SAMSUNG

Together for Tomorrow!
Enabling People
Education for Future Generations

Samsung Innovation Campus

Artificial Intelligence Course

Together for Tomorrow!
Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

AI Course

Samsung Innovation Campus

Update History

Version	Update Details
v1.0	• 1st draft completed
v1.01	• 'Guide for Instructors' page inserted • 'Enabling People' logo replaced
v1.02	• Front and Cover page layout updated • 'Coding Exercise' replaced 'Coding Practice' • 'Quiz.' replaced 'Coding Problem'
V1.03	• Samsung, SIC loge updated
V1.04	• Chapter info. updated

Course Description (1/2)

- Duration: 24 Hours.
- Purpose:
  - Introduction to probability and statistics.
  - Introduction to the estimation theory and hypothesis testing.
  - Introduction to the SciPy library.
- Target Audience:
  - This course is for those who would like to gain deeper understanding of the mathematical background of data analysis and modeling.
- Prerequisite:
  - A student taking this course should be able to program in Python.
  - A student taking this course should be familiar with the general aspects of the Pandas library.

Course Description (2/2)

Objectives:

- Make probabilistic assessment.
- Carry out descriptive statistics and exploratory data analysis.
- Apply hypothesis testing.

Together for Tomorrow! Enabling People

Education for Future Generations

Chapter 4.

Probability and Statistics

UNIT 1.

Understanding of Probability

Samsung Innovation Campus

Unit 1.
Understanding of Probability

What this unit is about:

- You will gain understanding of the probability theory.
- You will learn about the random variables.
- You will learn about the discrete probability distributions.

Expected outcome:

- Ability to define and calculate the probability.
- Ability to calculate the expected values.
- Ability to model random events with the appropriate probability distribution.

How to check your progress:

- Coding Exercises.
- Quiz.

Chapter 4.

Probability and Statistics

UNIT 1. Understanding of Probability

1.1. Probability Theory.

1.2. Probability Rules.

1.3. Random Variable.

1.4. Discrete Probability Distribution.

Unit 2. Understanding of Statistics I

2.1. Continuous Probability Density.

2.2. Conjoint Probability.

Unit 3. Understanding of Statistics II

3.1. Descriptive Statistics.

3.2. Central Limit Theorem.

3.3. Estimation Theory.

Unit 4. Statistical Hypothesis Testing

4.1. Principles of Hypothesis Testing.

4.2. Hypothesis Testing in Action.

Probability Theory

Definition 1.1. Outcome: An *outcome* of an experiment is any possible observation of that experiment.

Definition 1.2. Sample Space: The *sample space* of an experiment is the finest-grain, mutually exclusive, collectively exhaustive set of all possible outcomes.

Definition 1.3. Event: An *event* is a set of outcomes of an experiment.

Flip a coin three times. Observe the sequence of heads and tails.

$$S = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$$

Flip a coin three times. Observe the number of heads.

$$S = \{0, 1, 2, 3\}$$

Probability Theory - Contd.

Set Algebra	Probability
set	event
universal set	sample space
element	outcome

If a Six-Sided Dice is tossed: Sample Space ??

- *The event* $E_1 = \{\textit{Roll 4 or higher}\} = \{4, 5, 6\}$.
- *The event* $E_2 = \{\textit{The roll is even}\} = \{2, 4, 6\}$.
- $E_3 = \{\textit{The roll is the square of an integer}\} = \{1, 4\}$.

Probability Theory - Contd.

Definition 1.4. Event Space: An *event space* is a collectively exhaustive, mutually exclusive set of events.

The members of a sample space are *outcomes*. By contrast, the members of an event space are *events*.

The event space is a set of events (sets), while the sample space is a set of outcomes (elements).

Together for Tomorrow! Enabling People
Education for Future Generations

Random Experiment and Event (1/6)

Random experiment:

- Experiment where all possible outcomes can be observed.
- Can be repeated as many times as possible under the exact same conditions.
- The result or outcome is produced randomly.

Random Experiment and Event (2/6)

Sample space:

- Suppose that the possible outcomes of a random experiment are $e_1, e_2, e_3, \dots, e_N$.
- You denote the sample space of the random experiment by S. This can be expressed as a set.

$$S = \{e_1, e_2, e_3, \dots, e_N\}$$

Ex) Rolling a dice, $S = \{1, 2, 3, 4, 5, 6\}$.

Ex) Flipping a coin, $S = \{T, H\}$ where T = tail and H = head.

Random Experiment and Event (3/6)

- An event is a subset of the sample space:

  Ex) The whole of S → "total event"

  Ex) Empty set ϕ → "empty event"

  Ex) Subsets composed of individual elements of S: $\{e_1\}$, $\{e_2\}$, … → "elementary events"

  Ex) Other subsets: $\{e_1, e_2\}$, $\{e_7, e_9, e_{13}\}$, …

Random Experiment and Event (4/6)

Events of rolling a dice:

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

  Ex) Event of outcome three: $E_1 = \{3\}$→ "Elementary event"

  Ex) Event of outcome equal or larger than three: $A = \{3, 4, 5, 6\}$

  Ex) Event of odd outcome: $O = \{1, 3, 5\}$

  Ex) Event of even outcome: $E = \{2, 4, 6\}$

  Ex) Event of even outcome that is equal or larger than three: $A \cap E = \{4, 6\}$

  Ex) Event of even outcome or outcome that is equal or larger than three: $A \cup E = \{2, 3, 4, 5, 6\}$

Random Experiment and Event (5/6)

- Events of rolling a dice (continued):

  Ex) We notice that $E \cap O = \phi$ → "E and O are mutually exclusive events".

  Ex) Event of even outcome that is not equal or larger than three: $E - A = \{2\}$

  Ex) Event of outcome equal or larger than three that is not even: $A - E = \{3, 5\}$

  Ex) Event of outcome that is not equal or larger than three: $A^c = S - A = \{1, 2\}$

  → "Complementary event"

Together for Tomorrow! Enabling People
Education for Future Generations

Random Experiment and Event (6/6)

Events of flipping two coins: one dime (10 cents) and one quarter (25 cents).

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

  Ex) Event of both coins showing the tail side: $E_1 = \{(T,T)\}$

  Ex) Event of both coins showing the same side: $E_2 = \{(H,H),(T,T)\}$

  Ex) Event of at least one coin showing the tail side: $E_3 = \{(T,T),(T,H),(H,T)\}$

Definition of Probability (1/5)

Mathematical definition of probability:

- Suppose that N is the number of elementary events of the sample space.
- Suppose that N_A is the number of all possible outcomes corresponding to the event A.

$$P(A) = \frac{N_A}{N}$$

Definition of Probability (2/5)

Probabilities of rolling two dices:

- Sample space: $S = \{(1,1),(1,2),(1,3),\dots,(6,6)\}$, N=36.

Ex) The probability of dices showing different numbers.

Ex) The probability of dices showing different numbers.

$$P = \frac{36-6}{36} = \frac{5}{6}$$

Ex) The probability of dices showing numbers that sum up to 7.

The event we are interested in is $A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$. So, $N_A = 6$.

$$P = \frac{6}{36} = \frac{1}{6}$$

Together for Tomorrow! Enabling People
Education for Future Generations

Definition of Probability (3/5)

Probabilities of rolling two dices:

- Sample space: $S = \{(1,1),(1,2),(1,3),\dots,(6,6)\}$, $N = 36.$

Ex) The probability of dices showing numbers the product of which is even.

There are 27 combinations where the product can be an even number.

$$\left.\begin{aligned}(\text{Even}) \times (\text{Even}) = (\text{Even}) &\Rightarrow 3 \times 3 = 9 \text{ cases.}\\ (\text{Even}) \times (\text{Odd}) = (\text{Even}) &\Rightarrow 3 \times 3 = 9 \text{ cases.}\\ (\text{Odd}) \times (\text{Even}) = (\text{Even}) &\Rightarrow 3 \times 3 = 9 \text{ cases.}\end{aligned}\right\} \text{27 in total.}$$

$$P = \frac{27}{36} = \frac{3}{4}$$

Definition of Probability (4/5)

Statistical (empirical) definition of probability:

- Total number of observations of a random experiment is N.
- The number of observations corresponding to the event A is N_A.
- Ideally, the probability is given by the following limit:

$$P(A) = \lim_{N \to \infty} \frac{N_A}{N}$$

- However, the following is more realistic:

$$P(A) = \frac{N_A}{N}$$

Definition of Probability (5/5)

Statistical (empirical) definition of probability:

Ex) According to the data, out of 1,000 newborns 985 live beyond the first year.

What is the survival probability at one year?

$$P = \frac{985}{1000} = 0.985 \approx 0.98$$

Ex) A dice was rolled 10,000 times out of which 1,650 times the outcome ‘one’ was obtained.

What is the probability of the outcome one?

$$P = \frac{1650}{10000} = 0.165 \approx \frac{1}{6}$$

Theorems of Probability

- For an arbitrary event A, sample space S and empty event ϕ:

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

$$P(S) = 1$$

$$P(\phi) = 0$$

- Given the elementary events e_i and the corresponding probabilities p_i, the “normalization” holds:

$$p_1 + p_2 + \cdots + p_N = 1$$

Probability Axioms

Axioms of Probability: A probability measure $P[\cdot]$ is a function that maps events in the sample space to real numbers such that

Axiom 1. For any event A, $P[A] \geq 0$.

Axiom 2. $P[S] = 1$.

Axiom 3. For any countable collection $A_1, A_2, \ldots$ of mutually exclusive events

$$P[A_1 \cup A_2 \cup \cdots] = P[A_1] + P[A_2] + \cdots$$

Probability Notations

P[A] = Probability of Event A

P[s] = Probability of Outcome s

$$P[A \cap B] = P[A, B] = P[AB]$$

Probability Theorems

Theorem 1.3. *Given events A and B such that $A \cap B = \phi$, then*

$$P[A \cup B] = P[A] + P[B]$$

Theorem 1.4. *If $B = B_1 \cup B_2 \cup \cdots \cup B_m$ and $B_i \cap B_j = \phi$ for $i \neq j$, then*

$$P[B] = \sum_{i=1}^{m} P[B_i]$$

Theorem 1.5. *The probability of an event $B = \{s_1, s_2, \ldots, s_m\}$ is the sum of the probabilities of the outcomes contained in the event:*

$$P[B] = \sum_{i=1}^{m} P[\{s_i\}]$$

Theorem 1.6. *For an experiment with sample space $S = \{s_1, \ldots, s_n\}$ in which each outcome s_i is equally likely,*

$$P[s_i] = 1/n \qquad 1 \leq i \leq n$$

Together for Tomorrow! Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

UNIT 1. Understanding of Probability

1.1. Probability Theory.

1.2. Probability Rules.

1.3. Random Variable.

1.4. Discrete Probability Distribution.

Unit 2. Understanding of Statistics I

2.1. Continuous Probability Density.

2.2. Conjoint Probability.

Unit 3. Understanding of Statistics II

3.1. Descriptive Statistics.

3.2. Central Limit Theorem.

3.3. Estimation Theory.

Unit 4. Statistical Hypothesis Testing

4.1. Principles of Hypothesis Testing.

4.2. Hypothesis Testing in Action.

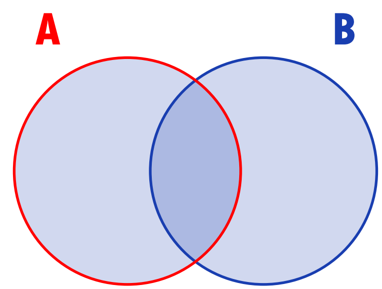
Probability Rules (1/12)

Sum of probabilities:

- For arbitrary events A and B:

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

"**Sum rule**"



Together for Tomorrow! Enabling People
Education for Future Generations

Probability Rules (2/12)

Sum of probabilities:

Ex) When we roll a dice, what is the probability that the outcome is odd or larger than three?

If the event $A =$ 'odd outcome' and $B =$ 'outcome larger than three',

$A = \{1, 3, 4\}$

$B = \{4, 5, 6\}$

a) The union of A and B is the event we are interested in.

So, $A \cup B = \{1, 3, 4, 5, 6\} \rightarrow P(A \cup B) = \frac{5}{6}$

Together for Tomorrow! Enabling People
Education for Future Generations

Probability Rules (3/12)

Sum of probabilities:

Ex) When we roll a dice, what is the probability that the outcome is odd or larger than three?

If the event A = 'odd outcome' and B = 'outcome larger than three',

$A = \{1, 3, 4\}$

$B = \{4, 5, 6\}$

b) We have $P(A) = \frac{3}{6}$, $P(B) = \frac{3}{6}$, $P(A \cap B) = P(\{4\}) = \frac{1}{6}$.

Using the sum rule, $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{6} + \frac{3}{6} - \frac{1}{6} = \frac{5}{6}$.

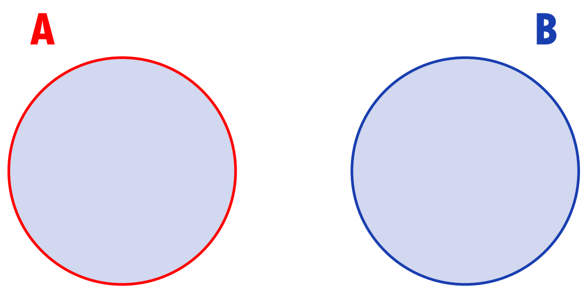
Together for Tomorrow! Enabling People
Education for Future Generations

Probability Rules (4/12)

Sum of probabilities:

- If the events A and C are mutually exclusive ($A \cap B = \phi$):

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



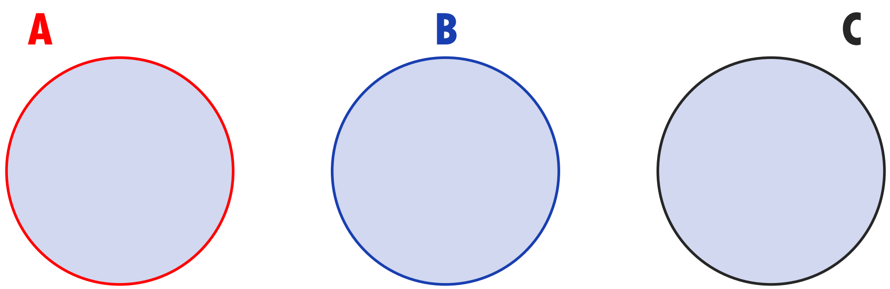
Together for Tomorrow! Enabling People
Education for Future Generations

Probability Rules (5/12)

Sum of probabilities:

- If the events A, B and C are mutually exclusive ($A \cap B = \phi$", " $A \cap C = \phi$", " $B \cap C = \phi$):

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



Probability Rules (6/12)

Sum of probabilities:

Ex) Out of 100 people, 30 are of blood type A, 28 of B, 17 of O and 25 of AB.

If we randomly pick one person, what is the probability of that person's blood type being A or AB?

a) Blood types are mutually exclusive events

So, $P(A \cup AB) = P(A) + P(AB) = \frac{30}{100} + \frac{25}{100} = \frac{11}{20}$.

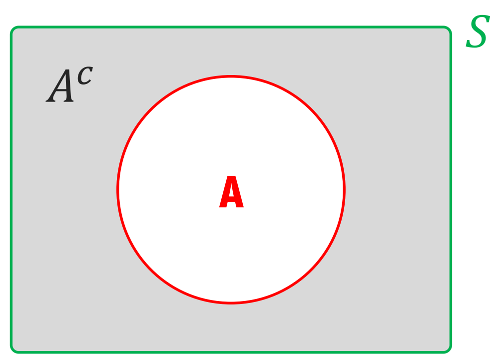
Together for Tomorrow! Enabling People
Education for Future Generations

Probability Rules (7/12)

Complementary events:

- For an event A and its complementary event A^c , the following relation holds:

$$A \cup A^c = S \;\Rightarrow\; P(A) + P(A^c) = 1$$
$$\Rightarrow\; P(A) = 1 - P(A^c)$$



Probability Rules (8/12)

Complementary events:

Ex) When we flip five coins, what is the probability that at least one comes out head?

a) If the event A= 'at least one coin comes out head', its complementary is A^c= 'all the coins come out tail'. Thus,

$$P(A^c) = \frac{1}{2^5} = \frac{1}{32}$$

b) Between the event and its complementary, we have $P(A)$+$P(A$^c)=1. So,

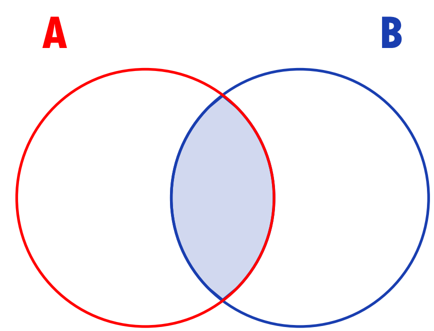
$$P(A) = 1 - P(A^c) = 1 - \frac{1}{32} = \frac{31}{32}$$

Probability Rules (9/12)

Conditional probability:

- Suppose that you have the events A and B of which probabilities are non-zero.
- The probability of the B conditional on the A is denoted by $P(B|A)$.
- The following relation holds among the probabilities.

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



- We can also rearrange the above relation as following.

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

Together for Tomorrow! Enabling People
Education for Future Generations

Probability Rules (10/12)

Conditional probability:

Ex) In a sack, there are two white balls and three red balls.

Suppose that two balls are taken out one after another.

Calculate the probability of both balls being white.

a) If the first ball is not put back into the sack.

Let us denote by A = event that the first ball is white, and

B = event that the second ball is white.

$$P(A \cap B) = P(B|A)P(A) = \frac{1}{4} \times \frac{2}{5} = \frac{1}{10}$$

Probability Rules (11/12)

Conditional probability:

Ex) In a sack, there are two white balls and three red balls.

Suppose that two balls are taken out one after another.

Calculate the probability of both balls being white.

b) If the first ball is put back into the sack before the second ball is withdrawn.

This is like a "reset". The events A and B are independent from each other.

$$P(A \cap B) = P(B|A)P(A) = P(B)P(A) = \frac{2}{5} \times \frac{2}{5} = \frac{4}{25}$$

Probability Rules (12/12)

For the given events A and B, you can think of the following cases.

1) If the events A and B are dependent on each other, then:

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

and

$$P(A \cap B) \neq P(A)P(B)$$

2) If the events A and B are independent from each other, then:

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

3) If the events A and B are mutually exclusive, then:

$$P(A \cap B) = 0$$

Bayes' Theorem (1/9)

- Let us remember the following relation:

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

- Using the above relation, we can state the so-called Bayes' theorem:

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

Law of Total Probability

For an event space $\{B_1, B_2, \ldots, B_m\}$ *with* $P[B_i] > 0$ *for all* i,

$$P[A] = \sum_{i=1}^{m} P[A|B_i]\, P[B_i].$$

Together for Tomorrow! Enabling People
Education for Future Generations

Bayes' Theorem (2/9)

- We may transform the Bayes' theorem in the following way:

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

$$\downarrow$$

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

$$= \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^c)}$$

Together for Tomorrow! Enabling People
Education for Future Generations

Bayes' Theorem (3/9)

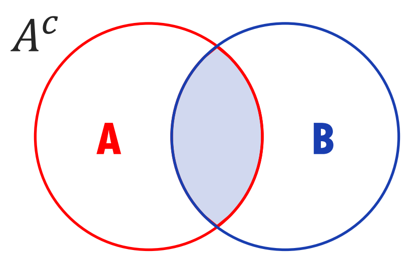
- We may transform the Bayes' theorem in the following way:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

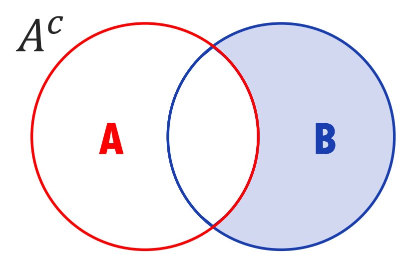
Bayes' Theorem (4/9)

- We may transform the Bayes' theorem in the following way:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$



$P(B \cap A)$



$P(B \cap A^c)$

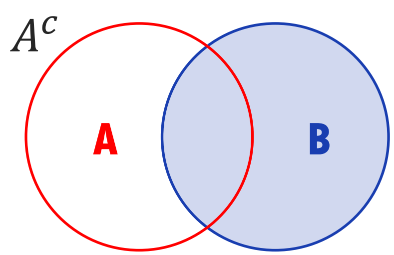
Together for Tomorrow! Enabling People
Education for Future Generations

Bayes' Theorem (5/9)

- We may transform the Bayes' theorem in the following way:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

$$\parallel$$



$P(B)$

Together for Tomorrow! Enabling People
Education for Future Generations

Bayes' Theorem (6/9)

- Bayes' theorem example:

  Ex) There are 100 coins out of which one is abnormal with both sides as heads (H).

  The remaining 99 coins are normal with head (H) on one side and tail (T) on the other.

  A coin is randomly drawn and then thrown five times in a row.

  In all five times, the coin lands showing the head side only.

  What is the probability that this coin is that abnormal one?

  a) Let us define as A the event that the coin is abnormal.

  Then A^c is the event that the coin is normal.

  Let us also define B the event that the coin lands five times showing the head side.

Together for Tomorrow! Enabling People
Education for Future Generations

Bayes' Theorem (7/9)

- Bayes' theorem example:

Ex) There are 100 coins out of which one is abnormal with both sides as heads (H).

The remaining 99 coins are normal with head (H) on one side and tail (T) on the other.

A coin is randomly drawn and then thrown five times in a row.

In all five times, the coin lands showing the head side only.

What is the probability that this coin is that abnormal one?

b) The answer we seek is given by $P(A|B)$:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{1 \times \frac{1}{100}}{1 \times \frac{1}{100} + \left(\frac{1}{2}\right)^5 \times \frac{99}{100}} \cong \mathbf{0.244}$$

Together for Tomorrow! Enabling People
Education for Future Generations

Bayes' Theorem (8/9)

Bayes' theorem example:

Ex) A particular type of cancer affects 0.1% of the population.

A new diagnostic method has been devised.

→ This method can diagnose as positive (+) in 99% of the true cases (D).

→ This method can diagnose as negative (-) in 95% of the false cases (D^c).

This diagnostic method was administered to a patient giving the positive (+) result.

What is the probability that this patient actually has the cancer?

a) We have $P(D) = 0.001$, $P(D^c) = 0.999$, $P(+|D) = 0.99$ and $P(-|D^c) = 0.95$.

We can also derive that $P(+|D^c) = 1 - P(-|D^c) = 0.05$.

Bayes' Theorem (9/9)

- Bayes' theorem example:

Ex) A particular type of cancer affects 0.1% of the population.

A new diagnostic method has been devised.

→ This method can diagnose as positive (+) in 99% of the true cases (D).

→ This method can diagnose as negative (-) in 95% of the false cases (D^c).

This diagnostic method was administered to a patient giving the positive (+) result.

What is the probability that this patient actually has the cancer?

b) The answer we seek is given by $P(D|+)$:

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|D^c)P(D^c)} = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.05 \times 0.999} \cong \mathbf{0.02}$$

Chapter 4.

Probability and Statistics

Together for Tomorrow! Enabling People
Education for Future Generations

UNIT 1. Understanding of Probability

1.1. Probability Theory.
1.2. Probability Rules.
1.3. Random Variable.
1.4. Discrete Probability Distribution.

Unit 2. Understanding of Statistics I

2.1. Continuous Probability Density.
2.2. Conjoint Probability.

Unit 3. Understanding of Statistics II

3.1. Descriptive Statistics.
3.2. Central Limit Theorem.
3.3. Estimation Theory.

Unit 4. Statistical Hypothesis Testing

4.1. Principles of Hypothesis Testing.
4.2. Hypothesis Testing in Action.

Random Variable (1/3)

- Random variable: a function that assigns a number to the outcome of a random experiment.

  Ex) In a random experiment of coin flipping, assign 1 to the head (H) and assign 0 to the tail (H).

- There are two types of random variables:

  1) Discrete: for random experiments with a finite number of the possible outcomes.

     Ex) Random variable that represents coin flipping experiment. Possible outcomes $= \{H, T\}$.

     Ex) Random variable that represents dice rolling experiment. Possible outcomes $= \{1,2,3,4,5,6\}$.

  2) Continuous: for random experiments with an infinite number of the possible outcomes.

     Ex) Random variable that represents the heights of people.

     Ex) Random variable that represents the wages of people.

Random Variable (2/3)

Discrete probability distribution function $P(x)$:

- Maps the values of a discrete random variable to the corresponding probabilities.
- You will denote in upper case a random variable and in lower case a particular value of it.
  Ex) Given a random variable X,

  the probability of X taking on a value x is denoted by $P(X=x)$ or $P(x)$.

Continuous probability density function $f(x)$:

- When integrated, gives the interval probabilities $\Leftarrow$ More about this in the next Unit.

Random Variable (3/3)

- Properties of discrete probability distribution:

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

2) $\sum_{all\ x_i} P(x_i) = 1$

Population and Sample (1/2)

Population:

- The entirety of the data set subject to the analysis.
- Can be either "real" or "idealized".
- The properties of a population are called parameters:

    Ex) Mean, standard deviation, variance, etc. of a population.

Sample:

- A subset of the population.
- The properties of a sample are called statistics:

    Ex) Mean, standard deviation, variance, etc. of a sample.

Population and Sample (2/2)

- We could calculate the mean, standard deviation, variance, etc. using a probability distribution function.
  - These quantities are properties of an "idealized" group.
  - They can be interpreted as parameters of an idealized population.

Population Mean (1/2)

- Population mean is denoted by μ.

  - It is also called the expected value of a random variable: $E[X]$.

    1) For the discrete case with $P(x)$= probability distribution function:

$$\mu = E[X] = \sum_{all\ x} x\, P(x)$$

    2) For the continues case with $f(x)$= probability density function:

$$\mu = E[X] = \int x\, f(x)dx$$

Population Mean (2/2)

Properties of the population mean:

1) $E[c] = c$

2) $E[c\,X] = c\,E[X]$

3) $E[X + c] = E[X] + c$

4) $E[X + Y] = E[X] + E[Y]$

Note: c stands for a constant.

Population Variance (1/2)

- The population variance is denoted by σ^2 or $Var(X)$.

1) For the discrete case with $P(x)$= probability distribution function:

$$\sigma^2 = Var(X) = \sum_{all\ x} (x - \mu)^2 \, P(x)$$

2) For the continues case with $f(x)$= probability density function:

$$\sigma^2 = Var(X) = \int (x - \mu)^2 f(x) dx$$

Note: In both cases $\sigma^2 = E[X^2] - (E[X])^2$

Note: The population standard deviation $\sigma = \sqrt{\sigma^2}$

Together for Tomorrow! Enabling People
Education for Future Generations

Population Variance (2/2)

Properties of the population variance:

1) $Var(c) = 0$

2) $Var(X + c) = Var(X)$

3) $Var(c\, X) = c^{2}\, Var(X)$

4) $Var(X + Y) = Var(X) + Var(Y) + 2\, Cov(X, Y)$

$= Var(X) + Var(Y)$ $\Leftarrow$ Only when X and Y are independent to each other!

Note: c stands for a constant.

Chapter 4.

Probability and Statistics

Together for Tomorrow! Enabling People

Education for Future Generations

UNIT 1. Understanding of Probability

1.1. Probability Theory.

1.2. Probability Rules.

1.3. Random Variable.

1.4. Discrete Probability Distribution.

Unit 2. Understanding of Statistics I

2.1. Continuous Probability Density.

2.2. Conjoint Probability.

Unit 3. Understanding of Statistics II

3.1. Descriptive Statistics.

3.2. Central Limit Theorem.

3.3. Estimation Theory.

Unit 4. Statistical Hypothesis Testing

4.1. Principles of Hypothesis Testing.

4.2. Hypothesis Testing in Action.

Discrete Probability Distribution (1/19)

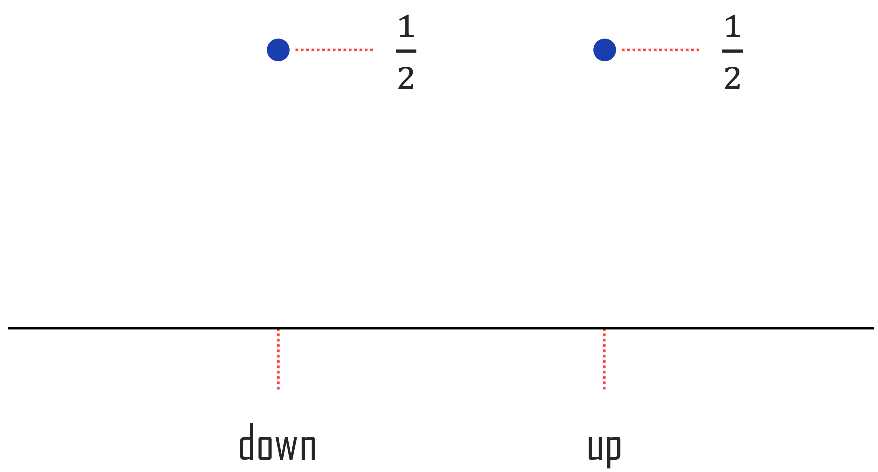
- Bernoulli random variable and probability distribution:

Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (2/19)

Bernoulli random variable and probability distribution:

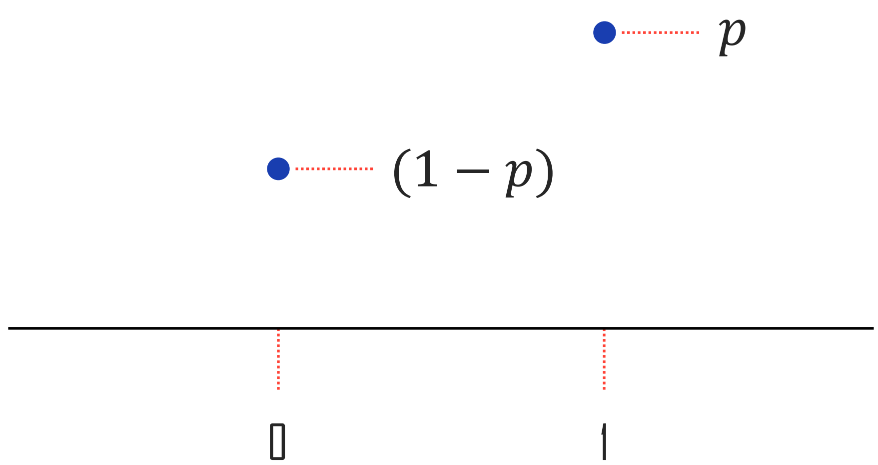
Ex) Coin flipping:



Discrete Probability Distribution (3/19)

Bernoulli random variable and probability distribution:

Ex) General binary outcome situation:



Discrete Probability Distribution (4/19)

Bernoulli random variable and probability distribution:

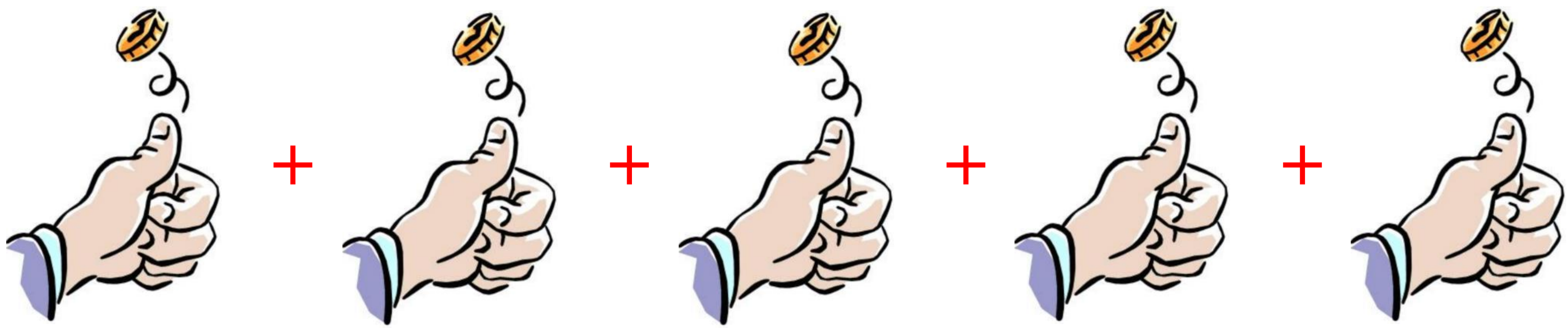
1) "X is a Bernoulli random variable" $\Leftrightarrow$ $X \sim Ber(p)$

2) $P(x) = p^x(1-p)^{1-x}$

- Mean = p
- Variance = $p\,(1-p)$
- Standard deviation = $\sqrt{p\,(1-p)}$

Discrete Probability Distribution (5/19)

- Binomial random variable and probability distribution:



Discrete Probability Distribution (6/19)

Binomial random variable and probability distribution:

1) "X is a Binomial random variable" $\Leftrightarrow X\sim Bin(n,p)$

2) $X_{bin} = X_{Ber} + X_{Ber} + \cdots + X_{Ber}$

$\leftarrow \quad n \quad \rightarrow$

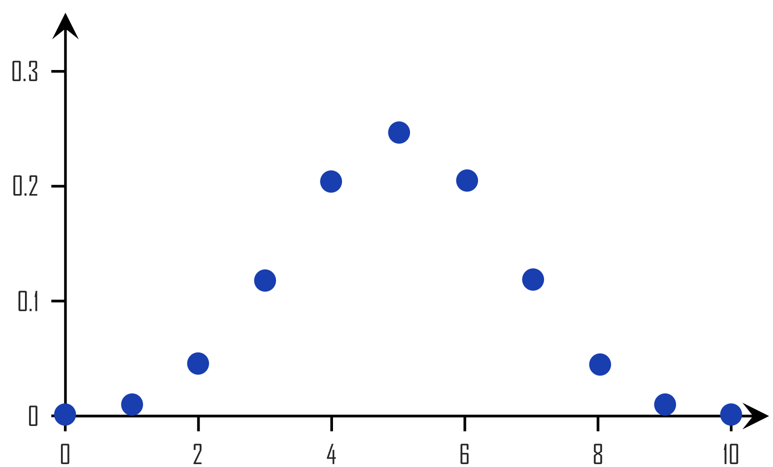
3) $P(x) = \binom{n}{x} p^{x}(1-p)^{n-x}$ $\quad \Leftarrow 0 \le x \le n$

- Mean $= n\,p$
- Variance $= n\,p\,(1-p)$
- Standard deviation = $\sqrt{n\,p\,(1-p)}$

Discrete Probability Distribution (7/19)

- Binomial random variable and probability distribution:

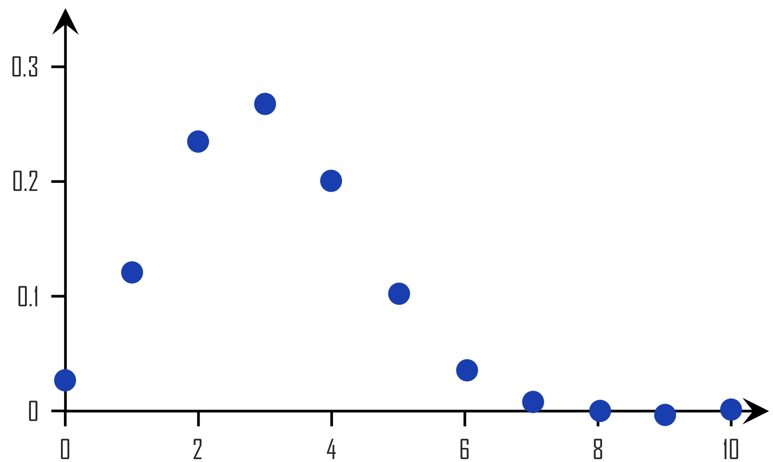
$$n = 10, p = 0.5$$



Discrete Probability Distribution (8/19)

Binomial random variable and probability distribution:

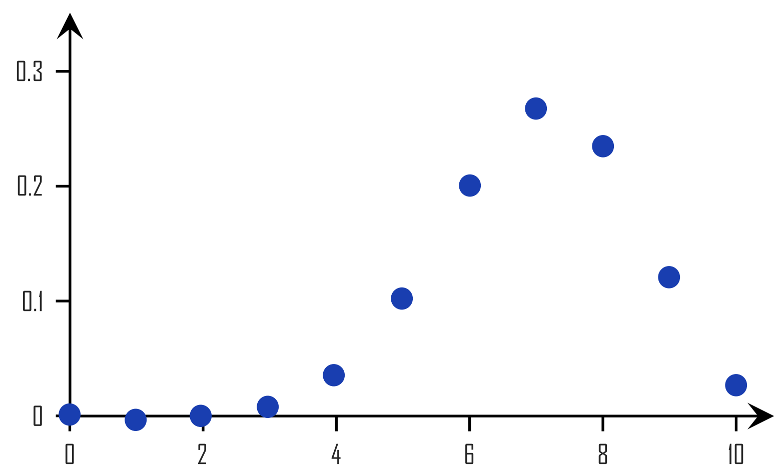
$$n = 10, p = 0.3$$



Discrete Probability Distribution (9/19)

- Binomial random variable and probability distribution:

$$n = 10, p = 0.7$$



Discrete Probability Distribution (10/19)

Binomial random variable and probability distribution:

Ex) It is known that the daily probability of raining is 20% for the following five days.

What is the probability of **zero** rainy day?

$$P(0) = \binom{5}{0} 0.2^0 (1-0.2)^{5-0} = \frac{5!}{0!\ 5!} 0.8^5 = 0.8^5 = 0.328$$

Discrete Probability Distribution (11/19)

Binomial random variable and probability distribution:

Ex) It is known that the daily probability of raining is 20% for the following five days.

What is the probability of **exactly two** rainy days?

$$P(2) = \binom{5}{2} 0.2^2 (1-0.2)^{5-2} = \frac{5!}{2!\,3!} 0.2^2 \times 0.8^3 = 10 \times 0.2^2 \times 0.8^3 = 0.205$$

Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (12/19)

Binomial random variable and probability distribution:

Ex) It is known that the daily probability of raining is 20% for the following five days.

What is the probability of **two or fewer** rainy days?

$$P(X \leq 2) = P(0) + P(1) + P(2)$$

$$= \binom{5}{0} 0.2^0 (1-0.2)^{5-0} + \binom{5}{1} 0.2^1 (1-0.2)^{5-1} + \binom{5}{2} 0.2^2 (1-0.2)^{5-2}$$

$$= 1 \times 1 \times 0.8^5 + 5 \times 0.2 \times 0.8^4 + 10 \times 0.2^2 \times 0.8^3$$

$$= 0.328 + 0.41 + 0.205$$

$$= 0.942$$

Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (13/19)

Poisson random variable and probability distribution:

- Named after the French scientist Simeon D. Poisson.
- Describes the frequency of events for a given interval of time or space.

Ex) Number of emails received per hour.

Ex) Number of earthquakes per year,

Ex) Number of chocolate chips in a cookie.

Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (14/19)

Poisson random variable and probability distribution:

1) "X is a Poisson random variable" $\Leftrightarrow X \sim Pois(\lambda)$

2) $P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \Leftarrow 0 \leq x$

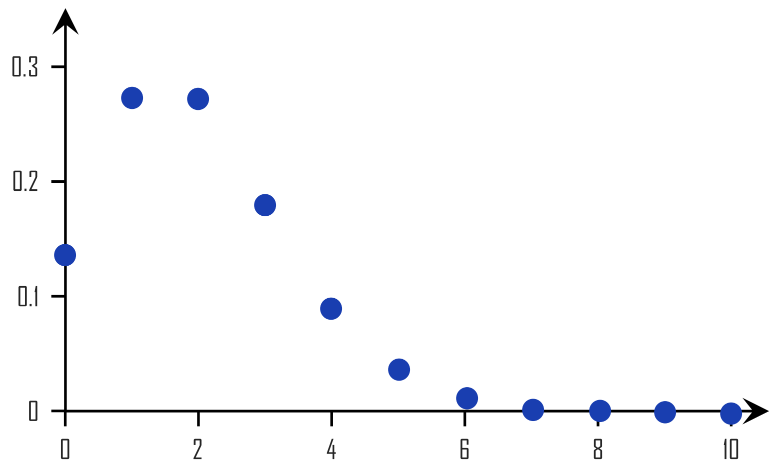
- Mean =λ
- Variance =λ
- Standard deviation = $\sqrt{\lambda}$

3) By increasing n while decreasing p the Binomial distribution converges to the Poisson with $\lambda = n\ p$.

Discrete Probability Distribution (15/19)

- Poisson random variable and probability distribution:

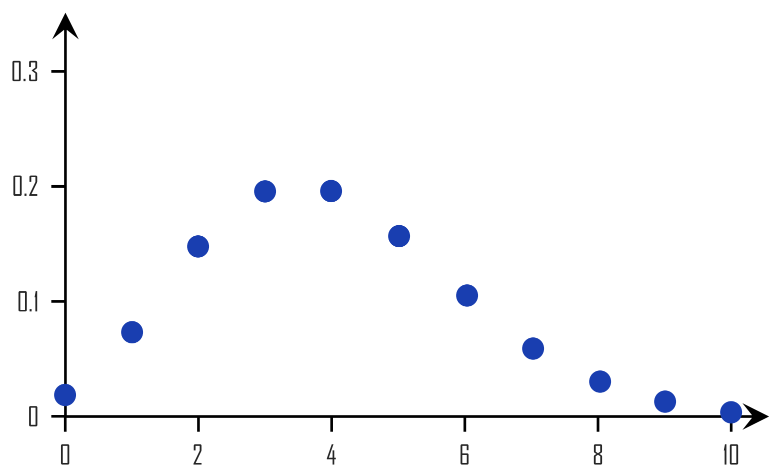
$lambda = 2$



Discrete Probability Distribution (16/19)

- Poisson random variable and probability distribution:

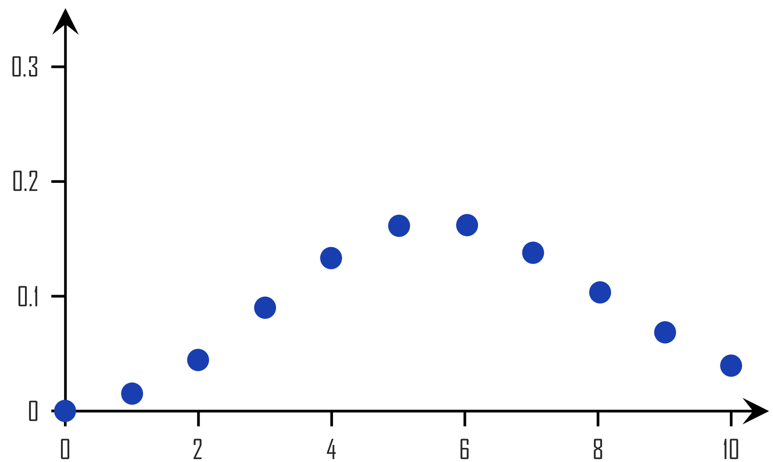
$$lambda = 4$$



Discrete Probability Distribution (17/19)

- Poisson random variable and probability distribution:

$lambda = 6$



Discrete Probability Distribution (18/19)

Poisson random variable and probability distribution:

Ex) During the last 100 days 40 spam mails were received.

What is the probability that tomorrow there will be zero spam mail?

You can calculate that $\lambda = \frac{40}{100} = 0.4$

Then, $P(0) = \frac{0.4^0 e^{-0.4}}{0!} = e^{-0.4} = 0.6703$

Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (19/19)

Poisson random variable and probability distribution:

Ex) During the last 100 days 40 spam mails were received.

What is the probability that tomorrow there will be more than one spam mails?

You can calculate that $\lambda = \frac{40}{100} = 0.4$

Then, do $P(1) + P(2) + P(3) + \cdots$?? $\Rightarrow$ Hard!

Instead, you can calculate by doing $1 - P(0) = 1 - 0.6703 = 0.3297$

Together for Tomorrow! Enabling People
Education for Future Generations

Coding Exercise #0301

Follow practice steps on 'ex_0301.ipynb' file.

Chapter 4.

Probability and Statistics

UNIT 2.

Understanding of Statistics I

Unit 2.
Understanding of Statistics I

What this unit is about:

- You will learn about the continuous probability densities.
- You will learn about the conjoint probability.
- You will learn about the correlation and linear relationship.

Expected outcome:

- Ability to define and calculate the probability.
- Ability to calculate the expected values.
- Ability to model random events with the appropriate probability density.

How to check your progress:

- Coding Exercises.
- Quiz.

Together for Tomorrow! Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

UNIT 1. Understanding of Probability

1.1. Probability Theory.
1.2. Probability Rules.
1.3. Random Variable.
1.4. Discrete Probability Distribution.

Unit 2. Understanding of Statistics I

2.1. Continuous Probability Density.
2.2. Conjoint Probability.

Unit 3. Understanding of Statistics II

3.1. Descriptive Statistics.
3.2. Central Limit Theorem.
3.3. Estimation Theory.

Unit 4. Statistical Hypothesis Testing

4.1. Principles of Hypothesis Testing.
4.2. Hypothesis Testing in Action.

Continuous Probability Density (1/24)

Continuous random variable and probability density:

- Infinite number of possible values.
- The probability at a specific value is zero: $P(X = \boldsymbol{x_0}) = 0$
- Non-zero probability only for intervals: $P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$
- Cumulative probability: $CDF(x) = P(-\infty < X \leq x)$

$$= \int_{-\infty}^{x} f(y)dy$$

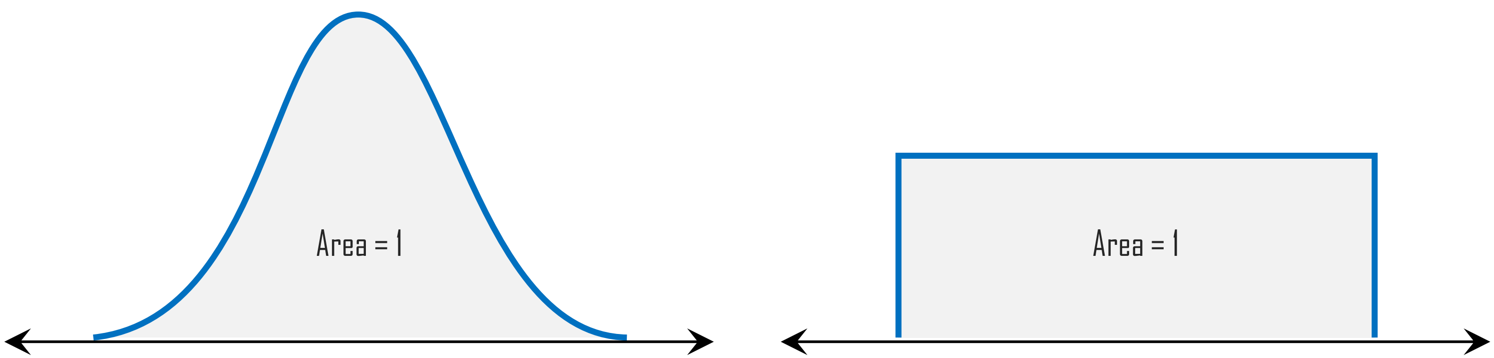
- We can calculate $P(x_1 \leq X \leq x_2) = CDF(x_2) - CDF(x_1)$

Continuous Probability Density (2/24)

Properties of **continuous** probability density:

1) $0 \leq f(x)$

2) $\int f(x)dx = 1$



Continuous Probability Density (3/24)

Uniform random variable and probability density:

1) "X is a Uniform random variable in the interval $[a, b]$" $\Leftrightarrow X \sim Unif(a, b)$

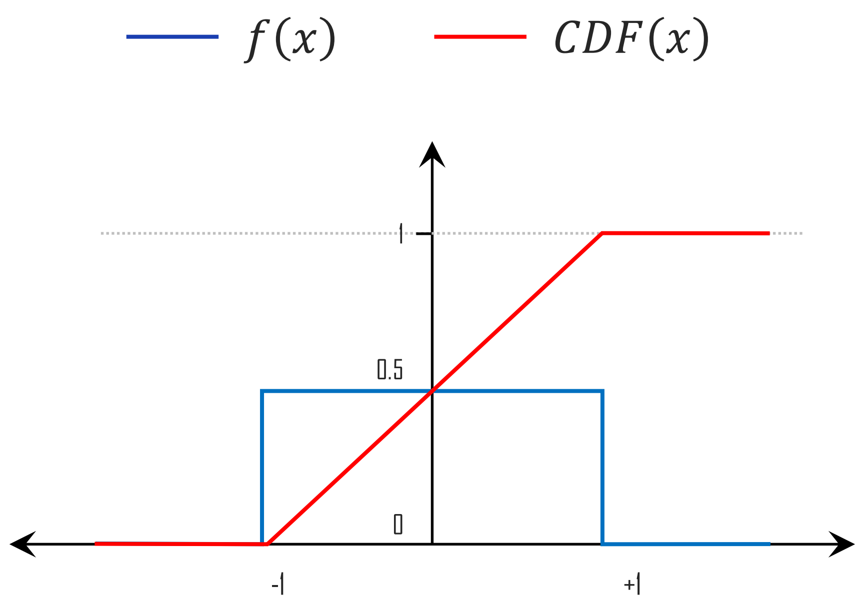
2) $f(x) = \frac{1}{(b-a)}$ $\Leftarrow$ defined in the interval $[a, b]$ and zero elsewhere.

- Mean$= \frac{1}{2}(a + b)$
- Variance$= \frac{1}{12}(b - a)^2$
- Standard deviation$= \frac{1}{\sqrt{12}}(b - a)$

Continuous Probability Density (4/24)

Uniform random variable and probability density:

Ex) $a = -1, b = +1$



Continuous Probability Density (5/24)

Uniform random variable and probability density:

Ex) The lifespan of a light bulb is uniformly distributed between 10,000 and 20,000 hours.

What is the probability that the light bulb will last between 12,000 and 15,000 hours of usage?

$$P(12000 \leq X \leq 15000) = \int_{12000}^{15000} \frac{1}{20000 - 10000} dx = 0.3$$

Continuous Probability Density (6/24)

Uniform random variable and probability density:

Ex) The lifespan of a light bulb is uniformly distributed between 10,000 and 20,000 hours.

What is the probability that the light bulb will last 15,000 hours or more?

$$P(15000 \le X \le 20000) = \int_{15000}^{20000} \frac{1}{20000 - 10000} dx = 0.5$$

Continuous Probability Density (7/24)

Normal random variable and probability density:

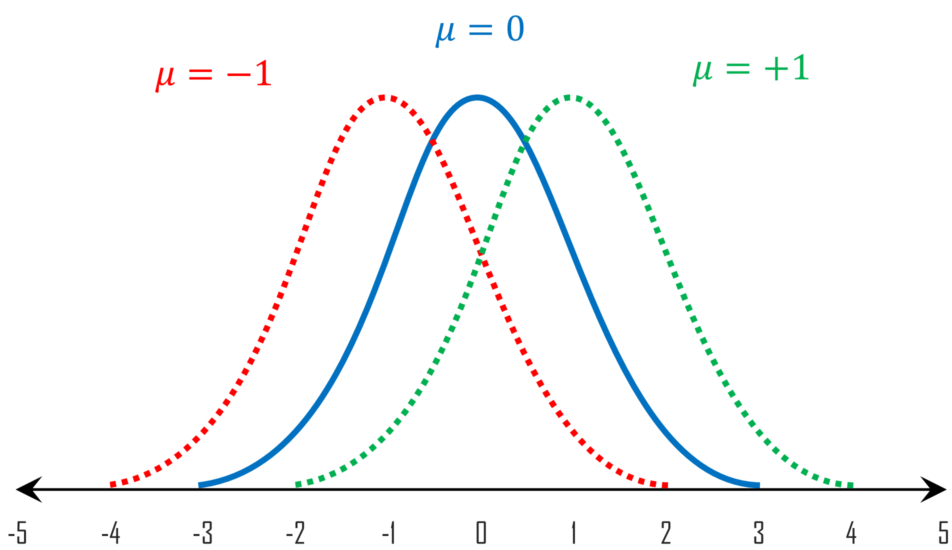
1) "X is a Normal random variable with mean μ and variance σ^2" $\Leftrightarrow X \sim N(\mu, \sigma^2)$

2) $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $\Leftarrow$ defined in the interval $(-\infty, +\infty)$

- Mean $= \mu$
- Variance $= \sigma^2$
- Standard deviation $= \sigma$

Continuous Probability Density (8/24)

Normal random variable and probability density:

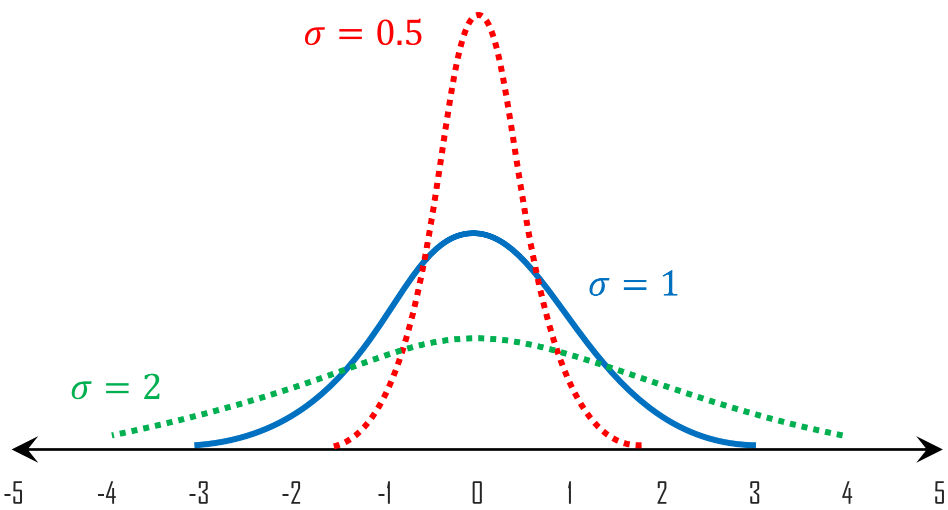


- μ is the so-called "**location**" parameter.

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (9/24)

Normal random variable and probability density:



- σ is the so-called "**shape**" parameter.

Continuous Probability Density (10/24)

Standard Normal random variable and probability density:

1) "X is a Standard Normal random variable" $\Leftrightarrow X \sim N(0,1)$

2) $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ $\Leftarrow$ defined in the interval $(-\infty, +\infty)$

- Mean = 0
- Variance = 1
- Standard deviation = 1