_{=1}$$

(area under a PDF = 1)

$$\boxed{P_X(s) = e^{(s^2\sigma_X^2/2)-sm_X}}$$

Fourier transform: $P_X(j\omega) = e^{\frac{-\omega^2\sigma_X^2}{2}} e^{-j\omega m_X}$