MA2002: Probability Theory

Shalu M A

IIITDM

Text Books

- Sheldon Ross. A First Course in Probability, 9th edition, Pearson.
- D. C. Montgomery and G. C. Runger, Applied Statistics and Probability for Engineers, Wiley India.
- R. A. Johnson, C. B. Gupta, Miller and Freund's Probability and Statistics for Engineers, Pearson Prentice-Hall.

To study/manage uncertainty

- Calculus $\longleftarrow$ limit (prediction)
- Probability and Statistics
- Fuzzy Logic/Mathematics

Real life problems

- Weather prediction $\longleftarrow$ economic /food security
- Investments $\longleftarrow$ Stock/share/commodity market
- Spread of a disease $\longleftarrow$ corona, ebola, TB

An example

- Bread $\longrightarrow$ Shelf life 4 days
- Mobile phone set $\longrightarrow$ Warranty period 6 months

Mr. X bought a bread on Monday (Manufactured on Monday) $\longrightarrow$ He got 4 days to consume it

Mr. Y bought a bread on Wednesday (Manufactured on Monday) $\longrightarrow$ He got 2 days to consume it

Mr. X bought a mobile phone set on January 1,2021 (Manufactured December 1, 2020) $\longrightarrow$ He got six months warranty

Mr. Y bought a mobile phone set on March 1,2021 (Manufactured December 1, 2020) $\longrightarrow$ He got six months warranty

Memoryless property !

Human beings are irrational (Richard Thaler)

- A patient meets the Dr. X to have heart surgery. Dr. X says the patient has 95 percentage survival rate.

Applications

Historically: In games of chance/gambling - for calculation of certain probabilities.

Human beings are irrational (Richard Thaler)

- A patient meets the Dr. X to have heart surgery. Dr. X says the patient has 95 percentage survival rate.
- A patient meets the Dr. Y to have heart surgery. Dr. Y says the patient has 5 percentage chance of death.

Who do you prefer ?

Applications

Historically: In games of chance/gambling - for calculation of certain probabilities.

Now:

- Engineering
- Science
- Economics
- Mathematics
- Computing
- Artificial Intelligence and Machine Learning

Applications

Historically: In games of chance/gambling - for calculation of certain probabilities.

Now:

- Engineering
- Science
- Economics
- Mathematics
- Computing
- Artificial Intelligence and Machine Learning

Probability versus Statistics

- **Probability Theory:** A branch of mathematics. Provides the framework, tools and techniques for computing *probabilities*.
- **Statistics:** The science of data. Deals with the collection, presentation, analysis, and use of data to make decisions, solve problems, and design products and processes.

In this Module

- Sample spaces, Events
- Operations on Events

Recall

Set

A set is a well-defined collection of objects.

Examples: $A = \{1, 2, 3\}$, $B = \{1, 2, 3, 4, 5\}$, $C = \{1, 2, 3, \dots\}$,
$D = \{p/q \mid p \text{ and } q \text{ are integers and } q \neq 0\}$.

Subset

A set A is a subset of another set B if every element (object) of A belongs to B. In this case, we write $A \subseteq B$.

Example: $\{1, 2, 3\} \subseteq \{1, 2, 3, 4, 5\} \subseteq \{1, 2, 3, \dots\}$.

Random experiment

An experiment with all possible outcomes are known in advance and the outcome of a trial is not known in advance.

Sample space

Definition

The set of all possible outcomes of a random experiment is called the *sample space* of the experiment.

Example: Experiment: Tossing a coin.

Outcome: H or T (head or tail)

Sample space: $S = \{H, T\}$

Example: Experiment: Rolling a die.

Sample space: $S = \{1, 2, 3, 4, 5, 6\}$.

More Examples

Example: Experiment: Tossing a coin three times.

Sample space: $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

Example: Experiment: Rolling a die twice.

Sample space: $S = \{(i, j) \mid 1 \leq i, j \leq 6\}$.

What is the size of the sample space?

More Example

Example: Experiment: Counting the number of phone calls received on a particular day.

Sample space: $S = \{0, 1, 2, 3, \ldots\}$

How many outcomes?

Example: Experiment: Waiting for a phone call.

Event

Definition

Any subset of a sample space is called an *event*.

Example: Tossing a coin: Sample space: $S = \{H, T\}$.

$A = \{H\}$ is a subset of the sample space. So, it is an event.

A is the event that a head is obtained.

Other events: $\{T\}, S = \{H, T\}$ and $\emptyset$.

Example: Rolling a die. Sample space: $S = \{1, 2, 3, 4, 5, 6\}$.

$A = \{1, 3, 5\}$ is an event. A is the event that an odd number is obained.

The experiment of rolling a die has 2^6 events.

Occurrence of an event

Definition

An event A is said to *occur* if the outcome of the experiment belongs to A.

Example: Tossing a coin. Sample space: $S = \{H, T\}$

Suppose H is the outcome.

Then the event $A = \{H\}$ has occurred.

The event $S = \{H, T\}$ has also occurred.

Example: Rolling a die. Sample space: $S = \{1, 2, 3, 4, 5, 6\}$.

Suppose 3 is the outcome.

Then the events $A = \{1, 3, 5\}$, $B = \{1, 2, 3\}$ and $S = \{1, 2, 3, 4, 5, 6\}$ all have occurred.

Example:

The sample space S is an event. It occurs always.

The sample space is called a *sure* event.

The empty set $\emptyset$ is an event. It never occurs.

The empty set is called an *improbable* event.

Algebra of Events

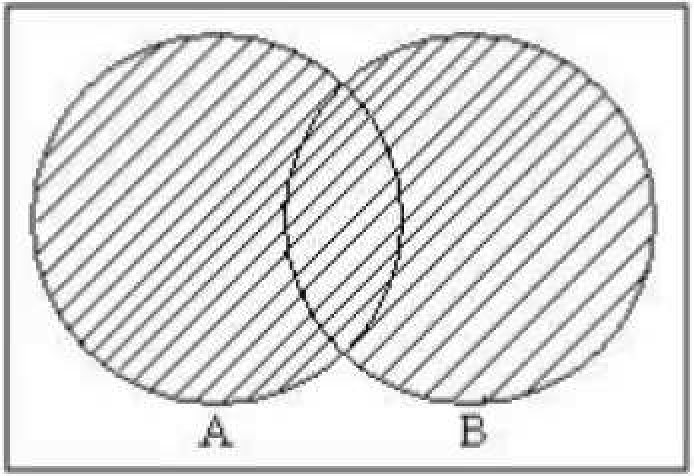
Events are sets!

So, it makes sense to take union, intersection, complement of events!

Union of Events

Let A and B be two events. Then their union $A \cup B$ is:

$$A \cup B = \{s \in S \mid s \in A \text{ or } s \in B\}.$$

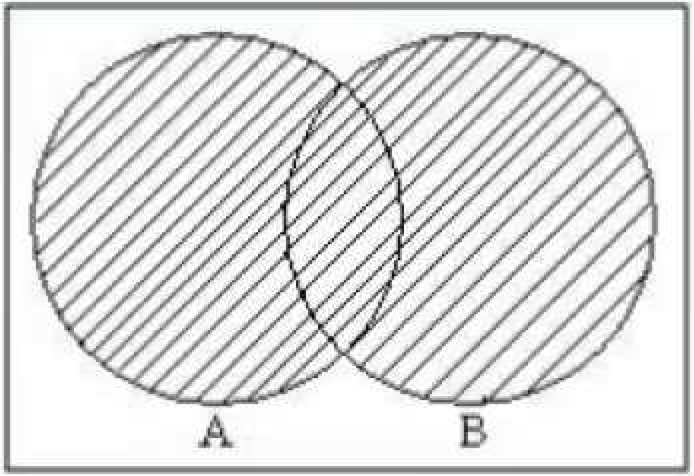


Thus event $A \cup B$ *occurs* if either A or B occurs.

Union of Events

Let A and B be two events. Then their union $A \cup B$ is:

$$A \cup B = \{s \in S \mid s \in A \text{ or } s \in B\}.$$

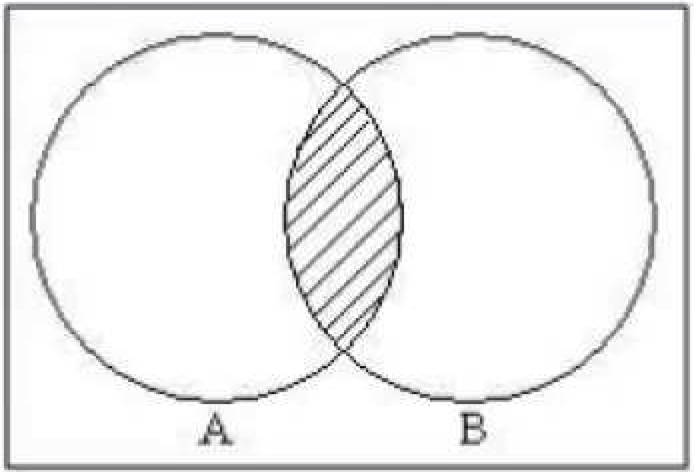


Thus event $A \cup B$ *occurs* if either A or B occurs.

Intersection of Events

Let A and B be two events. Then their intersection $A \cap B$ is:

$$A \cap B = \{s \in S \mid s \in A \text{ and } s \in B\}.$$



Thus event $A \cap B$ *occurs* if both A and B occur.

Mutually Exclusive Events

Let A and B be events such that $A \cap B = \emptyset$. Then the events are called disjoint or mutually exclusive.

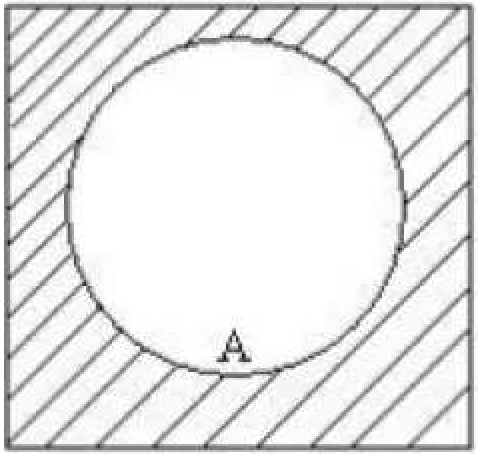
If A occurs, then B cannot occur.

If B occurs, then A cannot occur.

Complement of an Event

Let A be an event. Then its complement $\bar{A}$ (or A^c) is:

$$\bar{A} = \{s \in S \mid s \notin A\}.$$



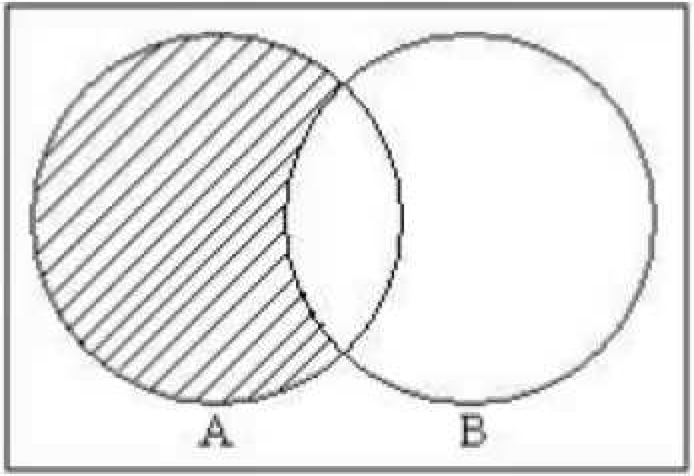
Thus $\bar{A}$ is the event that A does not occur.

(Note : Events A and $\bar{A}$ are mutually exclusive.)

Difference of Events

Let A and B be two events. Then their difference $A \setminus B$ (or $A - B$) is:

$$A \setminus B = \{s \in S \mid s \in A \text{ but } s \notin B\} = A \cap \bar{B}.$$

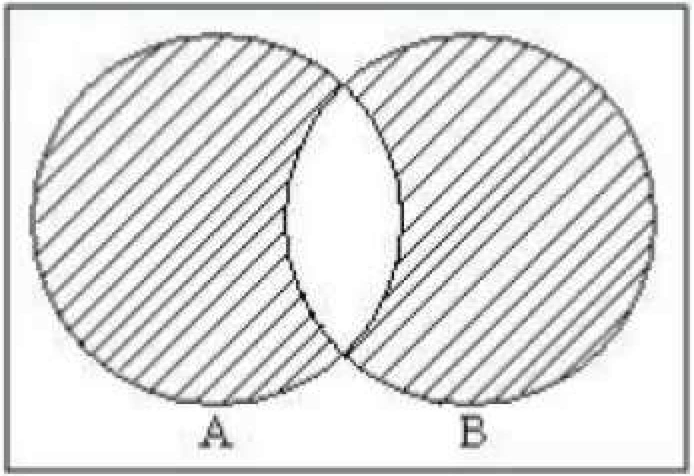


Thus event $A \setminus B$ is the event that A occurs but B does not occur.

Symmetric Difference of Events

Let A and B be two events. Then their symmetric difference $A \oplus B$ is:

$$A \oplus B = \{s \in S \mid s \in A \setminus B \text{ or } s \in B \setminus A\} = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$



Properties

- Commutative Laws: $A \cup B = B \cup A$
  $A \cap B = B \cap A$
- Associative Laws: $A \cup (B \cup C) = (A \cup B) \cup C$
  $A \cap (B \cap C) = (A \cap B) \cap C$
- Distributive Laws: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- De Morgan's Laws: $\overline{A \cup B} = \bar{A} \cap \bar{B}$
  $\overline{A \cap B} = \bar{A} \cup \bar{B}$

Problem

Let A, B, C be three events. Find expression for the event that of A, B, C,

1. Only A occurs.
2. Both A and B but not C occur.
3. All three occur.
4. At least one occurs.
5. At least two occur.
6. One and no more occurs.
7. Two and no more occur.
8. None occurs.
9. Not more than two occurs.

Countable Union and Intersection of Events

Countable Union

If $A_1, A_2, A_3, \ldots$ are events, then their union $\cup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup A_3 \cup \ldots$ is

$$\cup_{i=1}^{\infty} A_i = \{s \in S \mid s \in A_i \text{ for some } 1 \le i < \infty\}.$$

Thus $\cup_{i=1}^{\infty} A_i$ is the event that at least one of $A_1, A_2, A_3, \ldots$ occurs.

Countable Intersection

If $A_1, A_2, A_3, \ldots$ are events, then their intersection $\cap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap A_3 \cap \ldots$ is

$$\cap_{i=1}^{\infty} A_i = \{s \in S \mid s \in A_i \text{ for all } 1 \le i < \infty\}.$$

In this Module

- $\sigma-$ algebra (σ-field)
- Probability: The Axioms
- Properties of Probabilities

σ−algebra (σ-field)

Let S be a set. Let $\Sigma \subseteq P(S)$, the power set of S. Then Σ is called a σ-algebra if it satisfies the following properties.

(1) $S \in \Sigma$

(2) If $A \in \Sigma$, then $A^c = S - A \in \Sigma$ (Σ is closed under complementation)

(3) If $A_1, A_2, \ldots, A_n, \ldots \in \Sigma$, then $A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots \in \Sigma$
(Σ is closed under countable union)

Note: A σ−algebra is closed under countable intersection (Thanks to DeMorgan's law.)

Examples

(1) $\Sigma = \{S, \phi\}$

(2) Let $S = \{a, b, c\}$. $\Sigma = \{S, \phi, \{a\}, \{b, c\}\}$

$\sigma-$algebra (σ-field)

Let S be a set. Let $\Sigma \subseteq P(S)$, the power set of S. Then Σ is called a σ-algebra if it satisfies the following properties.

(1) $S \in \Sigma$

(2) If $A \in \Sigma$, then $A^c = S - A \in \Sigma$ (Σ is closed under complementation)

(3) If $A_1, A_2, \ldots, A_n, \ldots \in \Sigma$, then $A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots \in \Sigma$
(Σ is closed under countable union)

Note: A $\sigma-$algebra is closed under countable intersection (Thanks to DeMorgan's law.)

Examples

(1) $\Sigma = \{S, \phi\}$

(2) Let $S = \{a, b, c\}$. $\Sigma = \{S, \phi, \{a\}, \{b, c\}\}$

Note :

(1) Elements of a σ-algebra are caled measurable sets.

Probability

Probability means likelihood: The likelihood of the occurrence of an event.

Cricket: Between Australia and India.

Possible winning chances:

50-50

60-40

30-70

90-10

Probability: Coin Tossing

Is the coin fair? If biased, exactly how?

What is the probability of head?

Is it the limit of $\dfrac{\text{number of heads obtained}}{\text{total number of tosses made}}$?

There is no guarantee that this limit exists!

Can we find this probability at all?

Maybe not!

The goal of Probability Theory is different.

Axioms of Probability

Definition

Let S be a sample space and let Σ be a σ-algebra on S. Let $P : \Sigma \longrightarrow R^{+} \cup \{0\}$. For each event $A \in \Sigma$, assume that a number $P(A)$ is defined and satisfies the following 3 axioms:

1. $0 \leq P(A) \leq 1$.
2. $P(S) = 1$.
3. For any sequence of mutually exclusive (pairwise disjoint) events $A_1, A_2, A_3, \ldots$ ($A_i \cap A_j = \emptyset$ for $i \neq j$),

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Then the number $P(A)$ is called the *probability* of the event A.

Then (S, Σ, P) is called a probability space.

Example

Tossing a Coin. The sample space $S = \{H, T\}$. $\Sigma = \{S, \emptyset, \{H\}, \{T\}\}$

Model 1: Set $P(\{H\}) = P(\{T\}) = 1/2$. Also set $P(S) = 1$ and $P(\emptyset) = 0$.
These assignments satisfy all the axioms of probability. So, these numbers are probabilities.

Model 2: Set $P(\{H\}) = 2/3, P(\{T\}) = 1/3$. And $P(S) = 1$ and $P(\emptyset) = 0$.
These assignments also satisfy all the axioms of probability. So, these numbers are also probabilities.

How many models are possible? Let p be any fixed number, $0 \leq p \leq 1$.
Set $P(\{H\}) = p, P(\{T\}) = 1 - p$. And $P(S) = 1$ and $P(\emptyset) = 0$.

But there is only one correct model for a particular coin!

Properties of Probability

Property 1:

$P(\emptyset) = 0.$

Proof.

- Set $A_1 = S$ and $A_i = \emptyset$ for $i > 1$.
- Then $A_1, A_2, A_3, \ldots$ are mutually exclusive events and $\cup_{i=1}^{\infty} A_i = S$.
- So, $1 = P(S) = P(\cup_{i=1}^{\infty} A_i) = P(S) + \sum_{i=2}^{\infty} P(\emptyset) = 1 + \sum_{i=2}^{\infty} P(\emptyset)$.
- $\implies \quad P(\emptyset) = 0.$

$\square$

Property 2:

If $A_1, A_2, \ldots, A_n$ are mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Proof.

- Set $A_i = \emptyset$ for $i > n$.
- Then $A_1, A_2, A_3, \ldots$ are mutually exclusive and $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{n} A_i$.
- So, $P(\cup_{i=1}^{n} A_i) = P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(\emptyset)$

1.) *A box contains 3 marbles: 1 red, 1 green, and 1 blue. Consider an experiment that consists of taking 1 marble from the box and then replacing it in the box and drawing a second marble from the box. Describe the sample space. Repeat when the second marble is drawn without replacing the first marble.*

Solution

Let Red $\rightarrow$ R, Blue $\rightarrow B$, Green $\rightarrow$ G.

With replacement S_1={RR,RB,RG,BB,BR,BG,GG,GR,GB}.

Without replacement S_2={RB,RG,BR,BG,GR,GB}.

2.) *Two dice are thrown. Let E be the event that the sum of the dice is odd, let F be the event that at least one of the dice lands on 1, and let G be the event that the sum is 5. Describe the events* $E \cap F$, $E \cup F$, $F \cap G$, $E \cap F^c$, *and* $E \cap F \cap G$.

Solution

E = {(1,2),(1,4),(1,6), (2,1),(2,3),(2,5),(3,2),(3,4),(3,6),(4,1),(4,3),(4,5),(5,2),
(5,4),(5,6),(6,1),(6,3),(6,5)}
F = {(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(3,1),(4,1),(5,1),(6,1),(1,1)}
G = {(1,4),(4,1),(2,3),(3,2)}

$E \cap F$ ={at least one of the die is one and the sum of the dice is odd}
= {(1,2),(1,4),(1,6),(2,1),(4,1),(6,1)

Probability_2022_T1

mathcha.io/editor

Menu

Probability_2022_T1

Share

KIRUBAKARAN V K

$E \cap F =$ {at least one of the dice is one and the sum of the dice is odd}

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

$E \cup F =$ {at one of the dice is one or the sum is odd}

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

$F \cap G =$ {at least one of the dice is one and the sum is 5}

$= \{(1,4),(4,1)\}$

$E \cap F^c =$ {the sum of the dice is odd and no dice lands in 1}

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

$E \cap F \cap G =$ *{the sum of the dice is 5 and at least one of the dice is 1}*

$= \{(1,4),(4,1)\}$

125% Lang: None, Words: 499, Math: 24 More

3.) *A system is comprised of 5 components, each of which is either working or failed. Consider an experiment that consists of observing the status of each component, and let the outcome of the experiment be given by the vector* $(x_1, x_2, x_3, x_4, x_5)$, *where* x_i *is equal to 1 if component* i *is working and is equal to 0 if component* i *is failed.*

(a) How many outcomes are in the sample space of this experiment?

$|S| = 2^5 = 32$

(b) Suppose that the system will work if components 1 an

$|S| = 2^5 = 32$

(b) Suppose that the system will work if components 1 and 2 are both working, or if components 3 and 4 are both working, or if components 1, 3, and 5 are all working. Let W be the event that the system will work. Specify all the outcomes in W.

$$
\begin{aligned}
W = \{&(1,1,0,0,0),(1,1,1,1,0),(1,1,1,1,1),(1,1,1,0,1),(1,1,0,1,1),\\
&(1,1,0,0,1),(1,1,1,0,0),(1,1,0,1,0),\ (0,0,1,1,0),\ (0,0,1,1,1),\\
&(0,1,1,1,0),(0,1,1,1,1),(1,0,1,1,0),(1,0,1,1,1),(1,0,1,0,1)\}
\end{aligned}
$$

(c) Let A be the event that components 4 and 5 are both failed. How many outcomes are contained in the event A?

$$|A| = 2\times 2\times 2 = 8$$

(d) Write out all the outcomes in the event $A \cap W$.

system will work. Specify all the outcomes in W.

$$W = \{(1,1,0,0,0),(1,1,1,1,0),(1,1,1,1,1),(1,1,1,0,1),(1,1,0,1,1),$$
$$(1,1,0,0,1),(1,1,1,0,0),(1,1,0,1,0),\ (0,0,1,1,0),\ (0,0,1,1,1),$$
$$(0,1,1,1,0),(0,1,1,1,1),(1,0,1,1,0),(1,0,1,1,1),(1,0,1,0,1)\}$$

(c) Let A be the event that components 4 and 5 are both failed. How many outcomes are contained in the event A?

$$|A| = 2\times2\times2 = 8$$

(d) Write out all the outcomes in the event $A\cap W$.

$$A\cap W = \{(1,1,0,0,0),(1,1,1,0,0)\}$$

$A \cap W = \{(1,1,0,0,0),(1,1,1,0,0)\}$

4.) *A hospital administrator codes incoming patients suffering gunshot wounds according to whether they have insurance (coding 1 if they do and 0 if they do not) and according to their condition, which is rated as good (g), fair (f), or serious (s). Consider an experiment that consists of the coding of such a patient.*

(a) Give the sample space of this experiment.

$S = \{(1,g),(1,f),(1,s),(0,g),(0,f),(0,s)\}$

(b) Let A be the event that the patient is in serious condition. Specify the outcomes in A.

$A = \{$

(c) Let B be the event that the patient is uninsured. Speci

(b) Let A be the event that the patient is in serious condition. Specify the outcomes in A.

$A = \{(1, s), (0, s)\}$

(c) Let B be the event that the patient is uninsured. Specify the outcomes in B.

B=$\{(0, g)(0, f), (0, s)\}$

(d) Give all the outcomes in the event $B^c \cup A$.

$B^c = \{(1,g),(1,f),(1,s)\}$

$B^c \cup A = \{(1,g),(1,f),(1,s),(0,s)\}$

Property 2:

If $A_1, A_2, \ldots, A_n$ are mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Proof.

- Set $A_i = \emptyset$ for $i > n$.
- Then $A_1, A_2, A_3, \ldots$ are mutually exclusive and $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{n} A_i$.
- So, $P(\cup_{i=1}^{n} A_i) = P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(\emptyset) = \sum_{i=1}^{n} P(A_i)$ by Property 1.

$\square$

Example

Rolling a Die: The sample space is $S = \{1, 2, 3, 4, 5, 6\}$.Assume that all the outcomes are equally likely:

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}).$$

By Property 2,

$$P(\{1\} \cup \{2\} \cup \ldots \cup \{6\}) = P(\{1\}) + P(\{2\}) + \ldots + P(\{6\}).$$

Also

$$P(\{1\} \cup \{2\} \cup \ldots \cup \{6\}) = P(\{1, 2, 3, 4, 5, 6\}) = P(S) = 1.$$

Thus

$$P(\{1\}) + P(\{2\}) + \ldots + P(\{6\}) = 1.$$

and

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}.$$

Rolling a Die: Example *Contd.*

Thus the outcomes are equally likely means:

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}.$$

The probability of the event $A = \{1, 2, 3, 4\}$ is

$$P(A) = P(\{1\} \cup \{2\} \cup \{3\} \cup \{4\}) = P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) = \frac{4}{6}.$$

The probability of the event $B = \{1, 3, 5\}$ is

$$P(B) = P(\{1\} \cup \{3\} \cup \{5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) = \frac{3}{6}.$$

More generally, the probability of any event E is

$$P(E) = \frac{|E|}{|S|}.$$

Equally Likely Outcomes

Let the sample space be $S = \{1, 2, 3, \ldots, n\}$. The outcomes are equally likely means:

$$P(\{1\}) = P(\{2\}) = \ldots = P(\{n\}).$$

We have the following also true:

$$P(\{1\} \cup \{2\} \cup \ldots \cup \{n\}) = P(\{1\}) + P(\{2\}) + \ldots + P(\{n\}).$$

$$P(\{1\} \cup \{2\} \cup \ldots \cup \{n\}) = P(\{1, 2, 3, \ldots, n\}) = P(S) = 1.$$

$$\therefore\ P(\{1\}) + P(\{2\}) + \ldots + P(\{n\}) = 1.$$

So,

$$P(\{1\}) = P(\{2\}) = \ldots = P(\{n\}) = \frac{1}{n}.$$

If A is an event with m elements, then

$$P(A) = m \cdot \frac{1}{n} = \frac{m}{n} = \frac{|A|}{|S|}.$$

Problem

We are rolling a pair of dice. If all the 36 outcomes are equally likely, what is the probability that the sum of the upturned faces equals 7?

Problem

We are rolling a pair of dice. If all the 36 outcomes are equally likely, what is the probability that the sum of the upturned faces equals 7?

Solution: The event that the sum of the upturned faces equals 7 is

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

Example

A coin is tossed 6 times. Assume that all the $2^6 = 64$ outcomes are equally likely. Find the probability of obtaining 3 heads.

Example

A coin is tossed 6 times. Assume that all the $2^6 = 64$ outcomes are equally likely. Find the probability of obtaining 3 heads.

Solution: Let A be the event that exactly 3 heads are obtained.

$$\underline{\quad}\ \underline{\quad}\ \underline{\quad}\ \underline{\quad}\ \underline{\quad}\ \underline{\quad}$$

Then

$$|A| = \binom{6}{3} = \frac{6!}{3!3!} = 20.$$

So,

$$P(A) = \frac{20}{64} = \frac{5}{16}.$$

Note: Thus we must refresh our knowledge of counting techniques.

Discrete Probability Spaces

Definition:

A sample space S is called discrete if either S is finite or countably infinite:

$$S = \{s_1, s_2, \ldots, s_n\} \text{ or } S = \{s_1, s_2, s_3, \ldots\}.$$

Remark : A set A is countable if there exists a 1-1 map from the set A onto $\mathbb{N} = \{1, , 2, 3, \ldots\}$

Definition:

A probability space (S, Σ, P) is called discrete if its sample space S is discrete.

Discrete Probability Spaces

Note:

Let (S, Σ, P) be a discrete probability space and let A be an event. Then

$$P(A) = \sum_{s \in A} P(\{s\}).$$

For example, the event $A = \{s_1, s_2, s_3\}$ is the union of the disjoint events $\{s_1\}, \{s_2\}, \{s_3\}$.

So, $P(A) = P(\{s_1, s_2, s_3\}) = P(\{s_1\} \cup \{s_2\} \cup \{s_3\}) = P(\{s_1\}) + P(\{s_2\}) + P(\{s_3\})$ by Property 2.

If $|A|$ is infinite, then the identity holds by Axiom 3.

Discrete Probability Spaces

Note:

Let (S, Σ, P) be a discrete probability space and let A be an event. Then

$$P(A) = \sum_{s \in A} P(\{s\}).$$

For example, the event $A = \{s_1, s_2, s_3\}$ is the union of the disjoint events $\{s_1\}, \{s_2\}, \{s_3\}$.

So, $P(A) = P(\{s_1, s_2, s_3\}) = P(\{s_1\} \cup \{s_2\} \cup \{s_3\}) = P(\{s_1\}) + P(\{s_2\}) + P(\{s_3\})$ by Property 2.

If $|A|$ is infinite, then the identity holds by Axiom 3.

More Properties

Property 3

Let (S, Σ, P) be a probability space and let $A \in \Sigma$. Then $P(\bar{A}) = 1 - P(A)$.

Proof.

Note that $A \cup \bar{A} = S$ and $A \cap \bar{A} = \emptyset$. $\implies$ A and $\bar{A}$ are mutually exclusive events. $\implies$

$$\begin{aligned} 1 &= P(S) \quad (\text{ Axiom 2}) \\ &= P(A \cup \bar{A}) \\ &= P(A) + P(\bar{A}) \ (\text{Property 2}) \end{aligned}$$

Property 4

Let (S, Σ, P) be a probability space and let $A, B \in \Sigma$. If $A \subseteq B$, then $P(A) \leq P(B)$.

Proof.

Since $A \subseteq B$, $B = A \cup (B \setminus A)$ and $A \cap (B \setminus A) = \emptyset$. $\Longrightarrow$ A and $(B \setminus A)$ are mutually exclusive events.

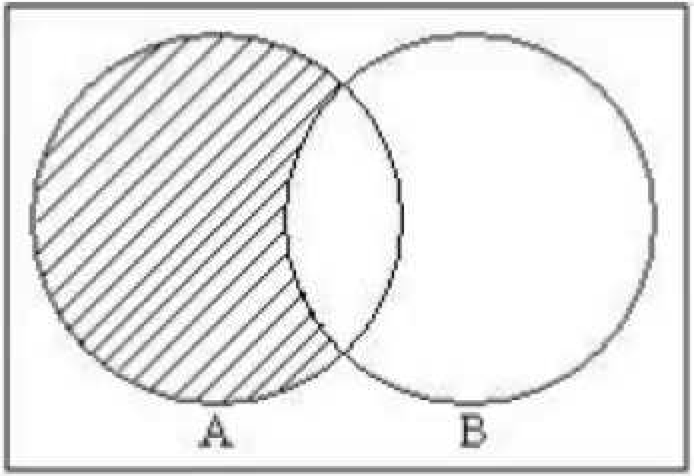
$$\begin{aligned} P(B) &= P(A \cup (B \setminus A)) \\ &= P(A) + P(B \setminus A) \quad (\textit{Property 2}) \\ &\geq P(A) \quad \text{since } P(B \setminus A) \geq 0 \quad (\textit{Axiom 1}) \end{aligned}$$

$\square$

Property 5

Let (S, Σ, P) be a probability space and let $A, B \in \Sigma$. Then

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



Proof.

$$A \cup B = (A \setminus B) \cup B$$

Since $(A \setminus B) \cap B = \phi$, by Property 2,

$$P(A \cup B) = P(A \setminus B) + P(B) \qquad --(1)$$

Note that

$$A = (A \setminus B) \cup (A \cap B)$$

Since $(A \setminus B) \cap (A \cap B) = \phi$, by Property 2,

$$P(A) = P(A \setminus B) + P(A \cap B)$$

$$\implies \quad P(A \setminus B) = P(A) - P(A \cap B) \qquad --(2)$$

From (1) and (2), $\quad P(A \cup B) = P(A) + P(B) - P(A \cap B).$

$\square$

Example: Tossing two coins

- Sample space $S = \{HH, HT, TH, TT\}$.
- Assume that all the four outcomes are equally likely.

$$P(HH) = P(HT) = P(TH) = P(TT) = \frac{1}{4}$$

- Let $A = \{HH, HT\}$ and $B = \{HT, TT\}.\Longrightarrow P(A) = P(B) = \frac{1}{2}$
- $A \cup B = \{HH, HT, TT\}$ and $A \cap B = \{HT\}$. $\Longrightarrow P(A \cup B) = \frac{3}{4}$ and $P(A \cap B) = \frac{1}{4}$
- Verify that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Problem (Homework)

Q1 : Certain fever is due to a bacterial infection with probability 0.7. And it is due to a viral infection with probability 0.4. The fever is always due to the bacterial or the viral infection mentioned. What is the probability that the fever is due to both the infections?
Q2 : A retail establishment accepts either the American Express or the VISA credit card. A total of 24 percent of its customers carry an American Express card, 61 percent carry a VISA card, and 11 percent carry both cards. What percentage of its customers carry a credit card that the establishment will accept?
Q3 : Suppose that A and B are mutually exclusive events for which $P(A) = 0.3$ and $P(B) = 0.5$. What is the probability that (a) either A or B occurs? (b) A occurs but B does not? (c) both A and B occur?

Prove that

Let (S, Σ, P) be a probability space and let $A, B \in \Sigma$. Then

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

Proof.

$$\begin{aligned}
P(A \cup B \cup C) &= P((A \cup B) \cup C) \\
&= P(A \cup B) + P(C) - P((A \cup B) \cap C) \quad (\textit{Property } 5) \\
&= P(A) + P(B) - P(A \cap B) + P(C) - P((A \cap C) \cup (B \cap C)) \\
&= P(A) + P(B) - P(A \cap B) + P(C) - [P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)] \\
&= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)
\end{aligned}$$

$\square$

Property 6 (Homework)

The principle of Inclusion-Exclusion

$$\begin{aligned} P(\cup_{i=1}^{n} A_i) &= \sum_{i=1}^{n} P(A_i) \\ &\quad - \sum_{i<j} P(A_i \cap A_j) \\ &\quad + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \\ &\quad - \dots \\ &\quad + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n) \end{aligned}$$

Proof.

By induction on n.
Basis Step: True for $n = 2$. (Property 5)
Induction Hypothesis: Assume true for $n \geq 2$.
Inductive Step: Prove the principle true for $n + 1$ assuming IH. $\square$

5.) *In an experiment, die is rolled continually until a 6 appears, at which point the experiment stops. What is the sample space of this experiment? Let E_n denote the event that n rolls are necessary to complete the experiment. What points of the sample space are contained in E_n? What is* $\left(\bigcup_{n=1}^{\infty} E_n\right)^c$?

Solution

Sample Space S={6,16,26,36,46,56,116,126,...,11111...16,...}

$$S=\{a_1a_2, \dots, a_{n-1}, 6 \,:\, a_i \in \{1,2,3,4,5\} \text{ and } n \in \mathbb{N}\}$$

Probability_2022_T1

Sample Space S={6,16,26,36,46,56,116,126,...,111111...16,...}

$$S=\{a_1a_2\cdots a_{n-1}6 \; : \; a_i \in \{1,2,3,4,5\} \; and \; n \in \mathbb{N}\}$$

S- infinite, countable.

E_n $-$ *the event that requires at least n rolls to complete the experiment*

$E_n = \{a_1a_2\cdots a_{j-1}6 \; : \; a_i \in \{1,2,3,4,5\} \; and \; j > n-1\}$

$E_n \; ? \; E_{n+1}$

$E_{n+1} \subseteq E_n$

$\left(\infty \qquad \right)^c$

E_n ? E_{n+1}

$E_{n+1} \subseteq E_n$

E_n – *the event that requires at least n rolls to complete the experiment* will contain the events that require at least $n+1$ rolls to complete the experiment.

$a_1 a_2 \cdots a_n 6 \in E_{n+1}$

Since there are n+1 terms in $a_1 a_2 \cdots a_n 6$ which is greater than n we have $a_1 a_2 \cdots a_n 6 \in E_n$

$E_{n+1}^c = E_n$

$|E_{n+1}| = 5|E_n|$

E_n is the complement of all subsets of size n of E_{n+1}

$$E_n = \{a_1 a_2 \cdots a_{j-1}6 \ : a_i \in \{1,2,3,4,5\} \ and \ j \ > n - 1\}$$

$$E_n \ ? \ E_{n+1}$$

$$E_{n+1} \ \subseteq E_n$$

E_n – *the event that requires at least n rolls to complete the experiment* will contain the events that require at least $n+1$ rolls to complete the experiment.

$$a_1 a_2 \cdots a_n 6 \in E_{n+1}$$

Since there are n+1 terms in $a_1 a_2 \ \cdots a_n 6$ which is greater than n we have $a_1 a_2 \cdots a_n 6 \in E_n$

$$E_{n+1}^c = E_n \ \cup \ E_{n-1} \cdots$$

$E_{n+1}^c \supset E_n \ \supset E_{n+1}$, a contradiction

Sample Space S={6,16,26,36,46,56,116,126,...,11111...16,...}

$$S=\{a_1a_2 \cdots a_{n-1}6 \ : \ a_i \in \{1,2,3,4,5\} \ and \ n \in \mathbb{N}\}$$

S- infinite, countable.

E_n $-$ *the event that requires at least n rolls to complete the experiment*

$$E_n = \{a_1a_2 \cdots a_{j-1}6 \ : \ a_i \in \{1,2,3,4,5\} \ and \ j > n-1\}$$

$E_n \ ? \ E_{n+1}$

$E_{n+1} \subseteq E_n$

E_n $-$ *the event that requires at least n rolls to complete the experiment* will contain the events that require at least $n+1$ rolls to complete the experiment.

$a_1a_2 \cdots a_n6 \in E_{n+1}$

E_n – *the event that requires at least n rolls to complete the experiment* will contain the events that require at least $n+1$ rolls to complete the experiment.

$a_1 a_2 \cdots a_n 6 \in E_{n+1}$

Since there are n+1 terms in $a_1 a_2 \cdots a_n 6$ which is greater than n we have $a_1 a_2 \cdots a_n 6 \in E_n$

$E_{n+1} \subseteq E_n \subseteq \ldots \subseteq E_1$=S since every outcome in S contains at least one entry, we have $E_1 = S$.

$$\left(\bigcup_{n=1}^{\infty} E_n \right)^c = \varnothing$$

6.) *A, B, and C take turns flipping a coin. The first one to get a head wins. The sample space of this experiment can be defined by*

$$S = \begin{cases} 1, 01, 001, 0001, \ldots, \\ 00000000 \cdots \end{cases}$$

(a) Interpret the sample space.

(b) Define the following events in terms of S:

(i) A wins = A.

(ii) B wins = B.

(iii) $(A \cup B)^c$.

Assume that A flips first, then B, then C, then A, and so on.

Solution

$$S = \begin{cases} 1, 01, 001, 0001, \dots, \\ 00000000 \cdots \end{cases}$$

(a) Interpret the sample space.

(b) Define the following events in terms of S:

(i) A wins = A.

(ii) B wins = B.

(iii) $(A \cup B)^c$.

Assume that A flips first, then B, then C, then A, and so on.

Solution

Sample space S is a binary string with every entry other than the last one is zero and the las.

Combinatorial Analysis

- Counting
- Permutations
- Combinations
- Multinomial coefficients

The Basic Principle of Counting

Suppose two experiments are performed. If experiment 1 has m possible outcomes and if, for each outcome of experiment 1, experiment 2 has n possible outcomes, then *the two experiments together* have mn possible outcomes.

Outline of the proof :

$$
\begin{array}{c}
(1,1)\ (1,2)\ldots(1,n) \\[1em]
(2,1)\ (2,2)\ldots(2,n) \\
\cdots \\
(m,1)\ (m,2)\ldots(m,n)
\end{array}
$$

Examples

A small community consists of 10 women, where each woman has 3 children. A woman and one of her children must be chosen for "The Mother and the Child of the Year" award. How many ways can we do this?

Solution: Consider the selection as two experiments ((i) woman and (ii) a child). The number of possible outcomes is $10 \times 3 = 30$.

Problem 2: A fancy number plate has a letter followed by a digit. What is the possible number of such plates?
Solution : $26 \times 10 = 260$.
Problem 3: A fancy number plate has two digits. What is the possible number of such plates?

Problem 4: A fancy number plate has two distinct digits. What is the possible number of such plates?

The Generalized Basic Principle of Counting

Suppose r experiments are as follows.Experiment 1 has n_1 possible outcomes.For each outcome of experiment 1, experiment 2 has n_2 possible outcomes. For each of these outcomes of the first two experiments, experiment 3 has n_3 possible outcomes and so on. Then the r experiments together have $n_1 n_2 \ldots n_r$ possible outcomes.

Exmaples

Problem 5 : A college planning committee consists 3 freshmen, 4 sophomores, 5 juniors and 2 seniors. A subcommittee of 4, consisting of 1 person from each class, is to be chosen. How many different subcommittees are possible?

Exmaples

Problem 5 : A college planning committee consists 3 freshmen, 4 sophomores, 5 juniors and 2 seniors. A subcommittee of 4, consisting of 1 person from each class, is to be chosen. How many different subcommittees are possible?

Solution: Consider choosing a freshman as experiment 1; choosing a sophomore as experiment 2; choosing a junior as experiment 3; and choosing a senior as experiment 4.

Note that a subcommittee is an outcome of performing these four experiments.

By the generalized basic principle of counting, there are $3 \times 4 \times 5 \times 2 = 120$ possible subcommittees.

Examples

Problem 6: How many different 6-place license plates are possible if the first two places are letters and the last four places are digits.

Solution: $26 \times 26 \times 10 \times 10 \times 10 \times 10 = 67,60,000.$

Problem 7 : How many such different license plates are possible if we cannot repeat letters as well as numbers.

Solution: $26 \times 25 \times 10 \times 9 \times 8 \times 7 = 32,76,000.$

Words/Strings/Sequences

Problem 8 : How many 10 letter words are possible using the English alphabet?

Solution: $26 \times 26 \times \ldots \times 26 = 26^{10}$.

Problem 9 : How many ways can 20 books of different titles be distributed among 5 children?

Solution: Let A, B, C, D, E be the 5 children. Then any word of length 20 formed using A, B, C, D, E corresponds to a way of distributing the books.

The answer is $5 \times 5 \times \ldots \times 5 = 5^{20}$.

Problem 10: How many *functions* are there from $\{1, 2, 3, \ldots, n\}$ to $\{0, 1\}$?

Solution: Each function can be viewed as a string 0s and 1s of length n.

The answer is $2 \times 2 \times \ldots \times 2 = 2^n$.

Permutations

Problem 11: How many different 4-place fancy license plates are possible if the four places have distinct digits?

Solution: $10 \times 9 \times 8 \times 7 = 5,040.$

Problem 12: How many queues are possible if 10 out of 20 people form the queue.

Solution: $20 \times 19 \times 18 \times \ldots 12 \times 11 = \dfrac{20!}{10!}.$

Problem 13: How many queues are possible if all the 20 people form the queue.

Solution: $20 \times 19 \times 18 \times \ldots 2 \times 1 = 20!.$

Permutations

A string/sequence/word is called a permutation if no element in it repeats.
A permutation on a set $S = \{1, 2, \ldots, n\}$ is a bijection from S to S.

Problem 14: How many permutations of length n can be formed using $1, 2, 3, \ldots, n$?

Solution: $n \times (n-1) \times (n-2) \times \ldots 3 \times 2 \times 1 = n!$

Problem 15: How many permutations of length r can be formed using $1, 2, 3, \ldots, n$?

Solution:

$$n \times (n-1) \times (n-2) \times \ldots \times (n-r+2) \times (n-r+1)$$

$$= \frac{n \times (n-1) \times (n-2) \times \ldots \times (n-r+2) \times (n-r+1) \times (n-r) \times (n-r-1) \times \ldots 2 \times 1}{(n-r) \times (n-r-1) \times \ldots 2 \times 1}$$

$$= \frac{n!}{(n-r)!}$$

Problem 16

If n people are present in a room, what is the probability that
at least two of them celebrate their birthday on the same day of the year?

Solution: The probability that all the n persons have different birthdays is

$$\frac{365 \times 364 \times \ldots (365-n+1)}{365^n} = \frac{365!}{365^n (365-n)!}.$$

The probability that at least two of celebrate their birthday on the same day is

$$1 - \frac{365!}{(365-n)! 365^n}.$$

Note: This probability is more than $\frac{1}{2}$ if $n \geq 23$. (Why?)

Combinations

A subset of r elements formed from a set of n elements is called a combination.

Fact: The number of r-element subsets of an n-element set is

$$\frac{n!}{r!(n-r)!}.$$

Proof: Let $S = \{1, 2, 3, \ldots, n\}$. The number of permutations of length r that can be formed using the n elements in S is

$$\frac{n!}{(n-r)!}.$$

Let k denote the number of r-element subsets of S. Using each such subset we can form $r!$ permutations. Thus the number of permutations of length r using the n elements of S is also equal to

$$k \times r!.$$

So,

$$\frac{n!}{(n-r)!} = k \times r! \Longrightarrow k = \frac{n!}{r!(n-r)!}.$$

Combinations

The number of r-element subsets of an n-element set is

$$\frac{n!}{r!(n-r)!}.$$

This number is called the binomial coefficient and is denoted by $\binom{n}{r}$:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Combinations

Note

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \binom{n}{n-r}.$$

Example

Problem: How many binary strings of length 6 are there with 3 zeros and 3 ones?

Answer: $\binom{6}{3} = 20$.

Binary Strings

The number of binary strings of length n with r 0's and $n-r$ 1's is

$$\binom{n}{r} = \binom{n}{n-r}.$$

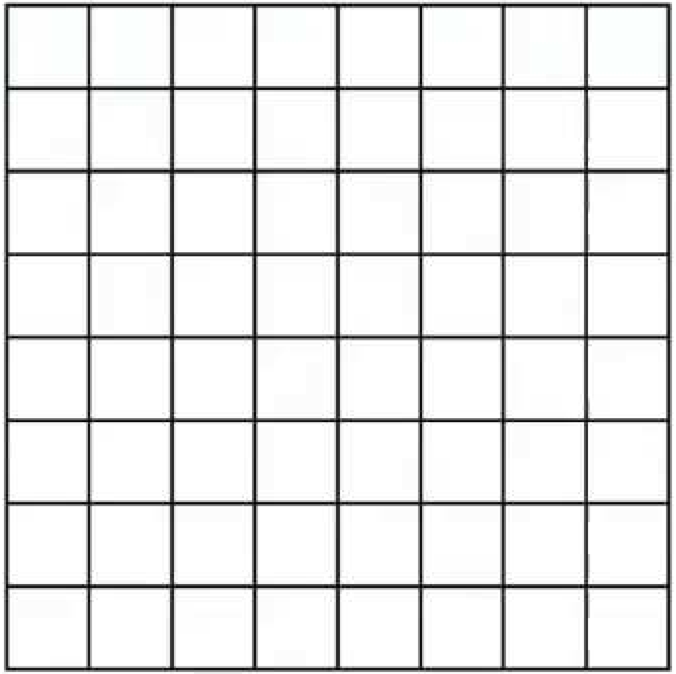
OR

The number of binary strings with m 0's and n 1's is

$$\binom{m+n}{n} = \binom{m+n}{m}.$$

Home work

How many ways can we go from the bottom-left corner to the top-right corner if we must always move either right or up along the grid?



Problem 17

A committee of 5 is to be selected from a group of 6 men and 9 women. If the selection is random, what is the probability that the committee consists of 3 men and 2 women?

Solution: The selection is random.

Note that each of the $\binom{15}{5}$ outcomes is equally likely.

There are $\binom{6}{3}\binom{9}{2}$ favourable outcomes.

Thus the required probability is

$$\frac{\binom{6}{3}\binom{9}{2}}{\binom{15}{5}} = \frac{240}{1001}.$$

Problem 18 :

Suppose that $m+n$ persons, of whom m are men and n are women, are arranged in a linear order in such a way that all the $(m+n)!$ orderings are equally likely.

If we record the result of this experiment by only listing the genders of the successive persons, show that all the possible results remain equally likely.

Each gender sequence corresponds to $m!n!$ permutations of persons.

The probability of any particular gender sequence is $\dfrac{m!n!}{(m+n)!}$.

Home work

Problem 19: How many ways can n identical balls be distributed into r distinct bins?

Problem 20: How many ways can n identical balls be distributed into r distinct bins so that no bin is left empty?

The Binomial Theorem

$$(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots + \binom{n}{n}y^n.$$

$$\begin{aligned}
(x+y)^3 &= (x+y)(x+y)(x+y) \\
&= (x+y)(xx+xy+yx+yy) \\
&= xxx+xxy+xyx+xyy+yxx+yxy+yyx+yyy \\
&= \binom{3}{0}x^3 + \binom{3}{1}x^2y + \binom{3}{2}xy^2 + \binom{3}{3}y^3 \\
&= x^3+3x^2y+3xy^2+y^3
\end{aligned}$$

The Number of Subsets

Problem 23: How many subsets are there for a set of n elements?

Solution: The number of subsets with k elements is $\binom{n}{k}$.

The answer is

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = (1+1)^n = 2^n.$$

Problem 24 : Prove that

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \ldots + (-1)^n \binom{n}{n}$$

The Matching Problem

Eample: There are N men at a party. Each is wearing a hat. They throw their hats into the center of a room and the hats are mixed up. Then each man randomly selects a hat.What is the probability that none of the men selects his own hat?

Solution: We will calculate the probability of the complementary event that at least one person selects his own hat.

Let A_i be the event that the ith man selects his own hat.Then $A_1 \cup A_2 \cup \ldots \cup A_N$ is the event that at least one person selects his own hat. And, by the principle of inclusion-exclusion,

$$\begin{aligned} P(\cup_{i=1}^{N} A_i) &= \sum_{i=1}^{N} P(A_i) \\ &- \sum_{i_1 < i_2} P(A_{i_1} \cap A_{i_2}) \\ &- \ldots \\ &+ (-1)^{n+1} \sum_{i_1 < i_2 < \ldots < i_n} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_n}) \\ &- \ldots \\ &+ (-1)^{N+1} P(A_1 \cap A_2 \cap \ldots \cap A_N) \end{aligned}$$

The Matching Problem...

We may regard the outcome of this experiment as a permutation of $1, 2, 3, \ldots, N$.

[So, the outcome $(1, 2, 3, \ldots, N)$ signifies that each man selects his own hat.]
Assume that all the $N!$ permutations are equally likely.

$$\therefore \; P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_n}) = \frac{(N-n)!}{N!}.$$

Also there are $\binom{N}{n}$ terms in $\sum_{i_1 < i_2 < \ldots < i_n} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_n})$.

$$\therefore \; \sum_{i_1 < i_2 < \ldots < i_n} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_n}) = \frac{N!}{n!(N-n)!} \frac{(N-n)!}{N!} = \frac{1}{n!}.$$

$$\therefore \; P(\cup_{i=1}^{N} A_i) = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{N+1} \frac{1}{N!}.$$

Hence the probability that none of the men selects his own hat is

$$1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots + (-1)^{N} \frac{1}{N!} \approx e^{-1} \approx 0.36788.$$

Problem

If 10 married couples are seated at random at a round table, compute the probability that no wife sits next to her husband.

The Multinomial Coefficient

How many ways can we partition n persons into 2 disjoint groups such that Group 1 has n_1 persons, Group 2 has n_2 persons, where $n_1 + n_2 = n$?

The Multinomial Coefficient

Full Screen
Close Full Screen

How many ways can we partition n persons into r disjoint groups such that Group 1 has n_1 persons, Group 2 has n_2 persons,..., Group r has n_r persons, where $n_1 + n_2 + \ldots + n_r = n$?

Solution: By the generalized basic principle of counting, the number of ways to form the r groups is

$$\binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \binom{n-n_1-n_2}{n_3} \times \ldots \times \binom{n-n_1-n_2-\ldots-n_{r-1}}{n_r}$$

$$\begin{aligned}
&= \frac{n!}{n_1!(n-n_1)!} \times \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \times \frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!} \times \ldots \times \frac{(n-n_1-n_2-\ldots-n_{r-1})}{n_r!0!} \\
&= \frac{n!}{n_1!n_2!\ldots n_r!} \\
&= \binom{n}{n_1, n_2, \ldots, n_r}.
\end{aligned}$$

Full Screen
Close Full Screen

The Multinomial Coefficient

The Multinomial Coefficient

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}.$$

Full Screen
Close Full Screen

Examples

1. How many ways can we split a club of 50 members into three committees A, B, C of 10, 20 and 20 people?

Answer:

$$\binom{50}{10, 20, 20} = \frac{50!}{10!20!20!}.$$

2. In the game of bridge, players A, B, C and D each gets 13 cards out of the 52 cards. What is the probability that A gets 6 hearts, B gets 4 hearts, C gets 2 hearts and D gets 1 heart?

Solution: The total number of outcomes is $\binom{52}{13, 13, 13, 13}$.

The total number of desired outcomes is $\binom{13}{6, 4, 2, 1}\binom{39}{7, 9, 11, 12}$.

Thus the required probability is

$$\frac{\binom{13}{6,4,2,1}\binom{39}{7,9,11,12}}{\binom{52}{13,13,13,13}}.$$

Special Strings

How many words can we form by rearranging the letters of the word PEPPER?

Solution: To form any such word we will have to fill 3 positions with a P, 2 positions with an E and one position with an R. Thus, out of the 6 positions available, the number of ways of choosing the positions for P, E and R is

$$\binom{6}{3, 2, 1} = \frac{6!}{3!2!1!} = 60.$$

The number of strings of length n with r different letters, where one letter appears n_1 times, one letter appears n_2 times, and so on, and one letter appears n_r times so that $n_1 + n_2 + \ldots + n_r = n$, is

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1!n_2!\ldots n_r!}.$$

Example

A die is rolled 12 times. What is the probability that 6 pairs are obtained?

Solution: The total number of outcomes is 6^{12}.

We have obtained 6 pairs means that we have obtained two 1's two 2's, ... , two 6's in some order.

Thus the number of desired outcomes is $\binom{12}{2,2,2,2,2,2} = \frac{12!}{(2!)^6}$.

Hence the required probability is

$$\frac{12!}{2^6 6^{12}}.$$

The Multinomial Theorem

$$(x_1 + x_2 + \ldots + x_r)^n = \sum_{n_1, n_2, \ldots, n_r : n_1 + n_2 + \ldots + n_r = n} \binom{n}{n_1, n_2, \ldots, n_r} x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}.$$

Outline

- Conditional Probability
- The Multiplication Rules
- Independence of Events
- Bayes' Formula

Conditional Probability

Motivation:

- Sometimes the prbabilities we are interested in are conditonal. Predict the agricultral productivity, provided we got a good rainy season. What is the probabilty that India wins the next test match against England at Chennai, provided we play 3 spinners
- Acts as a tool for computing probabilities.

Conditional Probability

Suppose we toss two dice. Assume that all the 36 possible outcomes are equally likely. What is the probability that the sum of the two dice equals 8?

Conditional Probability

Suppose we toss two dice. Assume that all the 36 possible outcomes are equally likely. What is the probability that the sum of the two dice equals 8?

Answer: $\frac{5}{36}$

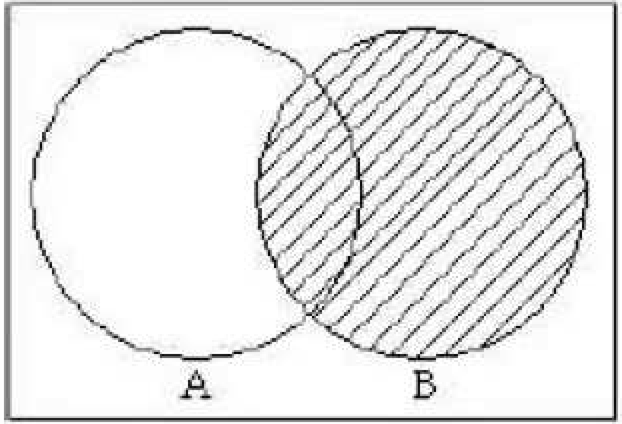
Suppose we observe that the first die is a 3. Given this information, what is the probability that the sum of the two dice equals 8?

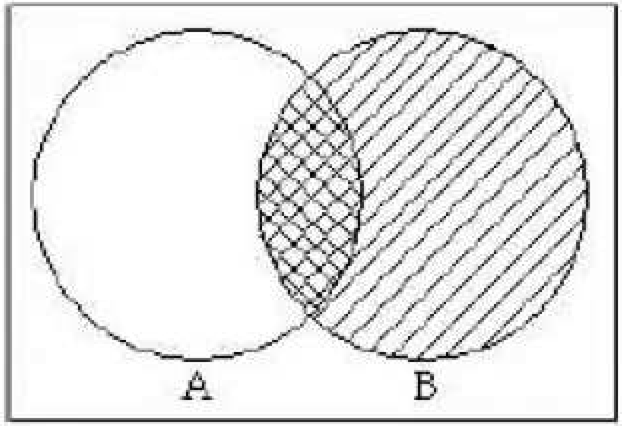
The Conditional Probability of A given B

Conditional Probability

The conditional probability $P(A|B)$ is the probability that the event A occurs given that the event B has occurred. If $P(B) > 0$, then

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





Problem 1

Suppose we toss two dice. Assume that all the 36 possible outcomes are equally likely. Given that the first die shows a 3, what is the probability that the sum of the two dice equals 8?

Problem 1

Suppose we toss two dice. Assume that all the 36 possible outcomes are equally likely. Given that the first die shows a 3, what is the probability that the sum of the two dice equals 8?

Solution: Let A be the event that the sum is 8 and B the event that the first die shows 3.

Then $A = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$ and $B = \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}$ and $A \cap B = \{(3,5)\}$.
So,

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

Problem 2

Suppose a coin is flipped twice. Assume that all the four outcomes are equally likely. What is the probability that both flips result in heads, given that the first flip does?

Problem 2

Suppose a coin is flipped twice. Assume that all the four outcomes are equally likely. What is the probability that both flips result in heads, given that the first flip does?

Solution: Let $A = \{HH\}$ and $B = \{HH, HT\}$. We are interested in $P(A|B)$.

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(\{HH\})}{P(\{HH, HT\})} \\ &= \frac{\frac{1}{4}}{\frac{1}{2}} \\ &= \frac{1}{2} \end{aligned}$$

Problem 3:

In the game bridge, the 52 cards are dealt out equally to 4 players called East, West, North, South. If North and South have a total of 8 spades among them, what is the probability that East has 3 of the remaining 5 spades?

Problem 3:

In the game bridge, the 52 cards are dealt out equally to 4 players called East, West, North, South. If North and South have a total of 8 spades among them, what is the probability that East has 3 of the remaining 5 spades?

Solution: Given: North-South have a total of 8 spades among their 26 cards.

Meaning: East-west have total of 5 spades among their 26 cards.

If East gets any 13 of these 26 cards, what is the probability that he or she will have exactly 3 spades?

$$\frac{\binom{5}{3}\binom{21}{10}}{\binom{26}{13}} \approx 0.339.$$

The Multiplication Rule

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

The Multiplication Rule

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

We obtain this by rearranging

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

Problem 4

Alice is undecided as to whether to take French or chemistry. Her estimates:

Probability of A in French: $P(A|F) = \frac{1}{2}$. Probability of A in chemistry: $P(A|C) = \frac{2}{3}$.

Alice decides to base her decision on the flip of a fair coin.What is the probability that she gets an A in chemistry?

Problem 4

Alice is undecided as to whether to take French or chemistry. Her estimates:

Probability of A in French: $P(A|F) = \frac{1}{2}$. Probability of A in chemistry: $P(A|C) = \frac{2}{3}$.

Alice decides to base her decision on the flip of a fair coin.What is the probability that she gets an A in chemistry?

Solution: We must compute $P(C \cap A)$.

$$P(C \cap A) = P(C)P(A|C) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}.$$

The Generalized Multiplication Rule

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\ P(A_2|A_1)\ P(A_3|A_1 \cap A_2)\ \ldots\ P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

Problem 5

An urn contains 8 red balls and 4 white balls. Two balls are drawn without replacement. In each draw, we assume that each ball in the urn is equally likely to be chosen.
What is the probability that both balls drawn are red?

Solution: Let R_1 and R_2 be the events that the first and second balls drawn are red. Then

$$P(R_1 \cap R_2) = P(R_1)P(R_2|R_1)$$

The Generalized Multiplication Rule

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\ P(A_2|A_1)\ P(A_3|A_1 \cap A_2)\ \ldots\ P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

Proof.

$$P(A_1)\ P(A_2|A_1)\ P(A_3|A_1 \cap A_2)\ \ldots\ P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

$$\begin{aligned} &= P(A_1)\,\frac{P(A_1 \cap A_2)}{P(A_1)}\,\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}\cdots\frac{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n)}{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1})} \\ &= P(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n). \end{aligned}$$

$\square$

The Generalized Multiplication Rule

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\ P(A_2|A_1)\ P(A_3|A_1 \cap A_2)\ \ldots\ P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

Proof.

$$\begin{aligned} & P(A_1)\ P(A_2|A_1)\ P(A_3|A_1 \cap A_2)\ \ldots\ P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \\ &= P(A_1)\ \frac{P(A_1 \cap A_2)}{P(A_1)}\ \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}\ \cdots\ \frac{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n)}{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1})} \\ &= P(A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap A_n). \end{aligned}$$

$\square$

Problem 6

An urn contains b black balls and r red balls. Four balls are randomly drawn without replacement. What is the probability that the first red is black, the second is black, the third is black and the fourth is red?

Problem 6

An urn contains b black balls and r red balls. Four balls are randomly drawn without replacement. What is the probability that the first red is black, the second is black, the third is black and the fourth is red?

Solution: Let R_1, B_2, B_3 and R_4 be the events that the the first ball is red, the second is black, the third is black and the fourth is red, respectively.

We are interested in the event that all these four events occur. So, the desired event is $R_1 \cap B_2 \cap B_3 \cap R_4$.

$$
\begin{aligned}
P(R_1 \cap B_2 \cap B_3 \cap R_4) &= P(R_1)\ P(B_2|R_1)\ P(B_3|R_1 \cap B_2)\ P(R_4|R_1 \cap B_2 \cap B_3) \\
&= \frac{r}{r+b} \times \frac{b}{r+b-1} \times \frac{b-1}{r+b-2} \times \frac{r-1}{r+b-3}
\end{aligned}
$$

Independence of Events

- $P(A|B) \neq P(A)$, in general.
  $\Longrightarrow$ knowing that B has occurred may change the probability of A.
- Sometimes: $P(A|B) = P(A)$

Note: Since $P(A|B) = \frac{P(A\cap B)}{P(B)}$,

$$P(A|B) = P(A) \Leftrightarrow P(A \cap B) = P(A)P(B).$$

Independence

Definition

Two events A and B are *independent* if $P(A \cap B) = P(A)P(B)$.

Definition

Two events A and B are *dependent* if they are not independent.

Problem 7

Consider the experiment of rolling a fair die. Consider the events $A = \{1, 3, 5\}$ and $B = \{1, 2, 3, 4\}$. Show that A and B are independent.

Problem 7

Consider the experiment of rolling a fair die. Consider the events $A = \{1, 3, 5\}$ and $B = \{1, 2, 3, 4\}$. Show that A and B are independent.

Solution: $P(A) = \frac{3}{6} = \frac{1}{2}$ and $P(B) = \frac{4}{6} = \frac{2}{3}$.

So, $P(A)P(B) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$.

Also $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$.

Thus $P(A \cap B) = P(A)P(B)$.

Hence the events A and B are independent.

Note:Independent $\neq$ Disjoint

Problem 8

A card is selected at random from a deck of 52 playing cards. Let A be the event that the selected card is an ace. Let B be the event that the selected card is a spade. Are A and B independent?

Problem 8

A card is selected at random from a deck of 52 playing cards. Let A be the event that the selected card is an ace. Let B be the event that the selected card is a spade. Are A and B independent?

Solution: $P(A) = \dfrac{4}{52} = \dfrac{1}{13}$ and $P(F) = \dfrac{13}{52} = \dfrac{1}{4}$.

So, $P(A)P(B) = \dfrac{1}{13} \times \dfrac{1}{4} = \dfrac{1}{52}$.

Also $P(A \cap B) = \dfrac{1}{52}$.

Thus $P(A \cap B) = P(A)P(B)$.

Hence the events A and B are independent.

Problem 9

Two coins are flipped. Assume that all the 4 outcomes are equally likely. Let A be the event that the first coin lands heads. Let B be the event that the second coin lands tails. Show that A and B are independent.

Problem 9

Two coins are flipped. Assume that all the 4 outcomes are equally likely. Let A be the event that the first coin lands heads. Let B be the event that the second coin lands tails. Show that A and B are independent.

Solution: $P(A) = P(\{HH, HT\}) = \frac{1}{2}$ and $P(B) = P(\{HT, TT\}) = \frac{1}{2}$.

$P(A \cap B) = P(\{HT\}) = \frac{1}{4}$.

Thus $P(A \cap B) = P(A)P(B)$.

Hence the events are independent.

Problem 10

Suppose two fair coins are flipped. What is the probability that the outcome is HT?

Solution: The desired event is $\{HT\}$.

Let A be the event that the first coin shows head. Let B be the event that the second coin shows tail.

Clearly the events are independent.

Proposition 1

If A and B are independent, then so are A and B^c.

Proof.

Let A and B be independent events.$\Longrightarrow P(A \cap B) = P(A)P(B)$. Note that $A = (A \cap B) \cup (A \cap B^c)$ and $(A \cap B) \cap (A \cap B^c) = \emptyset$. By Propery 2,
$P(A) = P(A \cap B) + P(A \cap B^c) = P(A)P(B) + P(A \cap B^c).\Longrightarrow$
$P(A \cap B^c) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c)$.
Hence A and B^c are independent.

$\square$

Proposition 2

If A and B are independent, then so are A^c and B^c.

Proof.

Let A and B be independent events.$\Longrightarrow P(A \cap B) = P(A)P(B)$. Note that $A^c \cap B^c = (A \cup B)^c$ (De Morgans law).

$$
\begin{aligned}
P(A^c \cap B^c) &= P\left((A \cup B)^c\right) \\
&= 1 - P(A \cup B) \\
&= 1 - \{P(A) + P(B) - P(A \cap B)\} \\
&= 1 - \{P(A) + P(B) - P(A)P(B)\} \\
&= (\ 1 - P(A)\ ) - P(B)(\ 1 - P(A)\ ) \\
&= (\ 1 - P(A)\ )(\ 1 - P(B)\ ) \\
&= P(A^c)P(B^c)
\end{aligned}
$$

Hence A^c and B^c are independent. (Please provide reasons for above steps !)

□

Independence of Three Events

Three events A, B, and C are said to be independent if

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

Independence of Several Events

Events $A_1, A_2, \ldots, A_n$ are said to be independent if for every subcollection $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$,

$$PA_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_k})$$

Events $A_1, A_2, A_3 \ldots$ are said to be independent if every finite subcollection $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ is independent.

Problem 11

Suppose a coin that shows heads with probability $p, 0 \leq p \leq 1$, is tossed n times. What is the probability that we get k heads?

Solution: Clearly the outcomes of the n tosses are independent.

The probability of any particular sequence with k heads, say $HH \ldots HTT \ldots T$, is $p^k(1-p)^{n-k}$.

So the required probability is

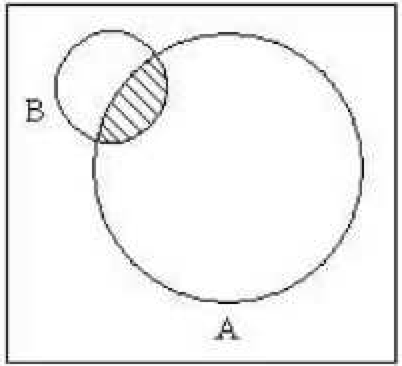
$$\binom{n}{k} p^k (1-p)^{n-k}.$$

Bayes' Formula

Let A and B be two events.Then B and B^c is a *partition* of the sample space. That is, B and B^c are disjoint and their union is equal to the sample space.

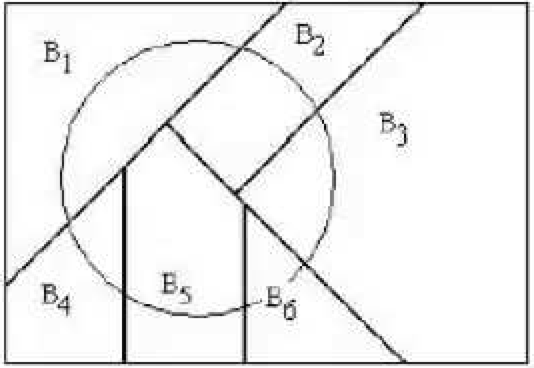
Thus $A \cap B$ and $A \cap B^c$ are disjoint and $A = (A \cap B) \cup (A \cap B^c)$. So,

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= P(A|B)P(B) + P(A|B^c)P(B^c) \\ &= P(A|B)P(B) + P(A|B^c)(1 - P(B)) \end{aligned}$$



Bayes' Formula

Let A be an event. Let $B_1, B_2, \ldots, B_n$ be pairwise disjoint events whose union is equal to S with $P(B_k) > 0$ for $1 \leq k \leq n$. That is $B_i \cap B_j = \emptyset (i \neq j)$ and $B_1 \cup B_2 \cup, \ldots, \cup B_n = S$.
($B_1, B_2, \ldots, B_n$ is called a partition of the sample space.)



Note that $B_i \cap B_j = \emptyset \Longrightarrow (A \cap B_i) \cap (A \cap B_j) = \emptyset$ for $i \neq j$.

$$A = A \cap S = A \cap (B_1 \cup B_2 \cup \ldots \cup B_n) = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_n)$$

By Property 2,

$$
\begin{aligned}
P(A) &= P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_n) \\
&= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_n)P(B_n) \\
&= \sum_{k=1}^{n} P(A|B_k)P(B_k)
\end{aligned}
$$

$$
\begin{aligned}
P(B_j|A) &= \frac{P(A \cap B_j)}{P(A)} \\
&= \frac{P(A|B_j)P(B_j)}{\sum_{k=1}^{n} P(A|B_k)P(B_k)}
\end{aligned}
$$

(Bayes Theorem)

Problem 12

In box 1, there are 60 blue balls and 40 red balls. In box 2, there are 10 blue balls and 20 red balls. A box is chosen at random; from it a ball is chosen at random. What is the probability that the selected ball is red?

Problem 12

In box 1, there are 60 blue balls and 40 red balls. In box 2, there are 10 blue balls and 20 red balls. A box is chosen at random; from it a ball is chosen at random. What is the probability that the selected ball is red?

Solution: Let B_1 be the event that box 1 is chosen and B_2 the event that box 2 is chosen.

$$P(B_1) = P(B_2) = \frac{1}{2}$$

If R is the event of choosing a red ball, then

$$P(R|B_1) = \frac{40}{100}, P(R|B_2) = \frac{20}{30}$$

Note that $R = (R \cap B_1) \cup (R \cap B_2)$.

contd.

By Bayes formula

$$P(R) = P(R|B_1)P(B_1) + P(R|B_2)P(B_2) = \frac{40}{100} \times \frac{1}{2} + \frac{20}{30} \times \frac{1}{2} = \tfrac{8}{15}.$$

Problem 13

In answering a multiple choice question, a student either knows the answer or guesses. Let p be the probability that the student knows the answer and $1-p$ the probability that the student guesses. Assume that a student who guesses will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the probability that a student knew the answer to a question given that he or she answered it correctly?

Solution: Let C denote the event that the student answers the question correctly.
Let K denote the event that the student actually knows the answer. $\Longrightarrow P(K)=p$. Then K^c denote the event that student doesn't know the answer (Student guesses the answer).
$\Longrightarrow P(K^c)=1-p$. Note that $P(C|K)=1$, $P(C|K^c)=\frac{1}{m}$ and $C=(C\cap K)\cup(C\cap K^c)$
By Bayes formula, $P(C)=P(C|K)P(K)+P(C|K^c)P(K^c)=1\times p+\frac{1}{m}(1-p)$

The required probability

$$\begin{aligned} P(K|C) &= \frac{P(K \cap C)}{P(C)} \\ &= \frac{P(C|K)P(K)}{P(C|K)P(K) + P(C|K^c)P(K^c)} \\ &= \frac{p}{p + (1/m)(1-p)} \\ &= \frac{mp}{1 + (m-1)p} \end{aligned}$$

Note (1) For $m = 4$ and $p = \frac{1}{2}$, this conditional probability is $\frac{4}{5}$.
Note: For $m = 5$ and $p = \frac{1}{2}$, this conditional probability is $\frac{5}{6}$.

Example

A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. But the test also yields a "false positive" result for 1 percent of the healthy persons tested. The fact is that 0.5 percent of the population actually has the disease. What is the probability that a person has the disease given that the test result is positive?

Random Variables

In this Module

- Random Variables
- Discrete Random Variables
- The Probability Mass Function
- Standard Discrete Random Variables

Random Variables

Motivation:

- While performing an experiment, we are often interested in some numerical value associated with the outcome rather than the outcome itself.

Random Variable (r.v.)

Definition

Let (S, Σ, P) be a probability space.

Example

Consider the experiment of tossing 3 fair coins. Let X denote the number of heads obtained. Then X is a random variable. $\implies X = 0, 1, 2, 3$.

Example

Consider the experiment of tossing 3 fair coins. Let X denote the number of heads obtained. Then X is a random variable. $\implies X = 0, 1, 2, 3$.

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

$$\{X = 0\} = \{\omega \in S : X(\omega) = 0\} = \{TTT\}$$

$$\{X = 1\} = \{\omega \in S : X(\omega) = 1\} = \{HTT, THT, TTH\}$$

$$\begin{aligned}
P(\{X = 0\}) &= P(\{TTT\}) = \frac{1}{8} \\
P(\{X = 1\}) &= P(\{HTT, THT, TTH\}) = \frac{3}{8} \\
P(\{X = 2\}) &= P(\{HHT, HTH, THH\}) = \frac{3}{8} \\
P(\{X = 3\}) &= P(\{HHH\}) = \frac{1}{8}
\end{aligned}$$

contd.

The probability mass function (pmf) of X is

x	0	1	2	3	Total
p(x) = P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

The cumulative distribution function (cdf) of X is

$$F(x) = P\{\omega \in S \ : \ X(\omega) \leq x\}$$

$$F(0) = P\{\omega \in S \ : \ X(\omega) \leq 0\} = P(X = 0) = \frac{1}{8}$$

$$F(0.75) = P\{\omega \in S \ : \ X(\omega) \leq 0.75\} = P(X = 0) = \frac{1}{8}$$

$$F(1) = P\{\omega \in S \ : \ X(\omega) \leq 1\} = P(X = 0) + P(X = 1)$$

contd.

The cdf of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{4}{8}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

contd.

The cdf of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \le x < 1 \\ \frac{4}{8}, & 1 \le x < 2 \\ \frac{7}{8}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

Sketch the cdf, $F(x)$, of X

Note that $F(-\infty) = 0$, $F(+\infty) = 1$, $F(x)$ is non-decreasing and $F(x)$ is right continuous.

Example

Consider the experiment of tossing n fair coins.

Example *Contd.*

Let $A = \{1, 2\}$. Then $\{X \in A\} = \{X = 1\} \cup \{X = 2\}$.

Thus $\{X \in A\}$ occurs if and only if $\{HTT, THT, TTH, HHT, HTH, THH\}$ occurs. So,

$$P(\{X \in A\}) = P(\{HTT, THT, TTH, HHT, HTH, THH\}) = \frac{6}{8}.$$

In particular, for $S^* = \{0, 1, 2, 3\}$

$$\begin{aligned} P(\{X \in S^*\}) &= P(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\} \cup \{X = 3\}) \\ &= P(\{X = 0\}) + P(\{X = 1\}) + P(\{X = 2\}) + P(\{X = 3\}) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} \\ &= 1. \end{aligned}$$

contd.

The cdf of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{4}{8}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

Sketch the cdf, $F(x)$, of X

Note that $F(-\infty) = 0$, $F(+\infty) = 1$, $F(x)$ is non-decreasing

contd.

The probability mass function (pmf) of X is

x	0	1	2	3	Total
p(x) = P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

The cumulative distribution function (cdf) of X is

$$F(x) = P\{\omega \in S \ : \ X(\omega) \leq x\}$$

$$F(0) = P\{\omega \in S \ : \ X(\omega) \leq 0\} = P(X = 0) = \frac{1}{8}$$

$$F(0.75) = P\{\omega \in S \ : \ X(\omega) \leq 0.75\} = P(X = 0) = \frac{1}{8}$$

$$F(1) = P\{\omega \in S \ : \ X(\omega) \leq 1\} = P(X = 0) + P(X = 1) = \frac{4}{8}$$

contd.

The cdf of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{4}{8}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

Example *Contd.*

Let $A = \{1, 2\}$. Then $\{X \in A\} = \{X = 1\} \cup \{X = 2\}$.

Example

Consider the experiment of tossing n fair coins. Let X denote the number of heads obtained. Then X is a random variable. $\Longrightarrow X = 0, 1, 2, \ldots, n$.

Note that $\{X = k\}$ is the event that k heads are obtained. Thus the probability of this event is

$$P(\{X = k\}) = \binom{n}{k} \frac{1}{2^n}.$$

We note that

$$\sum_{k=0}^{n} P(\{X = k\}) = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{2^n} = 1.$$

Discrete Random Variable

Definition

A random variable X that assumes only a finite or countably infinite number of values is called a discrete random variables.

Example 1: A binomial random variable X with parameters $(n, \frac{1}{2})$ is a discrete random variable as it assumes only $n+1$ values.

Example 2: Consider the experiment of tossing a fair coin until a head is obtained. Let X denote the number of tosses made. Then X is random variable that assumes one of $1, 2, 3, \ldots$. Thus X is a discrete random variable as it assumes only countably infinite number of values.

What is $P(X = n)$? What is $P(X > n)$?

The Probability Mass Function (pmf)

Definition

The probability mass function $p(a)$ of a discrete random variable X is given by

$$p(a) = P(X = a).$$

Note: The probability mass function $p(a)$ is positive for at most a countable number of values of a. If X assumes one of the values $x_1, x_2, x_3 \ldots$, then

$$\begin{aligned} p(x_i) &\geq 0 \quad \text{for } i = 1, 2, 3, \ldots. \\ p(x) &= 0 \quad \text{for all other values of } x. \end{aligned}$$

Note: Since X must necessarily assume one of $x_1, x_2, x_3 \ldots$, we also have that

$$\sum_{i=0}^{\infty} p(x_i) = 1.$$

Note: A discrete random variable is completely specified by its probability mass function (pmf).

The Discrete Uniform Random Variable

Definition

A random variable X is called a discrete uniform random variable if it is equally likely to assume any of the n values $1, 2, 3, \ldots, n$. It probability mass function is

$$p(k) = P(X = k) = \frac{1}{n}, \quad k = 1, 2, \ldots, n.$$

Note: In this case $p(a) = 0$ if $a \neq 1, 2, \ldots, n$.

Notation: We write $X \sim \text{Uniform}(n)$.

Example: The random number generators in computers are *almost uniform* over the given range of values.

Note

Suppose that we perform an experiment whose outcome is classified as a success or a failure with respective probabilities p and $(1-p)$.
Let $X=1$ if the experiment is a success and $X=0$ if it is a failure.
$\Longrightarrow X$ is the number of success in a single trial.

$$\begin{aligned} p(1) &= P(X=1)=p \\ p(0) &= P(X=0)=1-p \end{aligned}$$

The Bernoulli Random Variable

Definition

A random variable X is called a Bernoulli random variable with parameter $p, 0 \leq p \leq 1$, if its probability mass function (pmf) is given by

$$p(x) \;=\; P(X = x) = p^x(1-p)^{1-x} \text{ for } x = 0, 1$$

Note

Suppose that we perform an experiment whose outcome is classified as a success or a failure with respective probabilities p and $(1-p)$.
Let $X = 1$ if the experiment is a success and $X = 0$ if it is a failure.
$\Longrightarrow X$ is the number of success in a single trial.

$$\begin{aligned} p(1) &= P(X=1) = p \\ p(0) &= P(X=0) = 1-p \\ p(x) &= p^x(1-p)^{1-x} \text{ for } x = 0, 1 \end{aligned}$$

The Bernoulli Random Variable

Definition

A random variable X is called a Bernoulli random variable with parameter $p, 0 \leq p \leq 1$, if its probability mass function (pmf) is given by

$$p(x) \quad = \quad P(X = x) = p^x(1-p)^{1-x} \text{ for } x = 0, 1$$

Notation: We write $X \sim \mathrm{B}(p)$.

Note 2

Suppose that we perform an experiment whose outcome is classified as a success or a failure with respective probabilities p and $(1-p)$.

Let X be the number of successes out of n trials. $\Longrightarrow X = 0, 1, 2, \ldots, n$

$$\begin{aligned} p(0) &= P(X=0) = (1-p)^n \\ p(1) &= P(X=1) = \binom{n}{1} p(1-p)^{n-1} \\ p(x) &= \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x = 0, 1, \ldots, n \end{aligned}$$

Note 2

Suppose that we perform an experiment whose outcome is classified as a success or a failure with respective probabilities p and $(1-p)$.
Let X be the number of successes out of n trials. $\Longrightarrow X = 0, 1, 2, \ldots, n$

$$\begin{aligned}
p(0) &= P(X=0) = (1-p)^n \\
p(1) &= P(X=1) = \binom{n}{1} p(1-p)^{n-1} \\
p(x) &= \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x = 0, 1, \ldots, n
\end{aligned}$$

The Binomial Random Variable

Definition

A random variable X is called a binomial random variable with parameters $(n, p), 0 \leq p \leq 1$, if its probability mass function (pmf) is given by

$$p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n.$$

Notation: We write $X \sim \mathrm{B}(n, p)$.

Example

Screws produced by a certain company will be defective with probability 0.01 independently of one another. The company sells the screws in packages of 10 and offers a money-back guarantee that at most 1 of the 10 screws is defective. What proportion of packages sold must the company replace?

Example

Screws produced by a certain company will be defective with probability 0.01 independently of one another. The company sells the screws in packages of 10 and offers a money-back guarantee that at most 1 of the 10 screws is defective. What proportion of packages sold must the company replace?

Solution: Let X be the number of defective screws in a package. Then X is a binomial random variable with parameters $(10, 0.01)$. The pmf of X, $p(x) = P(X = x) = \binom{10}{x}(0.01)^x(0.99)^{10-x}$ for $x = 0, 1, 2, \ldots, 10$. The required probability is

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) \\ &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{10}{0}(.01)^0(.99)^{10} - \binom{10}{1}(.01)^1(.99)^9 \\ &\approx 0.04 \end{aligned}$$

Note that $0.04 = \frac{4}{100}$. Thus 4 out of each 100 packages may have to be replaced.

A simple mathematical model of crime and justice

- Siméon Denis Poisson, a French Mathematician , (early 18th century)
- Population size is large (n is large)
- p, probability of conviction is low
- $\lambda = np$, average number of convicted remains constant/moderate.
- Computing $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ is hard provided n is large.

Approximate the Binomial Random Variable (by Poisson)

Let X be a binomial random variable with parameters (n, p). Suppose n is large and p is small so that $\lambda = np$ is moderate. Then

$$
\begin{aligned}
P(X = i) &= \binom{n}{i} p^i (1-p)^{n-i} \\
&= \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\
&= \frac{n(n-1)\dots(n-i+1)}{i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}
\end{aligned}
$$

Examples

Example 1: The number of typographical errors on a single page of a book is a Poisson random variable with parameter $\lambda = \frac{1}{2}$. Calculate the probability that there is at least one error on page number 10.

Solution: Let X denote the number of errors on page 10. Then $X \sim P(\frac{1}{2})$ and $P(X = i) = e^{-\frac{1}{2}} \frac{(\frac{1}{2})^i}{i!}$ for $i = 0, 1, \ldots,$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-1/2} \approx 0.393.$$

Example 2: The probability that an item produced by a certain machine will be defective is 0.1. Find the probability that a sample of 10 items will contain at most 1 defective item.

Approximate the Binomial Random Variable (by Poisson)

Let X be a binomial random variable with parameters (n, p). Suppose n is large and p is small so that $\lambda = np$ is moderate. Then

$$\begin{aligned} P(X = i) &= \binom{n}{i} p^i (1-p)^{n-i} \\ &= \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\ &= \frac{n(n-1)\ldots(n-i+1)}{i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)\ldots(n-i+1)}{n^i} \left(\frac{\lambda^i}{i!}\right) \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^i} \end{aligned}$$

For n large and λ moderate,

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda},$$

and

$$\frac{n(n-1)\ldots(n-i+1)}{n^i} = 1 \cdot \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{i-1}{n}\right) \approx 1, \text{ and } \left(1 - \frac{\lambda}{n}\right)^i$$

Approximate the Binomial Random Variable

So,

$$P(X = i) = \frac{n(n-1)\ldots(n-i+1)}{n^i}\left(\frac{\lambda^i}{i!}\right)\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^i} \approx e^{-\lambda}\frac{\lambda^i}{i!}.$$

The Poisson Random Variable

Definition

A random variable X that assumes one of the values $0, 1, 2, 3, \ldots\ldots$ is called a Poisson random variable with parameter λ, for some $\lambda > 0$, if its probability mass function (pmf) is given by

$$p(i) = P(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}, \quad i = 0, 1, 2, \ldots.$$

Notation: We write $X \sim \text{Poisson}(\lambda)$.

Note: $\displaystyle\sum_{i=0}^{\infty} p(i) = e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = 1.$

Applications

Thus the following random variables follow Poisson distribution, each with a specific parameter λ.

- The number of misprints on a page of a book.
- The number of people in a community who live for 100 years.
- The number earthquakes on a given day.
- The number of α-particles emitted by a radioactive material in a fixed time interval.
- The number of customers entering a post office on a given day.

Examples

Example 1: The number of typographical errors on a single page of a book is a Poisson random variable with parameter $\lambda = \frac{1}{2}$. Calculate the probability that there is at least one error on page number 10.

Note

Suppose that we perform an experiment whose outcome is classified as a success or a failure with respective probabilities p and $(1-p)$.

Problem

Consider an experiment that consists of counting the number of α particles given off in a 1-second interval by 1 gram of radioactive material. From the past experience, it is known that on the average, 3.2 such α particles are given off. Find the probability that at most 2 α particles will appear.

Note

Suppose that we perform an experiment whose outcome is classified as a success or a failure with respective probabilities p and $(1-p)$.
Let X be the number of trials until a success occurs. $\Longrightarrow X = 1, 2, \ldots,$

$$\begin{aligned} p(1) &= P(X=1) = p \\ p(2) &= P(X=2) = (1-p)p \end{aligned}$$

The Geometric Random Variable

Definition

A random variable X that assumes one of the values $1, 2, 3, \ldots$ is called a geometric random variable with parameter p, $0 \le p \le 1$, if its pmf is given by

$$p(n) = P(X = n) = (1-p)^{n-1}p, \quad n = 1, 2, 3, \ldots .$$

$X \sim \text{Geometric}(p)$.

Problem: Compute $P(X \ge n)$ in two ways.

The Negative Binomial Random Variable

Definition

A random variable X that assumes one of the values $r, r+1, r+2, \ldots$ $(r \geq 1)$ is called a negative binomial random variable with parameters (r, p), where $0 \leq p \leq 1$, if its pmf is given by

$$p(n) = P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \quad n = r, r+1, r+2, \ldots.$$

Example: Suppose independent trials, where each trial has probability $p, 0 \leq p \leq 1$, of being a success, are performed until r successes occur. Let X denote the number of trials made. Then $X \sim$ Negative Binomial(r, p).

Problems

Problem 1: Independent trials, each resulting in a success with probability p, are performed. What is the probability that r successes occur before m failures?

Problem 2: A pipe-smoking scientist always carries 2 matchboxes–1 in his left-hand pocket and 1 in his right-hand pocket. Each time he needs a match, he is equally likely take it from either pocket. Consider the moment when the scientist discovers that one of his matchboxes is empty. If each matchbox initially contained N matches, what is the probability that there are exactly k matches, $k = 0, 1, 2, \ldots, N$, in the other box?

The Cumulative Distribution Function

Recall: The probability mass function $p(a)$ of a *discrete random variable* X is defined by $p(a) = P(X = a)$.

The Geometric Random Variable

Definition

A random variable X that assumes one of the values $1, 2, 3, \ldots$ is called a geometric random variable with parameter p, $0 \le p \le 1$, if its pmf is given by

$$p(n) = P(X = n) = (1-p)^{n-1}p, \quad n = 1, 2, 3, \ldots.$$

$X \sim \text{Geometric}(p)$.

Problem: Compute $P(X \ge n)$ in two ways.

Approximate the Binomial Random Variable (by Poisson)

Let X be a binomial random variable with parameters (n, p). Suppose n is large and p is small so that $\lambda = np$ is moderate.

The Discrete Uniform Random Variable

Definition

A random variable X is called a discrete uniform random variable if it is equally likely to assume any of the n values $1, 2, 3, \dots, n$. It probability mass function is

$$p(k) = P(X = k) = \frac{1}{n}, \quad k = 1, 2, \dots, n.$$

Note: In this case $p(a) = 0$ if $a \neq 1, 2, \dots, n$.

Notation: We write $X \sim \text{Uniform}(n)$.

contd.

The cdf of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{4}{8}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

Sketch the cdf, $F(x)$, of X
Note that $F(-\infty) = 0$, $F(+\infty) = 1$, $F(x)$ is non-decreasing

contd.

The probability mass function (pmf) of X is

x	0	1	2	3	Total
p(x) = P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

The cumulative distribution function (cdf) of X is

$$F(x) = P\{\omega \in S \,:\, X(\omega) \leq x\}$$

$$F(0) = P\{\omega \in S \,:\, X(\omega) \leq 0\} = P(X = 0) = \frac{1}{8}$$

$$F(0.75) = P\{\omega \in S \,:\, X(\omega) \leq 0.75\} = P(X = 0) = \frac{1}{8}$$

$$F(1) = P\{\omega \in S \,:\, X(\omega) \leq 1\} = P(X = 0) + P(X = 1) = \frac{4}{8}$$

contd.

The cdf of X is

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \le x < 1 \\ \frac{4}{8}, & 1 \le x < 2 \\ \frac{7}{8}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

Example *Contd.*

Let $A = \{1, 2\}$. Then $\{X \in A\} = \{X = 1\} \cup \{X = 2\}$.

Discrete Random Variable

Definition

A random variable X that assumes only a finite or countably infinite number of values is called a discrete random variables.

Note 2

Suppose that we perform an experiment whose outcome is classified as a success or a failure with respective probabilities p and $(1-p)$.
Let X be the number of successes out of n trials. $\Longrightarrow X = 0, 1, 2, \ldots, n$

$$p(0) \quad = \quad P(X = 0) = (1-p)^n$$

The Binomial Random Variable

Definition

A random variable X is called a binomial random variable with parameters $(n, p), 0 \leq p \leq 1$, if its probability mass function (pmf) is given by

$$p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n.$$

Notation: We write $X \sim \mathrm{B}(n, p)$.

Example

Screws produced by a certain company will be defective with probability 0.01 independently of one another. The company sells the screws in packages of 10 and offers a money-back guarantee that at most 1 of the 10 screws is defective. What proportion of packages sold must the company replace?

Solution: Let X be the number of defective screws in a package. Then X is a binomial random variable with parameters $(10, 0.01)$. The pmf of X, $p(x) = P(X = x) = \binom{10}{x}(0.01)^x(0.99)^{10-x}$ for $x = 0, 1, 2, \ldots, 10$. The required probability is

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) \\ &= 1 - P(X = 0) - P(X = 1) \end{aligned}$$

A simple mathematical model of crime and justice

- Siméon Denis Poisson, a French Mathematician , (early 18th century)
- Population size is large (n is large)
- p, probability of conviction is low

Approximate the Binomial Random Variable

So,

$$P(X=i)=\frac{n(n-1)\dots(n-i+1)}{n^i}\left(\frac{\lambda^i}{i!}\right)\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^i}\approx e^{-\lambda}\frac{\lambda^i}{i!}.$$

Applications

Thus the following random variables follow Poisson distribution, each with a specific parameter λ.

- The number of misprints on a page of a book.
- The number of people in a community who live for 100 years.
- The number earthquakes on a given day.
- The number of α-particles emitted by a radioactive material in a fixed time interval.

Examples

Example 1: The number of typographical errors on a single page of a book is a Poisson random variable with parameter $\lambda = \frac{1}{2}$. Calculate the probability that there is at least one error on page number 10.

Solution: Let X denote the number of errors on page 10. Then $X \sim P(\frac{1}{2})$ and

$P(X = i) = e^{-\frac{1}{2}} \frac{(\frac{1}{2})^i}{i!}$ for $i = 0, 1, \ldots,$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-1/2} \approx 0.393.$$

Example 2: The probability that an item produced by a certain machine will be defective is 0.1. Find the probability that a sample of 10 items will contain at most 1 defective item.

Solution: Let X denote the number of defective items in the sample of 10 items. Then $X \sim \text{Binomial}(10, 0.1)$. $\Longrightarrow P(X = x) = \binom{10}{x}(0.1)^x(0.9)^{10-x}$ for $x = 0, 1, \ldots, 10$

So, $P(X \leq 1) = P(X = 0) + P(X = 1) = \binom{10}{0}(.1)^0(.9)^{10} + \binom{10}{1}(.1)^1(.9)^9 = 0.7361.$

Poisson Approximation: We may think that X is Poisson with parameter $\lambda = np = (10)(0.1) = 1.$

Problem

Consider an experiment that consists of counting the number of α particles given off in a 1-second interval by 1 gram of radioactive material. From the past experience, it is known that on the average, 3.2 such α particles are given off. Find the probability that at most 2 α particles will appear.

The Geometric Random Variable

Definition

A random variable X that assumes one of the values $1, 2, 3, \ldots$ is called a geometric random variable with parameter $p, 0 \leq p \leq 1$, if its pmf is given by

$$p(n) = P(X = n) = (1 - p)^{n-1} p, \quad n = 1, 2, 3, \ldots.$$

Problems

Problem 1: Independent trials, each resulting in a success with probability p, are performed.

The Cumulative Distribution Function

Recall: The probability mass function $p(a)$ of a *discrete random variable* X is defined by $p(a) = P(X = a)$. It fully describes the random variable. There are others ways too.

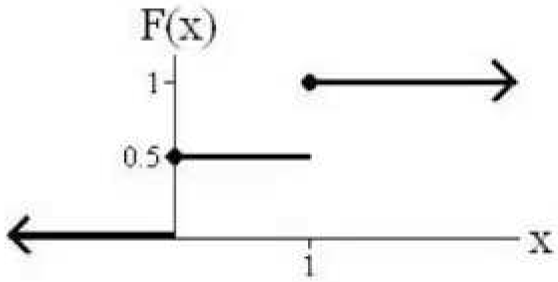
Definition

Let X be a random variable. Then the function $F(x)$ given by

$$F(x) = P(X \leq x), \quad -\infty < x < \infty,$$

is called the cumulative distribution function (cdf) of X.

Example: Let $X \sim \text{Bernoulli}(\frac{1}{2})$. Its cumulative distribution function (cdf) is given in the following figure:



Properties of the Cumulative Distribution Function

Let $F(x) = P(X \leq x)$ be the cumulative distribution function (cdf) of a random variable X. Then

1. $F(x)$ is a nondecreasing function: If $a < b$, then $F(a) \leq F(b)$.
2. $\lim_{b \to \infty} F(b) = 1$.

Why Cumulative Distribution Function?

Let X be a random variable with cdf $F(x)$. Then $F(x)$ contains all the information about the random variable X.

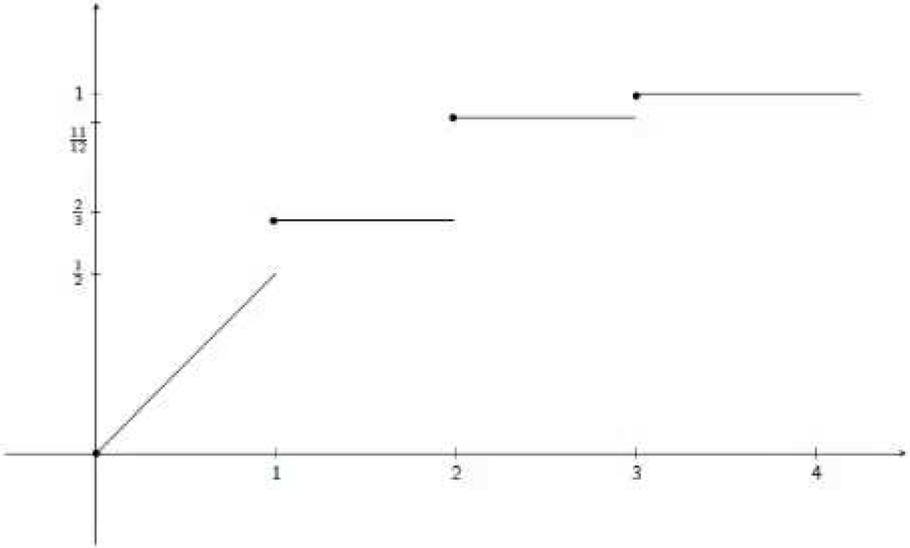
Example

The cumulative distribution function $F(x)$ of a random variable X (X is not purely discrete) is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x < 1 \\ \frac{2}{3} & 1 \leq x < 2 \\ \frac{11}{12} & 2 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

Draw the graph of $F(x)$. Compute (a) $P(X < 3)$, (b) $P(X = 1)$, (c) $P(X > \frac{1}{2})$

Solution:



(a) $P(X < 3) = \lim_{x \to 3^-} F(x) = \dfrac{11}{12}$.

(b) $P(X = 1) = F(1) - \lim_{x \to 1^-} F(x) = \dfrac{2}{3} - \dfrac{1}{2} = \dfrac{1}{6}$.

(c) $P(X > \frac{1}{2}) = 1 - P(X \leq \frac{1}{2}) = 1 - F(\frac{1}{2}) = 1 - \frac{1}{4} = \frac{3}{4}$.

(d) $P(2 < X \leq 4) = F(4) - F(2) =$

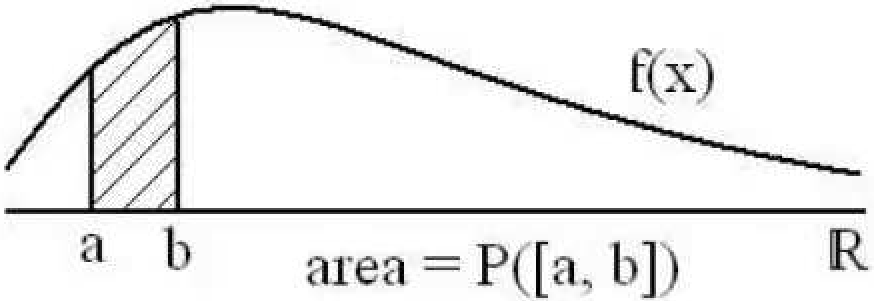
Continuous Random Variables

Definition

A random variable X is said to be a continuous random variable if there exists a non-negative function $f(x)$, defined for all real x, $-\infty < x < \infty$, such that for any set of B of real numbers

$$P(X \in B) = \int_B f(x)dx.$$

The function $f(x)$ is called the probability density function (pdf) of the random variable X.



Continuous Random Variables: Note

- $P(X \in (-\infty, \infty)) = \int_{-\infty}^{\infty} f(x)dx = 1.$
- $P(X \in [a, b]) = \int_{a}^{b} f(x)dx.$
- $P(X = a) = \int_{a}^{a} f(x)dx = 0.$
- $P(X < a) = P(X \leq a) = F(a) = \int_{-\infty}^{a} f(x)dx.$

Example

The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that

(a) the computer will function between 50 and 150 hours before breaking down?

(b) it will function for fewer than 100 hours?

Solution: Let X denote the life time of the computer. We are given the probability density function $f(x)$ (pdf) of X.

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} \lambda e^{-x/100} dx = \lambda(-100)\left[e^{-x/100}\right]_0^{\infty} = 100\lambda.$$

But $\int_{-\infty}^{\infty} f(x)dx = 1$. So, $100\lambda = 1 \Rightarrow \lambda = \frac{1}{100}$.

(a) $P(50 < X < 150) = \int_{50}^{150} \frac{1}{100} e^{-x/100} dx = \left[-e^{-x/100}\right]_{50}^{150} = e^{-1/2} - e^{-3/2} \approx 0.383.$

(b) $P(X < 100) = \int_0^{100} \frac{1}{100} e^{-x/100} dx = \left[-e^{-x/100}\right]_0^{100} = 1 - e^{-1} \approx 0.632.$

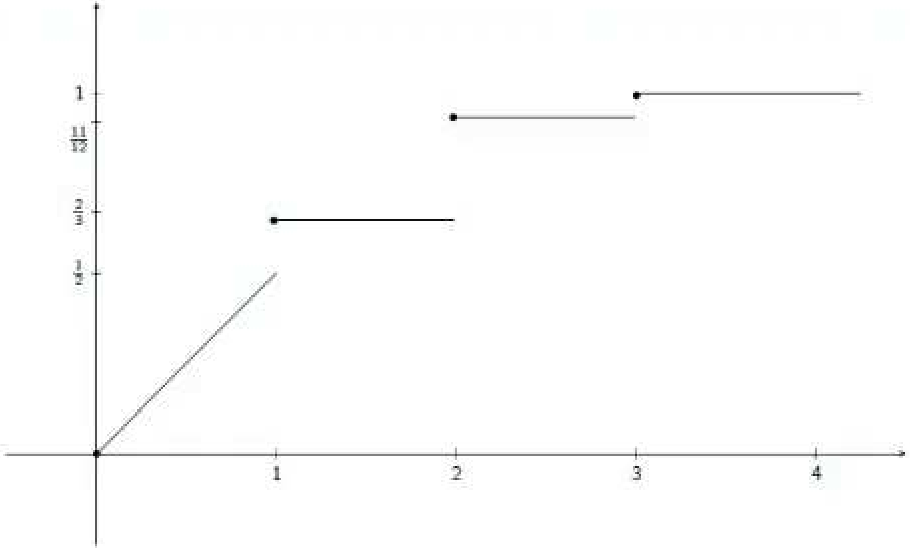
Continuous Random Variables

Definition

A random variable X is said to be a continuous random variable if there exists a non-negative function $f(x)$, defined for all real x, $-\infty < x < \infty$, such that for any set of B of real numbers

$$P(X \in B) = \int_B f(x)dx.$$

Solution:



(a) $P(X < 3) = \lim_{x \to 3^-} F(x) = \dfrac{11}{12}$.

(b) $P(X = 1) = F(1) - \lim_{x \to 1^-} F(x) = \dfrac{2}{3} - \dfrac{1}{2} = \dfrac{1}{6}$.

(c) $P(X > \frac{1}{2}) = 1 - P(X \leq \frac{1}{2}) = 1 - F(\frac{1}{2}) = 1 - \frac{1}{4} = \frac{3}{4}$.

(d) $P(2 < X \leq 4) = F(4) - F(2) = 1 - \frac{11}{12} = \frac{1}{12}$.

Continuous Random Variables

Definition

A random variable X is said to be a continuous random variable if there exists a non-negative function $f(x)$, defined for all real $x, -\infty < x < \infty$, such that for any set of B of real numbers

$$P(X \in B) = \int_B f(x)dx.$$

Properties of the Probability Density Function

Let X be a continuous random variable with pdf $f(x)$.

Property 1:

$$P(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}) = \int_{a-\frac{\epsilon}{2}}^{a+\frac{\epsilon}{2}} f(x)dx \approx \epsilon f(a).$$

Example

The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that

(a) the computer will function between 50 and 150 hours before breaking down?

(b) it will function for fewer than 100 hours?

Solution: Let X denote the life time of the computer. We are given the probability density function $f(x)$ (pdf) of X.

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} \lambda e^{-x/100} dx = \lambda(-100)\left[e^{-x/100}\right]_0^{\infty} = 100\lambda.$$

But $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$. So, $100\lambda = 1 \Rightarrow \lambda = \frac{1}{100}$.

(a) $P(50 < X < 150) = \displaystyle\int_{50}^{150} \frac{1}{100} e^{-x/100} dx = \left[-e^{-x/100}\right]_{50}^{150} = e^{-1/2} - e^{-3/2} \approx 0.383.$

(b) $P(X < 100) = \displaystyle\int_{0}^{100} \frac{1}{100} e^{-x/100} dx = \left[-e^{-x/100}\right]_{0}^{100} = 1 - e^{-1} \approx 0.632.$

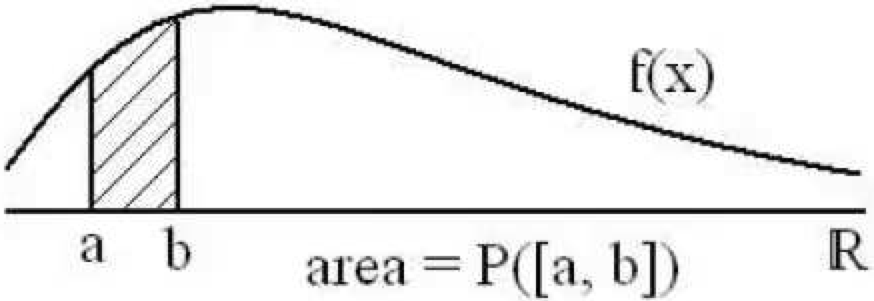
Continuous Random Variables

Definition

A random variable X is said to be a continuous random variable if there exists a non-negative function $f(x)$, defined for all real $x, -\infty < x < \infty$, such that for any set of B of real numbers

$$P(X \in B) = \int_B f(x)dx.$$

The function $f(x)$ is called the probability density function (pdf) of the random variable X.



Continuous Random Variables: Note

- $P(X \in (-\infty, \infty)) = \int_{-\infty}^{\infty} f(x)dx = 1.$
- $P(X \in [a, b]) = \int_{a}^{b} f(x)dx.$
- $P(X = a) = \int_{a}^{a} f(x)dx = 0.$
- $P(X < a) = P(X \leq a) = F(a) = \int_{-\infty}^{a} f(x)dx.$

Example

The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that

(a) the computer will function between 50 and 150 hours before breaking down?

(b) it will function for fewer than 100 hours?

Solution: Let X denote the life time of the computer. We are given the probability density function $f(x)$ (pdf) of X.

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} \lambda e^{-x/100} dx = \lambda(-100)\left[e^{-x/100}\right]_0^{\infty} = 100\lambda.$$

But $\int_{-\infty}^{\infty} f(x)dx = 1$. So, $100\lambda = 1 \Rightarrow \lambda = \frac{1}{100}$.

(a) $P(50 < X < 150) = \int_{50}^{150} \frac{1}{100} e^{-x/100} dx = \left[-e^{-x/100}\right]_{50}^{150} = e^{-1/2} - e^{-3/2} \approx 0.383.$

(b) $P(X < 100) = \int_0^{100} \frac{1}{100} e^{-x/100} dx = \left[-e^{-x/100}\right]_0^{100} = 1 - e^{-1}$

Properties of the Probability Density Function

Let X be a continuous random variable with pdf $f(x)$.

Property 1:

$$P(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}) = \int_{a-\frac{\epsilon}{2}}^{a+\frac{\epsilon}{2}} f(x)dx \approx \epsilon f(a).$$

That is, the probability that X assumes a value in an interval of length ϵ around the point a is approximately $\epsilon f(a)$.

Property 2:

Let $F(a)$ be the cumulative distribution of X:

$$F(a) = P(X \leq a) = P(X \in (-\infty, a]) = \int_{-\infty}^{a} f(x)dx.$$

This implies that

$$\frac{d}{da}F(a) = f(a).$$

That is, the density function $f(a)$ is the derivative of the cumulative distributive function $F(a)$.

Standard Continuous Distributions: Uniform Random Variables

Definition

A random variable X is said to be a uniform random variable on the interval $[a, b]$ if its probability density function is given by

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

For any c, d such that $a \leq c < d \leq b$,

$$P(c \leq X \leq d) = \int_c^d f(x)dx = \int_c^d \frac{1}{b-a}dx = \frac{d-c}{b-a}.$$

That is, the probability is proportional to the length of the subinterval $[c, d]$.

Homework: Plot the graph of the pdf of $X \sim \text{Uniform}([a, b])$. Also find its cdf and plot it.

The cdf of $X \sim U(a, b)$

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

Examples

Example 1: The random variable X is uniformly distributed over the interval $[0, 10]$. Calculate the probability that (a) $X < 3$, (b) $X > 6$ and (c) $3 < X < 8$.

Example 2: Buses arrive at a specified stop at 15-minute intervals starting at 7 AM. That is, they arrive at 7, 7:15, 7:30, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits (a) less than 5 five minutes for a bus; (b) more than 10 minutes for a bus.

Examples

Example 1: The random variable X is uniformly distributed over the interval $[0, 10]$. Calculate the probability that (a) $X < 3$, (b) $X > 6$ and (c) $3 < X < 8$.

Example 2: Buses arrive at a specified stop at 15-minute intervals starting at 7 AM. That is, they arrive at 7, 7:15, 7:30, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits (a) less than 5 five minutes for a bus; (b) more than 10 minutes for a bus.

Solution: Let X denote the amount of time (in minutes) required by the passenger to reach the bus stop between 7.00 and 7.30. $X \sim U(0, 30)$. Then the pdf of X is

$$f(x) = \begin{cases} \frac{1}{30}, & 0 \leq x \leq 30 \\ 0, & \text{otherwise} \end{cases}$$

(a) $P(10 < X < 15) + P(25 < X < 30) = \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx = \frac{1}{3}$.

Example

Consider a random chord of a circle of radius r.

Example

Consider a random chord of a circle of radius r. What is the probability that the length of the chord will be greater than the side of an equilateral triangle inscribed in that circle?

Solution: Formulation I. Consider a chord of the circle obtained by choosing the distance D of the chord from the center uniformly at random from 0 to r. Then the required probability is

$$P(D < \frac{r}{2}) = \frac{r/2}{r} = \frac{1}{2}.$$

Formulation II. Consider a chord of the circle obtained by fixing one end A and choosing the other end B randomly: This means that the angle θ made by the chord with the tangent at A varies uniformly from 0° to 180°.

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Note: The cumulative distribution function of $X \sim exp(\lambda)$ is (if $a > 0$)

$$\begin{aligned} F(a) &= P(X \leq a) \\ &= \int_0^a \lambda e^{-\lambda x} dx, \\ &= \left[-e^{-\lambda x}\right]_0^a \\ &= 1 - e^{-\lambda a} \end{aligned}$$

Note that (1) $F(\infty) = 1$. (2) $P(X > a) = 1 - P(X \leq a) = 1 - F(a) = e^{-\lambda a}$

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Solution: Let X denote the length of the phone call. $X \sim exp(\frac{1}{10})$. The cdf of X is

$$F(a) = \begin{cases} 0, & a < 0 \\ 1 - e^{-\frac{a}{10}}, & a \geq 0 \end{cases}$$

Then the required probabilities are

(a) $P(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) = e^{-1} \approx 0.368.$

(b) $P(10 < X < 20) = F(20) - F(10) = e^{-1} - e^{-2} \approx 0.233.$

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Note: The cumulative distribution function of $X \sim exp(\lambda)$ is (if $a > 0$)

$$\begin{aligned} F(a) &= P(X \leq a) \\ &= \int_0^a \lambda e^{-\lambda x} dx, \\ &= \left[-e^{-\lambda x}\right]_0^a \\ &= 1 - e^{-\lambda a} \end{aligned}$$

Note that (1) $F(\infty) = 1$. (2) $P(X > a) = 1 - P(X \leq a) = 1 - F(a) = e^{-\lambda a}$

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Solution: Let X denote the length of the phone call. $X \sim exp(\frac{1}{10})$. The cdf of X is

$$F(a) = \begin{cases} 0, & a < 0 \\ 1 - e^{-\frac{a}{10}}, & a \geq 0 \end{cases}$$

Then the required probabilities are

(a) $P(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) = e^{-1} \approx 0.368.$

(b) $P(10 < X < 20) = F(20) - F(10) =$

1. Two fair dice are thrown at once. Assume that the probability of occurance of all the 36 pair of outcomes are equally likely. Let X denote the sum of the outcomes. Then, X is a random variable. Find the probability mass function and cumulative distribution function of X.

$$P(\{X=2\}) = P(\{(1,1)\}) = \frac{1}{36}$$

$$P(\{X=3\}) = P(\{(1,2),(2,1)\}) = \frac{2}{36}$$

$$P(\{X=4\}) = P(\{(1,3),(3,1),(2,2)\}) = \frac{3}{36}$$

$$P(\{X=5\}) = P(\{(1,4),(4,1),(2,3),(3,2)\}) = \frac{4}{36}$$

$$P(\{X=6\}) = P(\{(1,5),(5,1),(2,4),(4,2),(3,3)\}) = \frac{5}{36}$$

$$P(\{X=5\}) = P(\{(1,4),(4,1),(2,3),(3,2)\}) = \frac{4}{36}$$

$$P(\{X=6\}) = P(\{(1,5),(5,1),(2,4),(4,2),(3,3)\}) = \frac{5}{36}$$

$$P(\{X=7\}) = P(\{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}) = \frac{6}{36}$$

$$P(\{X=8\}) = P(\{(2,6),(6,2),(3,5),(5,3),(4,4)\}) = \frac{5}{36}$$

$$P(\{X=9\}) = P(\{(4,5),(5,4),(3,6),(6,3)\}) = \frac{4}{36}$$

$$P(\{X=10\}) = P(\{(4,6),(6,4),(5,5\}) = \frac{3}{36}$$

$$P(\{X = 9\}) = P(\{(4,5),(5,4),(3,6),(6,3)\}) = \frac{}{36}$$

$$P(\{X = 10\}) = P(\{(4,6),(6,4),(5,5)\}) = \frac{3}{36}$$

$$P(\{X = 11\}) = P(\{(6,5),(5,6)\}) = \frac{2}{36}$$

$$P(\{X = 12\}) = P(\{(6,6)\}) = \frac{1}{36}$$

$$P(\{X = 1\}) = 0$$

Cumulative Distribution Function $F(x)$

$$F(X \leq 1) = 0$$

$$F(X \leq 2) = P(\{X = 2\}) = \frac{1}{}$$

$$F(X \leq 3) = \frac{1}{36} + \frac{2}{36} = \frac{3}{36}$$

$$F(X \leq 4) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} = \frac{6}{36}$$

$$F(X \leq 5) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} = \frac{10}{36}$$

$$F(X \leq 6) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} = \frac{15}{36}$$

$$F(X \leq 7) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} = \frac{21}{36}$$

$$F(X \leq 8) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} = \frac{26}{36}$$

Probability_2022_T3

mathcha.io/editor

Menu Save Probability_2022_T3 Share KIRUBAKARAN V K

$$F(X \leq 8) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} = \frac{26}{36}$$

$$F(X \leq 9) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} = \frac{30}{36}$$

$$F(X \leq 10) = F(X \leq 9) + P(X = 10) = \frac{30}{36} + \frac{3}{36} = \frac{33}{36}$$

$$F(X \leq 11) = F(X \leq 10) + P(X = 11) = \frac{33}{36} + \frac{2}{36} = \frac{35}{36}$$

$$F(X \leq 12) = F(X \leq 11) + P(X = 12) = \frac{35}{36} + \frac{1}{36} = \frac{36}{36} = 1$$

125% Lang: None, Words: 156, Math: 32 More

Probability_2022_T3

mathcha.io/editor

Menu Save Probability_2022_T3 Share KIRUBAKARAN V K

Imported Directoy
T2
T2
Imported Directoy
T3
Imported Directoy
T4
Imported Directoy
ms
Imported Directoy
ms
Probability_2022_T1
Probability_2022_T2
Probability_2022_T3
T6
Welcome To Mathcha

Shared by Others

Features Document
by mathcha.io [Snapshot]

$$F(12) = F(11) + P(X = 12) = \frac{35}{36} + \frac{1}{36} = \frac{36}{36} = 1$$

$$F(13) = P(X \leq 13) = F(12) + 0 = 1$$

$$F(14) = 1$$

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{36} & 2 \leq x < 3 \\ \vdots & \\ 1 & x \geq 12 \end{cases}$$

125% Lang: None, Words: 156, Math: 35 More

2. Classify the following random variables as Bernoulli or Binomial and find their pmf.

(a) A fair coin is thrown once and let X denote the occurance of head. (Bernoulli) p=$\frac{1}{2}$, n=1

P(x)=$\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{1-x}$ Let H be denoted as 1 and T as 0. Then $x=0,1$

(b) A fair coin is thrown five times and let X denote the number of heads occured. (Binomial) n=5, p=$\frac{1}{2}$.

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Solution: Let X denote the length of the phone call. $X \sim exp(\frac{1}{10})$. The cdf of X is

$$F(a) = \begin{cases} 0, & a < 0 \\ 1 - e^{-\frac{a}{10}}, & a \geq 0 \end{cases}$$

Then the required probabilities are

(a) $P(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) = e^{-1} \approx 0.368.$

(b) $P(10 < X < 20) = F(20) - F(10) = e^{-1} - e^{-2} \approx 0.233.$

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Note: The cumulative distribution function of $X \sim exp(\lambda)$ is (if $a > 0$)

$$\begin{aligned} F(a) &= P(X \leq a) \\ &= \int_0^a \lambda e^{-\lambda x} dx. \\ &= \left[-e^{-\lambda x}\right]_0^a \\ &= 1 - e^{-\lambda a} \end{aligned}$$

Note that (1) $F(\infty) = 1$. (2) $P(X > a) = 1 - P(X \leq a) = 1 - F(a) = e^{-\lambda a}$

The Memoryless Property of Exponetial Random Variable

Definition

A non-negative random variable X is said to be memoryless if

$$P(X > s + t | X > t) = P(X > s) \quad \text{for all } s, t \geq 0.$$

Note:

$$P(X > s + t | X > t) = \frac{P((X > s + t) \cap (X > t))}{P(X > t)} = \frac{P(X > s + t)}{P(X > t)}.$$

So, equivalently, X is memoryless if

$$P(X > s + t) = P(X > s)\ P(X > t).$$

If $X \sim exp(\lambda)$, then it is memoryless:

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Solution: Let X denote the length of the phone call. $X \sim exp(\frac{1}{10})$. The cdf of X is

$$F(a) = \begin{cases} 0, & a < 0 \\ 1 - e^{-\frac{a}{10}}, & a \geq 0 \end{cases}$$

Then the required probabilities are

(a) $P(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) =$

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Note: The cumulative distribution function of $X \sim exp(\lambda)$ is (if $a > 0$)

$$\begin{aligned} F(a) &= P(X \leq a) \\ &= \int_0^a \lambda e^{-\lambda x} dx. \\ &= \left[-e^{-\lambda x}\right]_0^a \\ &= 1 - e^{-\lambda a} \end{aligned}$$

Note that (1) $F(\infty) = 1$. (2) $P(X > a) = 1 - P(X \leq a) = 1 - F(a) = e^{-\lambda a}$

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Solution: Let X denote the length of the phone call.

The Memoryless Property of Exponetial Random Variable

Definition

A non-negative random variable X is said to be memoryless if

$$P(X > s + t | X > t) = P(X > s) \quad \text{for all } s, t \geq 0.$$

Note:

$$P(X > s + t | X > t) = \frac{P((X > s + t) \cap (X > t))}{P(X > t)} = \frac{P(X > s + t)}{P(X > t)}.$$

So, equivalently, X is memoryless if

$$P(X > s + t) = P(X > s)\, P(X > t).$$

If $X \sim exp(\lambda)$, then it is memoryless:

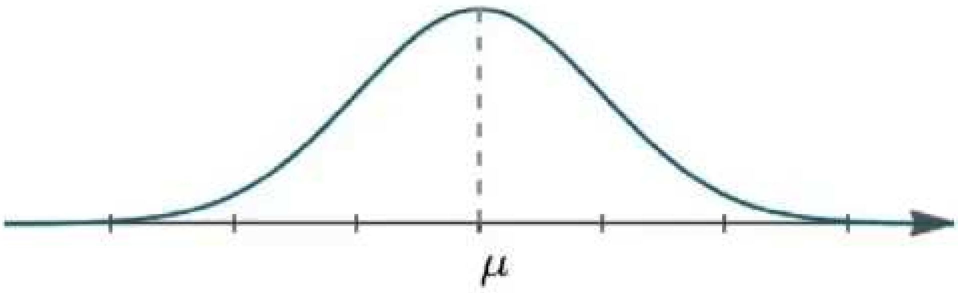
$$P(X > s + t) = e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t} = P(X > s)\, P(X > t) \quad \text{for all } s, t \geq 0.$$

Normal Random Variable, $X \sim N(\mu, \sigma^2)$

Definition

A continuous random variable X is called a normal random variable with parameters (μ, σ^2) if its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$



- Normal random variables were introduced by Abraham DeMoivre in 1733. He used it to approximate binomial probabilities when n is large.

Note

Check whether $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$ is indeed a probability density function.

Let $y = \frac{x-\mu}{\sigma}. \quad \implies \quad dy = \frac{dx}{\sigma}$.

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

Let $I = \int_{-\infty}^{\infty} e^{-y^2/2} dy$. Then show that

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2/2} e^{-y^2/2} dxdy = 2\pi.$$

(For more details, visit your differential equations text book by G.F. Simmons)

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy = 1$$

Lemma

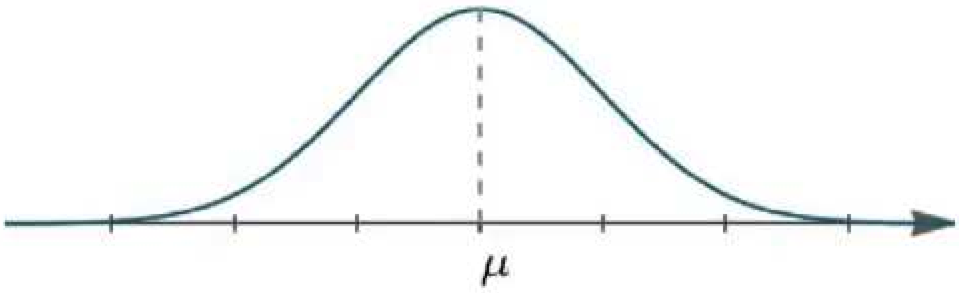
If X is a normal random variable with parameters (μ, σ^2), then $Y = aX + b$ is a normal random variable with parameters $(a\mu + b, a^2\sigma^2)$.

Normal Random Variable, $X \sim N(\mu, \sigma^2)$

Definition

A continuous random variable X is called a normal random variable with parameters (μ, σ^2) if its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$



Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Note: The cumulative distribution function of $X \sim exp(\lambda)$ is (if $a > 0$)

$$\begin{aligned} F(a) &= P(X \leq a) \\ &= \int_0^a \lambda e^{-\lambda x} dx, \\ &= \left[-e^{-\lambda x}\right]_0^a \\ &= 1 - e^{-\lambda a} \end{aligned}$$

Note that (1) $F(\infty) = 1$. (2) $P(X > a) = 1 - P(X \leq a) = 1 - F(a) = e^{-\lambda a}$

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Example

Consider a random chord of a circle of radius r. What is the probability that the length of the chord will be greater than the side of an equilateral triangle inscribed in that circle?

Solution: Formulation I. Consider a chord of the circle obtained by choosing the distance D of the chord from the center uniformly at random from 0 to r. Then the required probability is

$$P(D < \frac{r}{2}) = \frac{r/2}{r} = \frac{1}{2}.$$

Formulation II. Consider a chord of the circle obtained by fixing one end A and choosing the other end B randomly: This means that the angle θ made by the chord with the tangent at A varies uniformly from 0° to 180°.

Examples

Example 1: The random variable X is uniformly distributed over the interval $[0, 10]$. Calculate the probability that (a) $X < 3$, (b) $X > 6$ and (c) $3 < X < 8$.

Example 2: Buses arrive at a specified stop at 15-minute intervals starting at 7 AM. That is, they arrive at 7, 7:15, 7:30, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits (a) less than 5 five minutes for a bus; (b) more than 10 minutes for a bus.

Solution: Let X denote the amount of time (in minutes) required by the passenger to reach the bus stop between 7.00 and 7.30. $X \sim U(0, 30)$. Then the pdf of X is

$$f(x) = \begin{cases} \frac{1}{30}, & 0 \leq x \leq 30 \\ 0, & \text{otherwise} \end{cases}$$

(a) $P(10 < X < 15) + P(25 < X < 30) = \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx = \frac{1}{3}.$

Standard Continuous Distributions: Uniform Random Variables

Definition

A random variable X is said to be a uniform random variable on the interval $[a, b]$ if its probability density function is given by

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

For any c, d such that $a \leq c < d \leq b$,

$$P(c \leq X \leq d) = \int_c^d f(x)dx = \int_c^d \frac{1}{b-a}dx = \frac{d-c}{b-a}.$$

That is, the probability is proportional to the length of the subinterval $[c, d]$.
Homework: Plot the graph of the pdf of $X \sim \text{Uniform}([a, b])$. Also find its cdf and plot it.

Properties of the Probability Density Function

Let X be a continuous random variable with pdf $f(x)$.

Property 1:

$$P(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}) = \int_{a-\frac{\epsilon}{2}}^{a+\frac{\epsilon}{2}} f(x)dx \approx \epsilon f(a).$$

That is, the probability that X assumes a value in an interval of length ϵ around the point a is approximately $\epsilon f(a)$.

Property 2:

Let $F(a)$ be the cumulative distribution of X:

$$F(a) = P(X \leq a) = P(X \in (-\infty, a]) = \int_{-\infty}^{a} f(x)dx.$$

This implies that

$$\frac{d}{da}F(a) = f(a).$$

That is, the density function $f(a)$ is the derivative of the cumulative distributive function $F(a)$.

Examples

Example 1: The random variable X is uniformly distributed over the interval $[0, 10]$. Calculate the probability that (a) $X < 3$, (b) $X > 6$ and (c) $3 < X < 8$.

Example 2: Buses arrive at a specified stop at 15-minute intervals starting at 7 AM. That is, they arrive at 7, 7:15, 7:30, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits (a) less than 5 five minutes for a bus; (b) more than 10 minutes for a bus.

Solution: Let X denote the amount of time (in minutes) required by the passenger to reach the bus stop between 7.00 and 7.30. $X \sim U(0, 30)$. Then the pdf of X is

$$f(x) = \begin{cases} \frac{1}{30}, & 0 \leq x \leq 30 \\ 0, & \text{otherwise} \end{cases}$$

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Solution: Let X denote the length of the phone call. $X \sim exp(\frac{1}{10})$. The cdf of X is

$$F(a) = \begin{cases} 0, & a < 0 \\ 1 - e^{-\frac{a}{10}} , & a \geq 0 \end{cases}$$

Then the required probabilities are

(a) $P(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) = e^{-1} \approx 0.368.$

The Memoryless Property of Exponetial Random Variable

Definition

A non-negative random variable X is said to be memoryless if

$$P(X > s + t | X > t) = P(X > s) \quad \text{for all } s, t \geq 0.$$

Note:

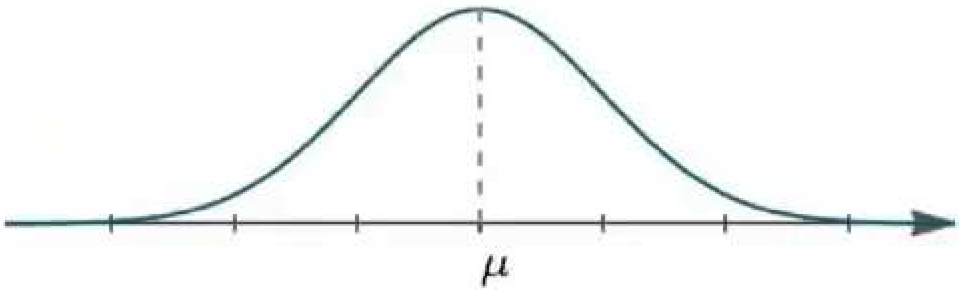
$$P(X > s + t | X > t) = \frac{P((X > s + t) \cap (X > t))}{P(X > t)} = \frac{P(X > s + t)}{P(X > t)}.$$

Normal Random Variable, $X \sim N(\mu, \sigma^2)$

Definition

A continuous random variable X is called a normal random variable with parameters (μ, σ^2) if its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$



- Normal random variables were introduced by Abraham DeMoivre in 1733. He used it to approximate binomial probabilities when n is large.
- The famous central limit theorem captures its relation to any distribution in an important sense.

Note

Check whether $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$ is indeed a probability density function.

Let $y = \frac{x-\mu}{\sigma}. \quad \Longrightarrow \quad dy = \frac{dx}{\sigma}$.

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

Let $I = \int_{-\infty}^{\infty} e^{-y^2/2} dy$. Then show that

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2/2} e^{-y^2/2} dx dy = 2\pi.$$

(For more details, visit your differential equations text book by G.F. Simmons)

Lemma

If X is a normal random variable with parameters (μ, σ^2), then $Y = aX + b$ is a normal random variable with parameters $(a\mu + b, a^2\sigma^2)$.

Proof: Let F_X and F_Y be the cumulative distribution functions of X and Y, respectively. Assume that $a > 0$ (Proof is similar for $a < 0$). Then

$$
\begin{aligned}
F_Y(x) &= P(Y \leq x) \\
&= P(aX + b \leq x) \\
&= P(X \leq \frac{x-b}{a}) \\
&= F_X(\frac{x-b}{a})
\end{aligned}
$$

Proof...

Let us differentiate w.r.t. x,

$$
\begin{aligned}
f_Y(x) &= \frac{1}{a} f_X\left(\frac{x-b}{a}\right) \\
&= \frac{1}{\sqrt{2\pi}\, a\sigma} e^{-\left(\frac{x-b}{a}-\mu\right)^2/2\sigma^2} \\
&= \frac{1}{\sqrt{2\pi}(a\sigma)} e^{-(x-(a\mu+b))^2/2(a\sigma)^2}
\end{aligned}
$$

$\square$

Note: Here in the place of 'μ' we have $a\mu + b$ and in the place of 'σ^2', we have $a^2\sigma^2$. Thus $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

Observation : $X \sim N(\mu, \sigma^2) \implies Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$

Let $X \sim N(\mu, \sigma^2)$. Let $Z = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} - \frac{\mu}{\sigma} = aX + b$. $\implies a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$.
$\implies a\mu + b = 0$ and $a^2\sigma^2 = 1$
$\implies Z \sim N(a\mu + b, a^2\sigma^2) = N(0, 1)$
This Z has a special name.

The Standard Normal Random Variable, $Z \sim N(0,1)$

Definition

If Z is a normal random variable with parameters $(0, 1)$, it is called the standard normal or unit normal random variable.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty.$$

Observations

1. If X is normal with parameters (μ, σ^2), then $Z = \frac{X-\mu}{\sigma}$ is a normal random variable with parameters (0,1) and so it is the standard normal random variable.

The Standard Normal Random Variable, $Z \sim N(0,1)$

Definition

If Z is a normal random variable with parameters $(0,1)$, it is called the standard normal or unit normal random variable.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty.$$

Observations

1. If X is normal with parameters (μ, σ^2), then $Z = \frac{X-\mu}{\sigma}$ is a normal random variable with parameters (0,1) and so it is the standard normal random variable.
2. Conversely, if Z is the standard normal random variable, then $X = \sigma Z + \mu$ is a normal random variable with parameters (μ, σ^2).

The cdf of the Standard Normal and The Normal Table

$\Phi(z)$

The cumulative distribution function of the standard normal random variable Z is usually denoted by $\Phi(z)$:

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx.$$

The Standard Normal Random Variable, $Z \sim N(0,1)$

Definition

If Z is a normal random variable with parameters $(0,1)$, it is called the standard normal or unit normal random variable.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty.$$

Observations

1. If X is normal with parameters (μ, σ^2), then $Z = \frac{X-\mu}{\sigma}$ is a normal random variable with parameters (0,1) and so it is the standard normal random variable.
2. Conversely, if Z is the standard normal random variable, then $X = \sigma Z + \mu$ is a normal random variable with parameters (μ, σ^2).

Observation : $X \sim N(\mu, \sigma^2) \Longrightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Let $X \sim N(\mu, \sigma^2)$.

Lemma

If X is a normal random variable with parameters (μ, σ^2), then $Y = aX + b$ is a normal random variable with parameters $(a\mu + b, a^2\sigma^2)$.

Proof: Let F_X and F_Y be the cumulative distribution functions of X and Y, respectively. Assume that $a > 0$ (Proof is similar for $a < 0$). Then

$$\begin{aligned} F_Y(x) &= P(Y \leq x) \\ &= P(aX + b \leq x) \\ &= P(X \leq \frac{x-b}{a}) \\ &= F_X(\frac{x-b}{a}) \end{aligned}$$

Note

Check whether $f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$ is indeed a probability density function.

Let $y = \frac{x-\mu}{\sigma}. \quad \Longrightarrow \quad dy = \frac{dx}{\sigma}$.

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

Let $I = \displaystyle\int_{-\infty}^{\infty} e^{-y^2/2} dy$. Then show that

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2/2} e^{-y^2/2} dx dy = 2\pi.$$

(For more details, visit your differential equations text book by G.F. Simmons)

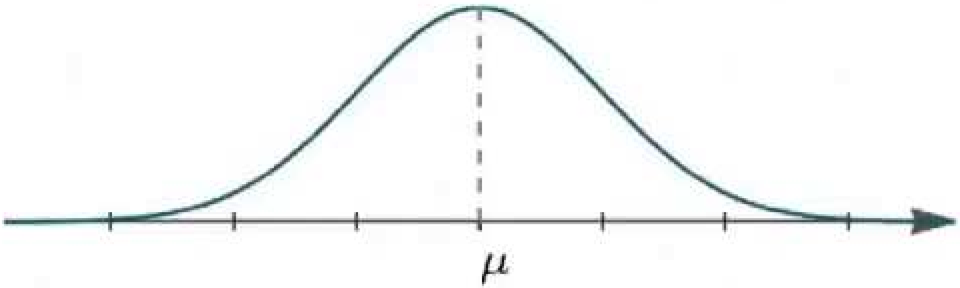
$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy = 1$$

Normal Random Variable, $X \sim N(\mu, \sigma^2)$

Definition

A continuous random variable X is called a normal random variable with parameters (μ, σ^2) if its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$



- Normal random variables were introduced by Abraham DeMoivre in 1733. He used it to approximate binomial probabilities when n is large.
- The famous central limit theorem captures its relation to any distribution in an important sense.

Note

Check whether $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$, $\quad -\infty < x < \infty \quad$ is indeed a probability density function.

Let $y = \frac{x-\mu}{\sigma}$. $\quad \Longrightarrow \quad dy = \frac{dx}{\sigma}$.

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

Let $I = \int_{-\infty}^{\infty} e^{-y^2/2} dy$.

Lemma

If X is a normal random variable with parameters (μ, σ^2), then $Y = aX + b$ is a normal random variable with parameters $(a\mu + b, a^2\sigma^2)$.

Proof: Let F_X and F_Y be the cumulative distribution functions of X and Y, respectively. Assume that $a > 0$ (Proof is similar for $a < 0$). Then

$$\begin{aligned}
F_Y(x) &= P(Y \leq x) \\
&= P(aX + b \leq x) \\
&= P(X \leq \frac{x-b}{a}) \\
&= F_X(\frac{x-b}{a})
\end{aligned}$$

Proof...

Let us differentiate w.r.t. x,

$$\begin{aligned} f_Y(x) &= \frac{1}{a} f_X(\frac{x-b}{a}) \\ &= \frac{1}{\sqrt{2\pi} a\sigma} e^{-(\frac{x-b}{a}-\mu)^2/2\sigma^2} \\ &= \frac{1}{\sqrt{2\pi}(a\sigma)} e^{-(x-(a\mu+b))^2/2(a\sigma)^2} \end{aligned}$$

$\square$

Note: Here in the place of 'μ' we have $a\mu + b$ and in the place of 'σ^2', we have $a^2\sigma^2$. Thus $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

Observation : $X \sim N(\mu, \sigma^2) \Longrightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Let $X \sim N(\mu, \sigma^2)$. Let $Z = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} - \frac{\mu}{\sigma} = aX + b$. $\Longrightarrow a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$.
$\Longrightarrow a\mu + b = 0$ and $a^2\sigma^2 = 1$
$\Longrightarrow Z \sim N(a\mu + b, a^2\sigma^2) = N(0,1)$
This Z has a special name.

The cdf of the Standard Normal and The Normal Table

$\Phi(z)$

The cumulative distribution function of the standard normal random variable Z is usually denoted by $\Phi(z)$:

$$\Phi(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx.$$

$$\Phi(-z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-z} e^{-x^2/2} dx.$$

Let $y = -x$. $x = -z \Longrightarrow y = z$ and $x = -\infty \Longrightarrow y = \infty$. $dy = -dx$

$$\begin{aligned}
\Phi(-z) &= \tfrac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{x=-z} e^{-x^2/2} dx \\
&= \tfrac{1}{\sqrt{2\pi}} \int_{y=\infty}^{y=z} e^{-y^2/2} (-dy) \\
&= \tfrac{1}{\sqrt{2\pi}} \int_{z}^{\infty} e^{-y^2/2} dy \\
&= \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy - \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2} dy
\end{aligned}$$

Note

$$\Phi(-z) = 1 - \Phi(z), \quad -\infty < z < \infty.$$

Note: The "Normal Table" usually gives the values of $\Phi(z)$ for $z \geq 0$. For negative z, the above identity can be used.
Page number 205 (Standard Normal Curve table)

Example

If X is a normal random variable with parameters $(\mu, \sigma^2) = (3, 9)$, find (a) $P(2 < X < 5)$; (b) $P(X > 0)$; (c) $P(|X - 3| > 6)$. Note that
$X \sim N(\mu = 3, \sigma^2 = 3^2) \Longrightarrow Z = \frac{X-\mu}{\sigma} = \frac{X-3}{3} \sim N(0,1)$

Example

If X is a normal random variable with parameters $(\mu, \sigma^2) = (3, 9)$, find (a) $P(2 < X < 5)$; (b) $P(X > 0)$; (c) $P(|X - 3| > 6)$. Note that
$X \sim N(\mu = 3, \sigma^2 = 3^2) \Longrightarrow Z = \frac{X-\mu}{\sigma} = \frac{X-3}{3} \sim N(0, 1)$

Solution: (a)

$$\begin{aligned} P(2 < X < 5) &= P(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}) \\ &= P(\frac{-1}{3} < Z < \frac{2}{3}) \\ &= \Phi(\frac{2}{3}) - \Phi(\frac{-1}{3}) = \Phi(\frac{2}{3}) - (1 - \Phi(\frac{1}{3})) \end{aligned}$$

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545

Example

If X is a normal random variable with parameters $(\mu, \sigma^2) = (3, 9)$, find (a) $P(2 < X < 5)$; (b) $P(X > 0)$; (c) $P(|X - 3| > 6)$. Note that $X \sim N(\mu = 3, \sigma^2 = 3^2) \Longrightarrow Z = \frac{X-\mu}{\sigma} = \frac{X-3}{3} \sim N(0,1)$

Solution: (a)

$$\begin{aligned} P(2 < X < 5) &= P(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}) \\ &= P(\frac{-1}{3} < Z < \frac{2}{3}) \\ &= \Phi(\frac{2}{3}) - \Phi(\frac{-1}{3}) = \Phi(\frac{2}{3}) - (1 - \Phi(\frac{1}{3})) \\ &\approx 0.7486 - 1 + 0.6293 = 0.3779 \end{aligned}$$

(b)

$$P(X > 0) = P(\frac{X-3}{3} > \frac{0-3}{3}) \quad = \quad P(Z > -1)$$

The DeMoivre-Laplace Theorem

Theorem

Let $0 \leq p \leq 1$ be fixed. Let X_n be a binomial random variable with parameters (n, p). Then for any $a < b$,

$$P\left(a \leq \frac{X_n - np}{\sqrt{np(p-1)}} \leq b\right) \to \Phi(b) - \Phi(a)$$

as $n \to \infty$.

Note: This is a special case of the Central Limit Theorem (proof later)

The Gamma Function

Definition

The gamma function $\Gamma(\alpha)$ is given by

$$\Gamma(\alpha) = \int_0^{\infty} e^{-x} x^{\alpha-1} dx, \quad \alpha > 0$$

Properties

1. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for $\alpha > 1$. (Prove!)
2. $\Gamma(1) = 1$. (Prove!)
3. $\Gamma(n+1) = n!$ for any interger $n \geq 0$:
   $\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \ldots = n(n-1)\ldots 2 \cdot \Gamma(1) = n!$.

Gamma Random Variables

Definition

A random variable X is called a gamma random variable with parameters (α, λ), where $\lambda > 0$ and $\alpha > 0$, if its density function is given by

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Note

1. If $\alpha = 1$, X is simply the exponential random variable with parameter λ.

Functions of Random Variables

Note: If X is a random variable, then $Y = g(X)$ is also a random variable for any (real-valued) function g of X.

Example: Let X be a discrete random variable assuming one of the three values $-1, 0$ and 1 with probability $\frac{1}{3}$. Then $Y = X^2$ is a random variable assuming the values 0 and 1 .

$$P(Y = 0) = P(X = 0)$$

Question: Given the pmf p_X of a random variable X, what is the pmf of p_Y of $Y = g(X)$?

Question: Given the pmf p_X of a random variable X, what is the pmf of p_Y of $Y = g(X)$?

Answer: Suppose that X assumes the values $x_1, x_2, \ldots$ with positive probabilities

Functions of Random Variables: Continuous Case

Problem: Let X be the uniform random variable on the interval $[0, 1]$. Find the density function (pdf) of the random variable $Y = X^2$.

Solution: We first find the cumulative distribution function F_Y (cdf) of Y. Then we differentiate it to obtain the density function (pdf) of Y. The cdf of X is

$$F_x(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

Question: Given the pmf p_X of a random variable X, what is the pmf of p_Y of $Y = g(X)$?

Answer: Suppose that X assumes the values $x_1, x_2, \ldots$ with positive probabilities

The cdf of the Standard Normal and The Normal Table

$\Phi(z)$

The cumulative distribution function of the standard normal random variable Z is usually denoted by $\Phi(z)$:

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx.$$

$$\Phi(-z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-z} e^{-x^2/2} dx.$$

Let $y = -x$. $x = -z \Longrightarrow y = z$ and $x = -\infty \Longrightarrow y = \infty$. $dy = -dx$

$$\begin{aligned}
\Phi(-z) &= \tfrac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{x=-z} e^{-x^2/2} dx \\
&= \tfrac{1}{\sqrt{2\pi}} \int_{y=\infty}^{y=z} e^{-y^2/2} (-dy) \\
&= \tfrac{1}{\sqrt{2\pi}} \int_{z}^{\infty} e^{-y^2/2} dy \\
&= \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy - \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2} dy
\end{aligned}$$

Lemma

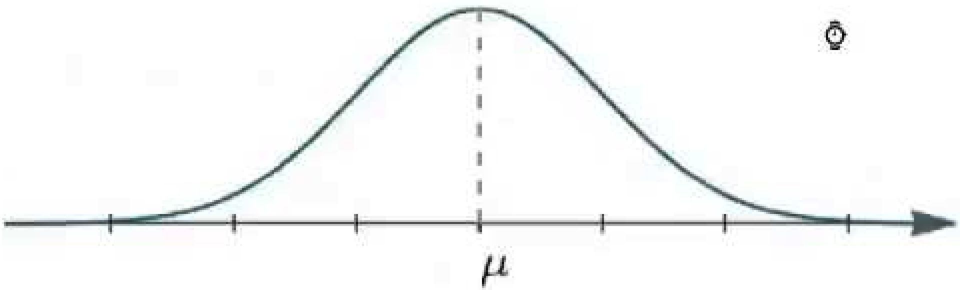
If X is a normal random variable with parameters (μ, σ^2), then $Y = aX + b$ is a normal random variable with parameters $(a\mu + b, a^2\sigma^2)$.

Normal Random Variable, $X \sim N(\mu, \sigma^2)$

Definition

A continuous random variable X is called a normal random variable with parameters (μ, σ^2) if its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$



- Normal random variables were introduced by Abraham DeMoivre in 1733. He used it to approximate binomial probabilities when n is large.

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Note: The cumulative distribution function of $X \sim exp(\lambda)$ is (if $a \geq 0$)

$$\begin{aligned} F(a) &= P(X \leq a) \\ &= \int_0^a \lambda e^{-\lambda x} dx. \\ &= \left[-e^{-\lambda x}\right]_0^a \\ &= 1 - e^{-\lambda a} \end{aligned}$$

Note that (1) $F(\infty) = 1$. (2) $P(X > a) =$

Examples

Example 1: The random variable X is uniformly distributed over the interval $[0, 10]$. Calculate the probability that (a) $X < 3$, (b) $X > 6$ and (c) $3 < X < 8$.

Example 2: Buses arrive at a specified stop at 15-minute intervals starting at 7 AM. That is, they arrive at 7, 7:15, 7:30, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits (a) less than 5 five minutes for a bus; (b) more than 10 minutes for a bus.

Solution: Let X denote the amount of time (in minutes) required by the passenger to reach the bus stop between 7.00 and 7.30. $X \sim U(0, 30)$. Then the pdf of X is

$$f(x) = \begin{cases} \frac{1}{30}, & 0 \leq x \leq 30 \\ 0, & \text{otherwise} \end{cases}$$

(a) $P(10 < X < 15) + P(25 < X < 30) = \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx = \frac{1}{3}$.

Standard Continuous Distributions: Uniform Random Variables

Definition

A random variable X is said to be a uniform random variable on the interval $[a.b]$ if its probability density function is given by

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

The cdf of $X \sim U(a, b)$

$$F(x) = \begin{cases} 0, & \mathrm{x} < \mathrm{a} \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1 & \mathrm{x} > \mathrm{b} \end{cases}$$

Question: Given the pmf p_X of a random variable X, what is the pmf of p_Y of $Y = g(X)$?

Answer: Suppose that X assumes the values $x_1, x_2, \ldots$ with positive probabilities and that $Y = g(X)$ assumes the values $y_1, y_2, \ldots$. Then

$$p_Y(y_j) = P(Y = y_j) = P(g(X) = y_j) = P(\{x_i \mid g(x_i) = y_j\}) = \sum_{x_i : g(x_i) = y_j} P(X = x_i) = \sum_{x_i : g(x_i) = y_j} p_X(x_i).$$

Gamma Random Variables

Definition

A random variable X is called a gamma random variable with parameters (α, λ), where $\lambda > 0$ and $\alpha > 0$, if its density function is given by

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The cdf of the Standard Normal and The Normal Table

$\Phi(z)$

The cumulative distribution function of the standard normal random variable Z is usually denoted by $\Phi(z)$:

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx.$$

$$\Phi(-z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-z} e^{-x^2/2} dx.$$

Let $y = -x$. $x = -z \Longrightarrow y = z$ and $x = -\infty \Longrightarrow y = \infty$. $dy = -dx$

$$\begin{aligned} \Phi(-z) &= \tfrac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{x=-z} e^{-x^2/2} dx \\ &= \tfrac{1}{\sqrt{2\pi}} \int_{y=\infty}^{y=z} e^{-y^2/2} (-dy) \end{aligned}$$

The Standard Normal Random Variable, $Z \sim N(0,1)$

Definition

If Z is a normal random variable with parameters $(0,1)$, it is called the standard normal or unit normal random variable.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty.$$

Observations

1. If X is normal with parameters (μ, σ^2), then $Z = \frac{X-\mu}{\sigma}$ is a normal random variable with parameters (0,1) and so it is the standard normal random variable.
2. Conversely, if Z is the standard normal random variable, then $X = \sigma Z + \mu$ is a normal random variable with parameters (μ, σ^2).

Lemma

If X is a normal random variable with parameters (μ, σ^2), then $Y = aX + b$ is a normal random variable with parameters $(a\mu + b, a^2\sigma^2)$.

Proof: Let F_X and F_Y be the cumulative distribution functions of X and Y, respectively. Assume that $a > 0$ (Proof is similar for $a < 0$). Then

$$\begin{aligned} F_Y(x) &= P(Y \leq x) \\ &= P(aX + b \leq x) \\ &= P(X \leq \frac{x-b}{a}) \end{aligned}$$

Normal Random Variable, $X \sim N(\mu, \sigma^2)$

Definition

A continuous random variable X is called a normal random variable with parameters (μ, σ^2) if its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$

Example

Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. Find the probability that (a) it lasts more than 10 minutes; (b) it gets over between 10 and 20 minutes.

Solution: Let X denote the length of the phone call. $X \sim exp(\frac{1}{10})$. The cdf of X is

$$F(a) = \begin{cases} 0, & a < 0 \\ 1 - e^{-\frac{a}{10}}, & a \geq 0 \end{cases}$$

Then the required probabilities are

(a) $P(X > 10) = 1 - F(10) = 1 - (1 - e^{-1}) = e^{-1} \approx 0.368.$

(b) $P(10 < X < 20) =$

Exponential Random Variable, $X \sim exp(\lambda)$

Definition

A continuous random variable whose probability density function (pdf) is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is called an exponential random variable with parameter λ.

Note: The cumulative distribution function of $X \sim exp(\lambda)$ is (if $a > 0$)

$$\begin{aligned} F(a) &= P(X \leq a) \\ &= \int_0^a \lambda e^{-\lambda x} dx, \\ &= \left[-e^{-\lambda x}\right]_0^a \\ &= 1 - e^{-\lambda a} \end{aligned}$$

Note that (1) $F(\infty) = 1$. (2) $P(X > a) = 1 - P(X \leq a) = 1 - F(a) = e^{-\lambda a}$

Standard Continuous Distributions: Uniform Random Variables

Definition

A random variable X is said to be a uniform random variable on the interval $[a, b]$ if its probability density function is given by

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

For any c, d such that $a \le c < d \le b$,

$$P(c \le X \le d) = \int_c^d f(x)dx = \int_c^d \frac{1}{b-a}dx = \frac{d-c}{b-a}.$$

The cdf of $X \sim U(a, b)$

$$F(x) = \begin{cases} 0, & \mathrm{x < a} \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1 & \mathrm{x > b} \end{cases}$$

Example

Consider a random chord of a circle of radius r. What is the probability that the length of the chord will be greater than the side of an equilateral triangle inscribed in that circle?

Solution: Formulation I. Consider a chord of the circle obtained by choosing the distance D of the chord from the center uniformly at random from 0 to r. Then the required probability is

$$P(D < \frac{r}{2}) = \frac{r/2}{r} = \frac{1}{2}.$$

Formulation II. Consider a chord of the circle obtained by fixing one end A and choosing the other end B randomly:

The Memoryless Property of Exponetial Random Variable

Definition

A non-negative random variable X is said to be memoryless if

$$P(X > s + t | X > t) = P(X > s) \quad \text{for all } s, t \geq 0.$$

Note:

$$P(X > s + t | X > t) = \frac{P((X > s + t) \cap (X > t))}{P(X > t)} = \frac{P(X > s + t)}{P(X > t)}.$$

The cdf of the Standard Normal and The Normal Table

$\Phi(z)$

The cumulative distribution function of the standard normal random variable Z is usually denoted by $\Phi(z)$:

Note

$$\Phi(-z) = 1 - \Phi(z), \quad -\infty < z < \infty.$$

Note: The "Normal Table" usually gives the values of $\Phi(z)$ for $z \geq 0$. For negative z, the above identity can be used.
Page number 205 (Standard Normal Curve table)

The DeMoivre-Laplace Theorem

Theorem

Let $0 \leq p \leq 1$ be fixed. Let X_n be a binomial random variable with parameters (n, p). Then for any $a < b$,

$$P\left(a \leq \frac{X_n - np}{\sqrt{np(p-1)}} \leq b\right) \to \Phi(b) - \Phi(a)$$

as $n \to \infty$.

Note: This is a special case of the Central Limit Theorem (proof later)

Gamma Random Variables

Definition

A random variable X is called a gamma random variable with parameters (α, λ), where $\lambda > 0$ and $\alpha > 0$, if its density function is given by

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ \end{cases}$$

Functions of Random Variables

Note: If X is a random variable, then $Y = g(X)$ is also a random variable for any (real-valued) function g of X.

Example: Let X be a discrete random variable assuming one of the three values $-1, 0$ and 1 with probability $\frac{1}{3}$. Then $Y = X^2$ is a random variable assuming the values 0 and 1 .

$$P(Y = 0) = P(X = 0) = \frac{1}{3}$$

$$P(Y = 1) = P(X^2 = 1) = P(X \in \{-1, 1\})$$

Functions of Random Variables: Continuous Case

Problem: Let X be the uniform random variable on the interval $[0, 1]$. Find the density function (pdf) of the random variable $Y = X^2$.

Functions of Random Variables: Continuous Case

Problem: Let X be the uniform random variable on the interval $[0,1]$. Find the density function (pdf) of the random variable $Y = X^2$.

Solution: We first find the cumulative distribution function F_Y (cdf) of Y. Then we differentiate it to obtain the density function (pdf) of Y. The cdf of X is

$$F_x(x) = \begin{cases} 0, & x \leq 0 \\ \text{x}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

The range of $Y = [0,1]$. For $0 < y < 1$,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(X^2 \leq y) \\ &= P(X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) = \sqrt{y} \end{aligned}$$

Thus the density function of $Y = X^2$ is $f_Y(y) = \dfrac{1}{2\sqrt{y}}$ for $0 < y < 1$ and 0 otherwise.

Joint Distributions

- Jointly distributed random variables
- Marginal distributions
- Independence of random variables
- Conditional Distributions

Motivation

- Often we are interested in more than one parameter of an outcome: In the experiment of selecting a particular stock randomly, we may be interested in current price, starting price and appreciation over the last few years.
- The joint distribution (joint pmf, joint pdf or joint cdf) of two (or more) random variables contains information about the individual random variables.

Joint Distributions: Discrete Case

Joint Probability Mass Function (jpmf)

Let X and Y be two discrete random variables. Then their joint probability mass function $p(x, y)$ is given by

$$p(x, y) = P(X = x, Y = y).$$

Example

Three notes are drawn randomly from a bag containing three 10 rupee notes, four 20 rupee notes and five 50 rupee notes. Let X and Y denote the number of 10 and 20 rupee notes drawn, respectively. Then find the joint probability mass function of X and Y, $\quad p(i,j) = P(X = i, Y = j)$.

Example

Three notes are drawn randomly from a bag containing three 10 rupee notes, four 20 rupee notes and five 50 rupee notes. Let X and Y denote the number of 10 and 20 rupee notes drawn, respectively. Then find the joint probability mass function of X and Y, $\quad p(i,j) = P(X = i, Y = j)$.

Note that

$$p(0,0) = P(X = 0, Y = 0) = \frac{\binom{5}{3}}{\binom{12}{3}} = \frac{10}{220},$$

$$p(0,1) = P(X = 0, Y = 1) = \frac{\binom{4}{1}\binom{5}{2}}{\binom{12}{3}} = \frac{40}{220}$$

etc.

The joint probability mass function $p(i,j) = P(X = i, Y = j)$

$i \setminus j$	0	1	2	3	$P(X = i)$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
$P(Y = j)$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	1

contd.

$$p_Y(j) = P(0,j) + P(1,j) + P(2,j) + P(3,j)$$

$$p_Y(0) = P(0,0) + P(1,0) + P(2,0) + P(3,0) = \frac{56}{220}$$

j	0	1	2	3	Total
$p_Y(j)$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	1

Marginal Probability Mass Functions

Let $p(x, y)$ be the joint probability mass function of X and Y. Then the probability mass function of X is

$$p_X(x) = P(X = x) = \sum_{y:p(x,y)>0} p(x, y)$$

and the probability mass function of Y is

$$p_Y(y) = P(Y = y) = \sum_{x:p(x,y)>0} p(x, y).$$

Joint Distributions: Continuous Case

Joint Probability Density Function

Two random variables X and Y are said to be jointly continuous if there exists a function $f(x, y)$, defined for all x and y, such that for every set C of ordered pairs of real numbers

$$P((X, Y) \in C) = \iint_C f(x, y) dxdy.$$

Note

Let X and Y be jointly continuous random variables and let A and B any sets of real numbers. Let $C = \{(x, y) \mid x \in A \text{ and } y \in B\} = A \times B$. Then

$$P((X, Y) \in C) = \int_B \int_A f(x, y)dxdy.$$

Joint Distributions: Continuous Case

Joint Probability Density Function

Two random variables X and Y are said to be jointly continuous if there exists a function $f(x, y)$, defined for all x and y, such that for every set C of ordered pairs of real numbers

$$P((X, Y) \in C) = \iint_C f(x, y) dx dy.$$

The function $f(x, y)$ is called the joint probability density function of the random variables X and Y.

Note

Let X and Y be jointly continuous random variables and let A and B any sets of real numbers.

Marginal Proability Density Functions

Theorem

If X and Y are jointly continuous, then they are individually continuous.

Proof.

Let A be any set of real numbers. Then

$$\begin{aligned} P(X \in A) &= P(X \in A, Y \in (-\infty, \infty)) \\ &= \int_A \int_{-\infty}^{\infty} f(x, y) dy dx \\ &= \int_A f_X(x) dx \end{aligned}$$

Marginal Probability Density Functions

Let X and Y be jointly continuous random variables with joint probability density function $f(x, y)$. Then the probability density function of X is given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

and the probability density function of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute (a) $P(X > 1, Y < 1)$, (b) $P(X < Y)$ and (c) $P(X < a)$.

Solution

(a)

$$
\begin{aligned}
P(X > 1, Y < 1) &= \int_{y=0}^{1} \int_{x=1}^{\infty} 2e^{-x} e^{-2y} \, dx dy \\
&= \int_{y=0}^{1} 2e^{-2y} [-e^{-x}]_1^{\infty} dy \\
&= e^{-1} \int_0^1 2e^{-2y} dy \\
&= e^{-1}(1 - e^{-2})
\end{aligned}
$$

Solution...

(b)

$$
\begin{aligned}
P(X < Y) &= \iint_{x<y} 2e^{-x}e^{-2y}\,dxdy \\
&= \int_{y=0}^{\infty}\int_{x=0}^{y} 2e^{-x}e^{-2y}\,dxdy \\
&= \frac{1}{3}
\end{aligned}
$$

Solution...

(c)

$$P(X < a) = \int_{x=0}^{a} \int_{y=0}^{\infty} 2e^{-x} e^{-2y} \, dy dx$$

Joint Cumulative Distribution Functions

Definition

Let X and Y be any two random variables. Then their joint cumulative distribution function $F(a, b)$ is given by

$$F(a, b) = P(X \leq a, Y \leq b), \quad -\infty < a, b < \infty.$$

Marginal Cumulative Distribution Functions

Definition

Let $F(a,b)$ be the joint distribution function of X and Y. Then the cumulative distribution function $F_X(a)$ of X is

$$\begin{aligned}
F_X(a) &= P(X \le a) \\
&= P(X \le a, Y < \infty) \\
&= P(\lim_{b\to\infty} \{X \le a, Y \le b\}) \\
&= \lim_{b\to\infty} P(X \le a, Y \le b) \\
&= \lim_{b\to\infty} F(a,b) \\
&= F(a,\infty)
\end{aligned}$$

Marginal Cumulative Distribution Functions

Definition

Similarly, the cumulative distribution function $F_Y(b)$ of Y is

$$\begin{aligned} F_Y(b) &= P(Y \leq b) \\ &= \lim_{a \to \infty} F(a, b) \\ &= F(\infty, b) \end{aligned}$$

Problem

Consider a circle of radius R and suppose that a point within the circle is chosen randomly. Let the center of the circle denote the origin and define X and Y to be the coordinates of the point chosen. Then the joint density of X and Y is given by

$$f(x,y) = \begin{cases} c & \text{if } x^2 + y^2 \leq R^2 \\ 0 & \text{if } x^2 + y^2 \geq R^2 \end{cases}$$

for some value of c.

(a) Determine c.

(b) Find the marginal density functions of X and Y.

(c) Compute the probability that D, the distance from the origin of the point selected, is less than or equal to a.

Solution

We know that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx = 1$$

$$\implies \quad c \iint_{x^2+y^2\leq R^2} dydx = 1 \implies c\pi R^2 = 1 \implies c = \frac{1}{\pi R^2}$$

$$f(x,y) = \begin{cases} \frac{1}{\pi R^2} & \text{if } x^2+y^2 \leq R^2 \\ 0 & \text{if } x^2+y^2 \geq R^2 \end{cases}$$

Solution contd.

The marginal probability density function of X is

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\
&= \frac{1}{\pi R^2}\int_{x^2+y^2\leq R^2} dy \\
&= \frac{1}{\pi R^2}\int_{y=-a}^{a} dy \quad \text{where } a=\sqrt{R^2-x^2} \\
&= \tfrac{2a}{\pi R^2} \\
&= \tfrac{2}{\pi R^2}\sqrt{R^2-x^2}, \quad x^2\leq R^2
\end{aligned}
$$

$$
\implies \quad f_X(x)=\begin{cases} \frac{2}{\pi R^2}\sqrt{R^2-x^2}, & x^2\leq R^2 \\ 0\,, & x^2>R^2 \end{cases}
$$

Solution contd.

The marginal probability density function of Y is

$$f_Y(y) = \begin{cases} \frac{2}{\pi R^2}\sqrt{R^2 - y^2}, & y^2 \leq R^2 \\ 0\ , & y^2 > R^2 \end{cases}$$

Note that $D = \sqrt{X^2 + Y^2}$, the distance from the origin. The cdf of D is obtained as follow: For $0 \leq a \leq R$

$$F_D(a) \quad = \quad P\left\{\sqrt{X^2 + Y^2} \leq a\right\}$$

Problem

The joint density of X and Y is given by

$$f(x, y) = \begin{cases} e^{-(x+y)} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the marginal density functions of X and Y respectively. Find the density function of the random variable X/Y.

Solution contd.

The marginal probability density function of Y is

$$f_Y(y) = \begin{cases} \frac{2}{\pi R^2}\sqrt{R^2 - y^2}, & y^2 \leq R^2 \\ 0\ , & y^2 > R^2 \end{cases}$$

Note that $D = \sqrt{X^2 + Y^2}$, the distance from the origin. The cdf of D is obtained as follow: For $0 \leq a \leq R$

$$\begin{aligned} F_D(a) &= P\left\{\sqrt{X^2 + Y^2} \leq a\right\} \\ &= P\left\{X^2 + Y^2 \leq a^2\right\} \end{aligned}$$

Solution contd.

The marginal probability density function of X is

Problem

Consider a circle of radius R and suppose that a point within the circle is chosen randomly. Let the center of the circle denote the origin and define X and Y to be the coordinates of the point chosen. Then the joint density of X and Y is given by

$$f(x,y) = \begin{cases} c & \text{if } x^2 + y^2 \leq R^2 \\ 0 & \text{if } x^2 + y^2 \geq R^2 \end{cases}$$

for some value of c.

(a) Determine c.

(b) Find the marginal density functions of X and Y.

(c) Compute the probability that D, the distance from the origin of the point selected, is less than or equal to a.

Marginal Cumulative Distribution Functions

Definition

Similarly, the cumulative distribution function $F_Y(b)$ of Y is

$$\begin{aligned} F_Y(b) &= P(Y \leq b) \\ &= \lim_{a \to \infty} F(a, b) \\ &= F(\infty, b) \end{aligned}$$

Marginal Cumulative Distribution Functions

Definition

Let $F(a, b)$ be the joint distribution function of X and Y.

Solution...

(b)

$$P(X < Y) = \iint_{x<y} 2e^{-x}e^{-2y}\,dxdy$$

Marginal Probability Density Functions

Let X and Y be jointly continuous random variables with joint probability density function $f(x, y)$. Then the probability density function of X is given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$$

and the probability density function of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx.$$

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ & \end{cases}$$

Joint Cumulative Distribution Functions

Definition

Let X and Y be any two random variables. Then their joint cumulative distribution function $F(a, b)$ is given by

$$F(a, b) = P(X \leq a, Y \leq b), \quad -\infty < a, b < \infty.$$

Marginal Cumulative Distribution Functions

Definition

Let $F(a, b)$ be the joint distribution function of X and Y. Then the cumulative distribution function $F_X(a)$ of X is

$$
\begin{aligned}
F_X(a) &= P(X \leq a) \\
&= P(X \leq a, Y < \infty) \\
&= P(\lim_{b \to \infty} \{X \leq a, Y \leq b\}) \\
&= \lim_{b \to \infty} P(X \leq a, Y \leq b)
\end{aligned}
$$

Solution

We know that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dydx = 1$$

Solution contd.

The marginal probability density function of X is

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\
&= \frac{1}{\pi R^2} \int_{x^2+y^2 \leq R^2} dy \\
&= \frac{1}{\pi R^2} \int_{y=-a}^{a} dy \quad \text{where } a = \sqrt{R^2 - x^2}
\end{aligned}
$$

Solution contd.

The marginal probability density function of Y is

$$f_Y(y) = \begin{cases} \frac{2}{\pi R^2}\sqrt{R^2 - y^2}, & y^2 \leq R^2 \\ 0\ , & y^2 > R^2 \end{cases}$$

Note that $D = \sqrt{X^2 + Y^2}$, the distance from the origin. The cdf of D is obtained as follow: For $0 \leq a \leq R$

$$\begin{aligned} F_D(a) &= P\left\{\sqrt{X^2 + Y^2} \leq a\right\} \\ &= P\left\{X^2 + Y^2 \leq a^2\right\} \end{aligned}$$

Problem

The joint density of X and Y is given by

$$f(x,y) = \begin{cases} e^{-(x+y)} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the marginal density functions of X and Y respectively. Find the density function of the random variable X/Y.

$$f_X(x) = \begin{cases} e^{-x}, & 0 < x < \infty \\ 0, & \text{otherwise} \end{cases} \quad , \quad f_Y(y) = \begin{cases} e^{-y}, & 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Problem

The joint density of X and Y is given by

$$f(x, y) = \begin{cases} e^{-(x+y)} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the marginal density functions of X and Y respectively. Find the density function of the random variable X/Y.

$$f_X(x) = \begin{cases} e^{-x}, & 0 < x < \infty \\ 0, & \text{otherwise} \end{cases}, \quad f_Y(y) = \begin{cases} e^{-y}, & 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Solution:

Let us compute the cumulative distribution function of X/Y. For $a > 0$,

$$
\begin{aligned}
F_{X/Y}(a) &= P(X/Y \leq a) \\
&= \iint_{x/y \leq a} f(x,y)\,dxdy \\
&= \iint_{x \leq ay} e^{-(x+y)}\,dxdy \\
&= \int_{y=0}^{\infty} \int_{x=0}^{ay} e^{-(x+y)}\,dxdy \\
&= \int_{y=0}^{\infty} (1 - e^{-ay}) e^{-y}\,dy = 1 - \frac{1}{a+1}
\end{aligned}
$$

Thus the density function of X/Y is

$$f_{X/Y}(a) = \begin{cases} \frac{1}{(a+1)^2}, & 0 < a < \infty \\ 0, & otherwise \end{cases}$$

Problems

Let $F(a, b)$ be the joint cdf of X and Y.

1. Let $a_1 < a_2$ and $b_1 < b_2$. Prove that

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F(a_2, b_2) + F(a_1, b_1) - F(a_1, b_2) - F(a_2, b_1).$$

2. Let a and b be any real numbers. Prove that

$$P(X > a, Y > b) = 1 - F_X(a) - F_Y(b) + F(a, b).$$

Generalization

Let $X_1, X_2, \ldots, X_n$ be n random variables, where $n \geq 2$.

- If $X_1, X_2, \ldots, X_n$ are discrete, then their joint probability mass function is given by $p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$.
- If $X_1, X_2, \ldots, X_n$ are jointly continuous, then they have a joint probability density function $f(x_1, x_2, \ldots, x_n)$ such that for any set $C \subseteq \mathbb{R}^n$, $P((X_1, X_2, \ldots, X_n) \in C)$ equals the integral of this joint pdf over C.
- Their joint cumulative distribution function is given by $F(a_1, a_2, \ldots, a_n) = P(X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n)$.

Independence of Random Variables

Definition

The random variables X and Y are independent if for any two sets of real numbers A and B

$$P(X \in A, Y \in B) = P(X \in A)\ P(Y \in B).$$

Theorem

The random variables X and Y are independent if and only if for all real numbers a and b

$$P(X \leq a, Y \leq b) = P(X \leq a)\ P(Y \leq b).$$

That is, X and Y are independent if and only if

$$F(a, b) = F_X(a)F_Y(b) \quad \text{for all } a, b.$$

Theorem

- Two discrete random variables X and Y are independent if and only if

$$p(x, y) = p_X(x)p_Y(y) \text{ for all } x, y.$$

- Two jointly continuous random variables X and Y are independent if and only if

$$f(x, y) = f_X(x)f_Y(y) \text{ for all } x, y.$$

Problem

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Prove that X and Y are independent random variables.

Theorem

- Two discrete random variables X and Y are independent if and only if

$$p(x, y) = p_X(x)p_Y(y) \text{ for all } x, y.$$

- Two jointly continuous random variables X and Y are independent if and only if

$$f(x, y) = f_X(x)f_Y(y) \text{ for all } x, y.$$

Problem

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Prove that X and Y are independent random variables.
Solution : Find the mpdf of (i) X and (ii) Y.

$$\begin{aligned} f_X(x) &= \int_{y=0}^{\infty} f(x,y)dy \\ &= \int_0^{\infty} 2e^{-x}e^{-2y}\,dy \\ &= e^{-x}\int_0^{\infty} 2e^{-2y}\,dy = e^{-x} \end{aligned}$$

solution contd.

The mpdf of X is

$$f_X(x) = \begin{cases} e^{-x}, & 0 < x < \infty \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} f_Y(y) &= \int_{x=0}^{\infty} f(x,y)dx \\ &= \int_0^{\infty} 2e^{-x}e^{-2y}dx \\ &= 2e^{-y}\int_0^{\infty} e^{-x}dx = 2e^{-2y} \end{aligned}$$

solution contd.

The mpdf of Y is

$$f_Y(y) = \begin{cases} 2e^{-2y}, & 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

$$f_X(x)f_Y(y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x, y < \infty \\ 0, & \text{otherwise} \end{cases} = f(x, y)$$

Hence X and Y are independent random variables.

Expectations

- Expectations and Variance of Standard Random Variables
- Linearity of Expection
- Independence and Expectation
- Moment Generating Functions

Expectation

A gambler pays Rs. 3 to play a game of dice. He makes Rs. i if the number i turns up. Will gambler make profit over a long period of time?

Expectation

Definition

If X is a discrete random variable with probability mass function $p(x)$, then the expectation or the expected value of X, denoted $E(X)$, is given by

$$E(X) = \sum_{x:p(x)>0} xp(x).$$

That is, the expectation of X is the weighted average of the possible values that X assumes.

Note: The expected value of a random variable is also called the *mean* or the *first moment*.

Example

T he pmf of X is given by

$$p(0) = \frac{1}{2} = p(1)$$

Expectation

Definition

If X is a discrete random variable with probability mass function $p(x)$, then the expectation or the expected value of X, denoted $E(X)$, is given by

$$E(X) = \sum_{x:p(x)>0} xp(x).$$

That is, the expectation of X is the weighted average of the possible values that X assumes.

Note: The expected value of a random variable is also called the *mean* or the *first moment*.

Example

T he pmf of X is given by

$$p(0) = \frac{1}{2} = p(1)$$

$$E(X) = \sum_{x:p(x)>0} xp(x) = 0 \times p(0) + 1 \times p(1)$$

$$E(X) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}.$$

contd.

Example

The pmf of X is given by

$$p(0) = \frac{1}{3} \text{ and } p(1) = \frac{2}{3}$$

$$E(X) = \sum_{x:p(x)>0} xp(x) = 0 \times p(0) + 1 \times p(1)$$

$$E(X) = 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}.$$

The Indicator Random Variable

Let A be an event and let I be a random variable defined as follows:

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

Then I is called the indicator random variable of the event A.
Its expectation is $P(A)$:

$$E(I) = 0 \times P(I = 0) + 1 \times P(I = 1) = 0 \times (1 - P(A)) + 1 \times P(A) = P(A).$$

Example

Let X be the outcome when we roll a fair die. Find $E(X)$.

Example

Let X be the outcome when we roll a fair die. Find $E(X)$.

Solution:

$$p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}.$$

$$E(X) = \sum_{x:p(x)>0} xp(x) = 1 \times p(1) + \ldots + 6 \times p(6)$$

$$E(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2}.$$

Expectation is the Same as the Center of Gravity

- Let X be a random variable with pmf $p(x_i)$, $i \geq 1$.
- Consider a weightless rod in which weights $p(x_i)$ are attached at locations x_i.
- The point about which the rod would be balancing is called the center of gravity of the rod.

Expectation of a Function of a Random Variable

Proposition

If X is a discrete random variable that takes on one of the values x_i, $i \geq 1$, with respective probabilities $p(x_i)$, then for any real-valued function g

$$E[g(X)] = \sum_i g(x_i)p(x_i).$$

Proof:

Suppose that $y_j, j \geq 1$, represent the different values of $g(x_i), i \geq 1$. Then grouping all the $g(x_i)$ having the same value gives

$$
\begin{aligned}
\sum_i g(x_i)p(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \\
&= \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i)
\end{aligned}
$$

Expectation of a Function of a Random Variable

Proposition

If X is a discrete random variable that takes on one of the values x_i, $i \geq 1$, with respective probabilities $p(x_i)$, then for any real-valued function g

$$E[g(X)] = \sum_i g(x_i)p(x_i).$$

Linearity of Expectation I

Corollary

If a and b are constants, then

$$E(aX + b) = aE(X) + b.$$

Proof.

$$\begin{aligned} E(aX + b) &= \sum_{x:p(x)>0} (ax + b)p(x) \\ &= a \sum_{x:p(x)>0} xp(x) + b \sum_{x:p(x)>0} p(x) \\ &= aE(X) + b \end{aligned}$$

The nth Moment $E[X^n]$

Definition

For any random variable X, the quantity $E[X^n], n \geq 1$, is called the nth moment of X.

Note

$$E[X^n] = \sum_{x:p(x)>0} x^n p(x).$$

$$E(1) = E(X^0) = 1$$

Generalizations

Proposition

If X and Y have a joint probability mass function $p(x, y)$, then

$$E[g(X, Y)] = \sum_y \sum_x g(x, y) p(x, y).$$

Linearity of Expectation II

Theorem

If X and Y are any random variables, then

$$E[X+Y] = E[X] + E[Y].$$

Theorem

If $X_1, X_2, \ldots, X_n$ are any random variables, then

$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n].$$

Variance

Definition

If X is a random variable with mean $E(X) = \mu$, then the variance of X, denoted $\mathrm{Var}(X)$, is given by

$$\mathrm{Var}(X) = E\left[(X - E(X))^2\right] = E[(X - \mu)^2]$$

The Standard Deviation

Definition

The standard deviation of a random variable X, denoted $\mathrm{SD}(X)$, is given by

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.$$

Variance: Alternative Formula

$$\begin{aligned}\mathrm{Var}(X) &= E[(X-\mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - 2\mu^2 + \mu^2 \\ &= E[X^2] - \mu^2\end{aligned}$$

$$\mathrm{Var}(X) = E[X^2] - (E(X))^2 .$$

The Standard Deviation

Definition

The standard deviation of a random variable X, denoted $\mathrm{SD}(X)$, is given by

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.$$

Independence and Expectation

If X and Y are independent random variables, then

$$E[XY] = E[X]\ E[Y].$$

More generally, in this case,

$$E[g(X)\ h(Y)] = E[g(X)]\ E[h(Y)].$$

Proof

Let us consider the discrete case. Let X and Y be two independent random variables with jpdf $p(x, y)$

$$
\begin{aligned}
E(XY) &= \sum_x \sum_y xyp(x, y) \\
&= \sum_x \sum_y xy \; p_X(x) \; p_Y(y) \quad (Independent \;\; R.Vs.) \\
&= \sum_x xp_X(x) \sum_y yp_Y(y) \\
&= E(X)E(Y)
\end{aligned}
$$