* Review 1st week
* Probability space , the sample space, Algebra, Probability measure

1. What is ?
2. Joint probability,
3. Marginal Probability If ,
4. Conditional Probability

* See the Week\_1 Supplement

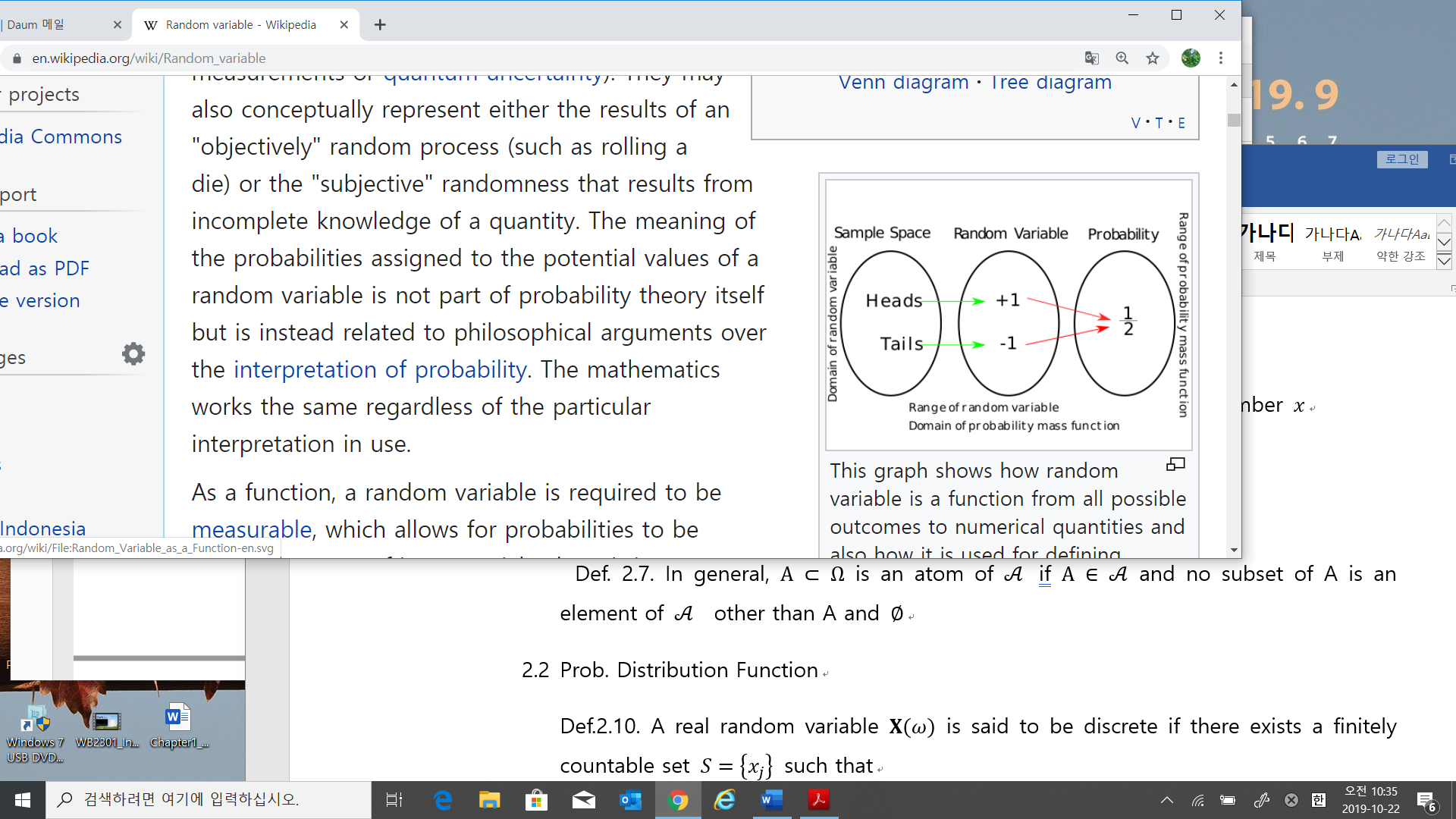
1. Random Variables and Stochastic Process
   1. Random Variables

Def. 2.1. Given a probability space,, a random variable is a real-(vector-) valued point function which carries a sample point, , into a point

in such a way that every sets, , of the form

is an element of the

* In the textbook, misspelled. As
* Random variable
* is a function such that associates to a real number
* Wiki: <https://en.wikipedia.org/wiki/Random_variable>



Ex.2.4. The Experiment: two coins flips

The sample space

The Event: at least one Head:

A candidate random variable

We may call this as the Indicator random variable

Now the event generated by may be defined (as you like)

Probability graph in

0 1

3/4

1/4

%%% Kim’s comment:

What is in this case? Well

And

Hence we may calculate for any %%%%

Def. 2.7. In general, is an atom of if and no subset of A is an element of other than A and .

Hence are atoms of

%%% Kim’s comment :

See the notation , %%%

* 1. Prob. Distribution Function
* **Probability Distribution function**

Def.2.10. A real random variable is said to be discrete if there exists a finitely countable set such that

* 1. Prob. Density Function

Suppose such that

Then is **the probability density function**,

Proposition 2.12.

%%% Kim’s comment : in the limit notation

%%%

%%% kim’s comment on

One of the special function in mathematics is function. The definition of is

In the shift form

What is the value of ? It may be called as an impulse function. As you see

Now for any constant , which is as large as possible

So to satisfy (1), the magnitude of , which is not defined at the real number. In fact the delta function is not a function. We may remember in the system theory, the Laplace transform of the delta function, i.e.,

%%%%

* Common (Probability) Distribution Functions for Random Variables

1. The uniform distribution function
2. The exponential distribution function
3. The Gaussian probability distribution(the normal distribution function)

A Gaussian random vector , the density is

where the mean vector,

the covariance matrix,

the determinant of

* 1. Probabilistic Concepts Applied to Random Variables
* Joint Probability Distribution
* Marginal Probability
* Joint Probability Density
* Marginal Probability Density
* The marginal probability distribution

%%% Kim’s Example

Given find the marginal

Ok. That is, by definition, a continuous case. In discrete case

Given a group of people, there are two experiments as measurement of temperature and blood pulse rate. Let us denote as one person

The probabilities of the outcomes are

Then the marginal probabilities are

And

You may compare the continuous case. %%%

%%% Kim’s Example for continuous case

Give the joint probability as

1. Is it a PDF ( or CDF) ?

One the necessary condition is



1. Find marginal probabilities
2. Is **independent**? No since

* HA\_2\_1 :

%%%

1. Is this a PDF?
2. Find marginal probabilities

%%%

Def 2.16. Two random variables and are called **independent** if any event of the form is independent of any event of the form where are sets in

* Fact
* The joint probability distribution
* The joint probability density function
  1. Functions of a Random Variable**-skip**

has the density function

Where stands for the absolute value of the determinant of the matrix

* 1. Expectations and Moments of a Random Variable

Def.

* The mean
* The sample mean

%% The sample mean is a Random Variable! It is an estimator of the mean of a random variable . If is an **independent identical distributed (iid)** random variable,i.e.,

Then the mean of the sample mean is

%%%Kim’s Comment

What is the difference between a) and b)? In order to use (a), it is needed know the probability density function, whereas in (b), not needed.

%%%

Examp. 2.19. is uniformly distributed from 0 to 1, i.e.,

Then

Examp. 2.22 The expectation of the value of one roll of one die?

Properties

1. The operator of expectation is linear
2. The square mean / second moment
3. The higher order moment
4. The variance
5. The standard deviation
6. The sample variance

This is a random variable. And the **unbiased** estimator of

%%% Kim’s comment: biased and un-biased estimator

What is the estimator? Let be a RV. I want to find a constant “C” as RV in some sense.

We may call C as an estimator of the RV . So there may be many estimators as you like.

We may classify the estimator as

1. Unbiased estimator / biased estimator

If , then C is the unbiased estimator, otherwise the biased estiamtor

1. The minimum variance estimator /the least square error estimator
2. **The mean of is the minimum variance estimator / the least square error estimator**.

Proof:

* , which minimizes the (c).

Examp. 2.24 The uniform distributed random variable

The Variance is

* 1. Characteristic Functions**-skip**

Lemma 2.27

Prop.2.28 If is a Gaussian random vector with mean, m, and covariance matrix P, then its characteristic function is

%%% Kim’s comment: **correlation**

Def: Two R.V. are **uncorrelated** if

%%%

Fact: Two GaussianR.Vectors are uncorrelated if is a diagonal matrix

* **Prop. 2.29. Uncorrelated Gaussian random variables are independent**

Theorem 2.30. If is a Gaussian random vector with mean , and covariance, , and if , where is a Gaussian random vector with zero mean and covariance, , then is a Gaussian random vector with mean, , and covariance, .

* **Theorem 2.30**

**A R.V , another R.V. and they are independent**. Find mean and covariance of

%%% Kim’s comment: Characteristic function is difficult to remember. In the text book, using the characteristic method. In this case we may apply basic theory.

Sol: Let’s apply the basic definition.

Hence

* In general, independency implies the uncorrelated, not vice versa
* However, in Gaussian Does satisfy the opposite direction. %%%

%%% Kim’s comment: covariance matrix

Sometimes, but most case in this course, we may deal with a random vector whose components are random variables, i.e.

is a random vector, its components are random variables Then the covariance of random vector is defined ad

where

hence by definition

Therefore the matrix is a symmetric matrix, i.e.,

The diagonal terms of the covariance matrix are variance of each random variable

%%%

* The covariance of a uncorrelated (so independent) Gaussian is a diagonal matrix,

%%% Kim’s comment: Linear matrix theory: similar transform

For any semi-positive symmetric matrix , there is a **similar transform matrix**such that

Hence the covariance for any gaussian Random vectors (correlated), there exits a such that

* Any Gaussian Random vectors, we can find a transformed Random Vectors which is uncorrelated(independent).
* Independency is important to calculate the probability. You know the Gaussian probability table, but it is a scalar. So it you want to calculate the joint probability which may be correlated, first find a similar transform matrix to generate a diagonal covariance matrix. Then you may calculate the joint probability as a separate probability.

%%%

* **The central limit theorem**

Theorem 2.31. Let be i.i.d. random variables with finite mean and variance,

and denote their sum as . Then the distribution of the normalized sum

is a Gaussian distribution with mean 0 and variance 1 in the limit as

* Proof : textbook P.52
* Remarks:

1. See, the condition, that means   
   the mean and the variance is constant, but the experiment is many time processing. For example,
2. A die, which is fair or not, you roll the same die many times. Then the mean of the sum () is a Gaussian if .
3. Some RV has no mean, then it will not be applicable.
   1. Conditional Expectations and Conditional Probabilities

* The conditional expectation
* Remarks
* is a constant, means it is not random variable.
* if is a constant, then is a constant
* if is a RV, then is a **Random Variable** of y
* **Iterated expectation(See the proof at p.57 and remember)**

%%% Kim’s comment

Even if we do not know.

I should say, this formula cannot emphasize too much! This very simple fact use diverse applications, big data, machine learning, and dynamic system analysis. We should **remember** this.

%%%

* Lemma 2.34.
  1. Stochastic Process
* Def. 2.36. A stochastic process is a family of random variables, , indexed by a real parameter and defined on a common probability space .

%%% Kim’s comment

A stochastic process (or random process) is a time varying random variable, i.e., for any fixed , the process is a random variable.

%%%

* Ex. 2.37
* Def. 2.38.

1. A stochastic process is said to be continuous in probability at t if

for all

1. Skip: A stochastic process is said to be separable if there exists a countable, dense set such that for any closed set

differ by a set such that

* Skip: Theorem 2.40. The rational numbers in provide a separating set S.
* Def. 2.42. Let X be a random process defined on the time interval, T. Let

be a partition of the time interval, T. If the increments,are mutually independent for any partition of T, then X is said to be a process with **independent increments**.

* Def. 2.43 We say that a random process, X, is a Gaussian process if for every finite collection, the corresponding density function,

is a Gaussian density function.

* Def. 2.44 We say that a random process X is a Gaussian process if every finite linear combination of the form

is a Gaussian random variable

* Def 2.45. A random process, where T is a subset of the real line, is said to be a **Markov process** if for any increasing collection

or, equivalently

* 1. Gauss-Markov Processes–**The fundamental**

1. Dynamics

* State , is a known matrix, is a Gaussian Random sequence.

1. Given Conditions
2. Noise

where

1. The states
2. The correlation

which implies

1. The mean and covariance

* The mean
* The covariance
  1. Non-linear Stochastic Difference Equations 🡪 skip