

Lecture 0: Introduction

Yi, Yung (이웅)

EE210: Probability and Introductory Random Processes
KAIST EE

MONTH DAY, 2021

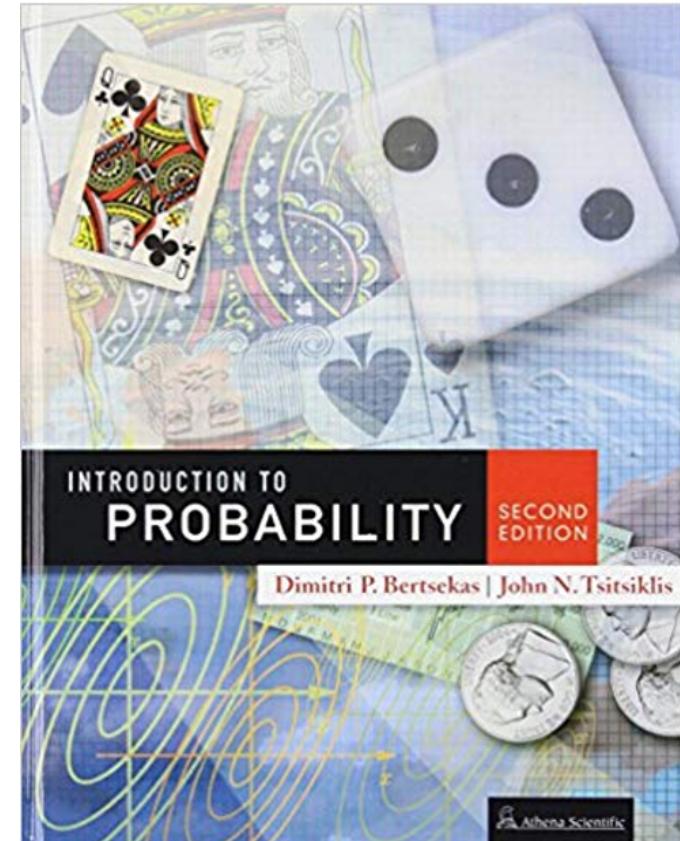
- Course logistics
- Why necessary to take the course of probability and random process?

- Yi, Yung (이융)
- Office: N1, 810
- <http://lanada.kaist.ac.kr>, yiyung@kaist.edu
- Computer Division
- A professor in KAIST EE since 2008
- Office hours: TBA

- A
 - B
 - C
-
- Mailing list: ee210@lanada.kaist.ac.kr
 - Please use KLMS for the questions about the lecture contents
 - This mailing list can be used for individual issues

- <http://klms.kaist.ac.kr/>
- To download course materials
- To ask questions about everything
- To check your score on each homework/exam
- To see all the announcements about the class

- Introduction to Probability
(2nd edition)
 - MIT course textbook
 - Dimitri P. Bertsekas and John N. Tsitsiklis



- Three Parts
 - Part I: Fundamentals of Probability
 - Part II: Inference and Limit Theorems
 - Part III: Random Processes
- On-line lectures at MIT and EdX
 - MIT: <http://bit.ly/2PkvYdr>
 - EdX: <http://bit.ly/3pHmZRd>
 - You can find older urls (2006, 2010, 2013) for this lecture, where there are many useful resources (recitation problems, homework problems, old exam problems, etc)
 - My lecture slides: based on theirs, but largely modified/reorganized/edited in many places for KAIST students

- In-class quiz (sometimes)
- Basically, weekly homework, but often bi-weekly
- 3 Exams (2 mid-terms and 1 final)
- Class participation
- Grading portions: A (X%), B (Y%), C(Z%), D(W%), F . . .
- Online lectures due to COVID-19 may change how to grade.

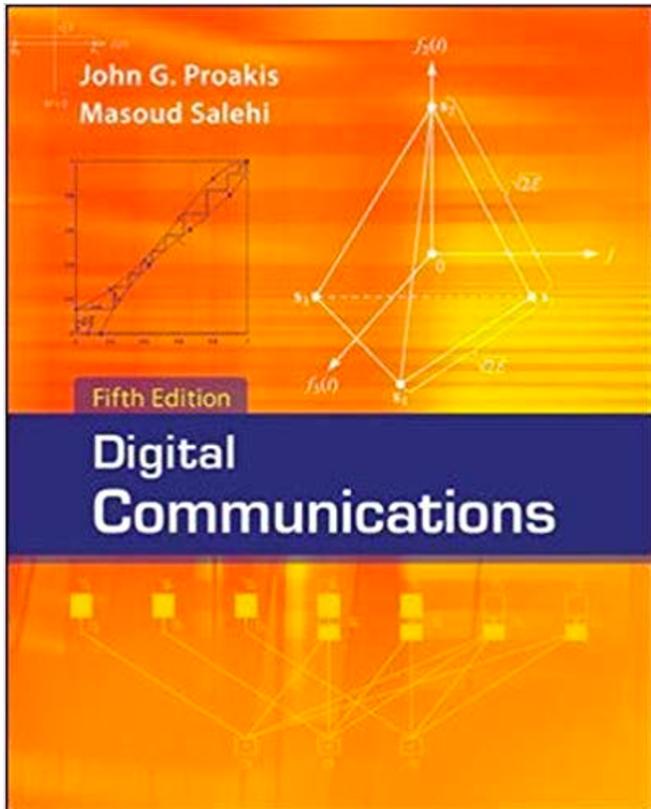
- Most should be via KLMS
 - Technical questions about lectures, homework, and etc
- Please DO NOT individually send emails to Prof. Yung Yi and TAs (or making calls or sending KakaoTalk msgs) about the technical questions (course contents, homework, etc)
 - All the questions need to be shared among the students.
 - TAs and Prof. Yung Yi will handle your questions as soon as possible.
 - But, you can send an email to Prof. Yung Yi for the things that need to be individually discussed.

Questions?

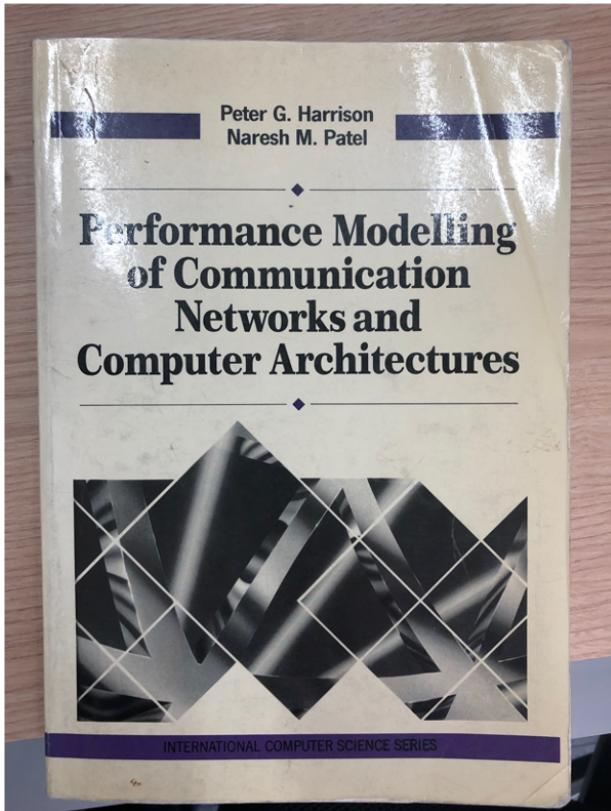
- Many things are "probabilistic"

- Many things are "probabilistic"
- Assume that you are a designer of the following engineering systems. Good design?
 - a web server
 - a communication device like mobile phones
 - an AI-based image classifier

- Many things are "probabilistic"
- Assume that you are a designer of the following engineering systems. Good design?
 - a web server
 - a communication device like mobile phones
 - an AI-based image classifier
- From an engineering point of view,
 - System input
 - Algorithms in systems
 - Analysis of systems



| | |
|--|----|
| Communications | 13 |
| 1-5 Overview of the Book | 16 |
| 1-6 Bibliographical Notes and References | 16 |
| | |
| 2 Probability and Stochastic Processes | 17 |
| 2-1 Probability | 17 |
| 2-1-1 Random Variables, Probability Distributions, and Probability Densities | 22 |
| 2-1-2 Functions of Random Variables | 28 |
| 2-1-3 Statistical Averages of Random Variables | 33 |
| 2-1-4 Some Useful Probability Distributions | 37 |
| 2-1-5 Upper bounds on the Tail Probability | 53 |
| 2-1-6 Sums of Random Variables and the Central Limit Theorem | 58 |
| 2-2 Stochastic Processes | 62 |
| 2-2-1 Statistical Averages | 64 |
| 2-2-2 Power Density Spectrum | 67 |
| 2-2-3 Response of a Linear Time-Invariant System to a Random Input Signal | 68 |
| 2-2-4 Sampling Theorem for Band-Limited Stochastic Processes | 72 |
| 2-2-5 Discrete-Time Stochastic Signals and Systems | 74 |
| 2-2-6 Cyclostationary Processes | 75 |
| 2-3 Bibliographical Notes and References | 77 |
| Problems | 77 |



Preface

Chapter 1 Essentials of Probability Theory

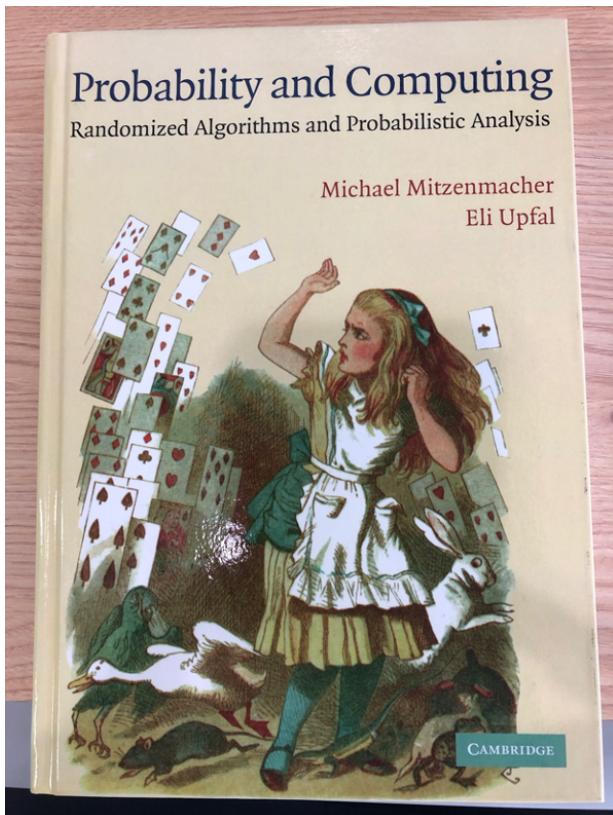
- 1.1 Sample space, events and probability
- 1.2 Conditional probability
- 1.3 Independence
- Exercises

Chapter 2 Random Variables and Distributions

- 2.1 Probability distribution functions
- 2.2 Discrete random variables
- 2.3 Continuous random variables
- 2.4 Joint random variables
- 2.5 Conditional distributions
- 2.6 Independence and sums
- Exercises

Chapter 3 Expected Values and Moments

- 3.1 Expectation
- 3.2 Generating functions and transforms
- 3.3 Asymptotic properties
- Exercises



Preface

1 Events and Probability

- 1.1 Application: Verifying Polynomial Identities
- 1.2 Axioms of Probability
- 1.3 Application: Verifying Matrix Multiplication
- 1.4 Application: A Randomized Min-Cut Algorithm
- 1.5 Exercises

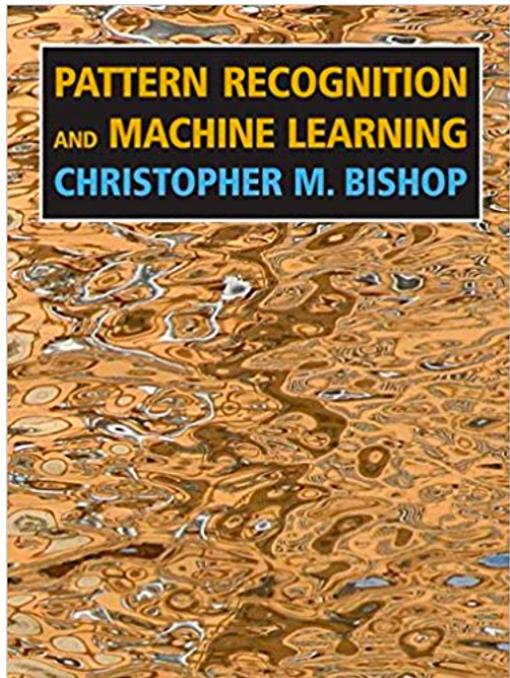
2 Discrete Random Variables and Expectation

- 2.1 Random Variables and Expectation
 - 2.1.1 Linearity of Expectations
 - 2.1.2 Jensen's Inequality
- 2.2 The Bernoulli and Binomial Random Variables
- 2.3 Conditional Expectation
- 2.4 The Geometric Distribution
 - 2.4.1 Example: Coupon Collector's Problem
- 2.5 Application: The Expected Run-Time of Quicksort
- 2.6 Exercises

3 Moments and Deviations

- 3.1 Markov's Inequality
- 3.2 Variance and Moments of a Random Variable
 - 3.2.1 Example: Variance of a Binomial Random Variable
- 3.3 Chebyshev's Inequality
 - 3.3.1 Example: Coupon Collector's Problem
- 3.4 Application: A Randomized Algorithm for Computing the
 - 3.4.1 The Algorithm
 - 3.4.2 Analysis of the Algorithm
- 3.5 Exercises

Textbook: Machine Learning



| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|---|--------------------------------|----|---------------------------------------|----|-------------------------------------|----|--|----|---|----|--|----|---|----|---|----|---|----|---------------------------------------|----|---|----|--|-----|------------------------------------|-----|---------------------------------------|-----|--------------------------------------|-----|--|-----|----------------------------------|-----|---------------------------------------|-----|-------------------------------------|-----|---|-----|---|-----|---------------------|-----|------|--|--|--|--|-----|--|-----|---|-----|-------------------------------------|-----|---|-----|-----------------|-----|
| Copyrighted Material xiv | <p>Copyrighted Material</p> <p>CONTENTS</p> <p>2 Probability Distributions 67</p> <table border="0"><tr><td>2.1 Binary Variables</td><td style="text-align: right;">68</td></tr><tr><td> 2.1.1 The beta distribution</td><td style="text-align: right;">71</td></tr><tr><td>2.2 Multinomial Variables</td><td style="text-align: right;">74</td></tr><tr><td> 2.2.1 The Dirichlet distribution</td><td style="text-align: right;">76</td></tr><tr><td>2.3 The Gaussian Distribution</td><td style="text-align: right;">78</td></tr><tr><td> 2.3.1 Conditional Gaussian distributions</td><td style="text-align: right;">85</td></tr><tr><td> 2.3.2 Marginal Gaussian distributions</td><td style="text-align: right;">88</td></tr><tr><td> 2.3.3 Bayes' theorem for Gaussian variables</td><td style="text-align: right;">90</td></tr><tr><td> 2.3.4 Maximum likelihood for the Gaussian</td><td style="text-align: right;">93</td></tr><tr><td> 2.3.5 Sequential estimation</td><td style="text-align: right;">94</td></tr><tr><td> 2.3.6 Bayesian inference for the Gaussian</td><td style="text-align: right;">97</td></tr><tr><td> 2.3.7 Student's t-distribution</td><td style="text-align: right;">102</td></tr><tr><td> 2.3.8 Periodic variables</td><td style="text-align: right;">105</td></tr><tr><td> 2.3.9 Mixtures of Gaussians</td><td style="text-align: right;">110</td></tr><tr><td>2.4 The Exponential Family</td><td style="text-align: right;">113</td></tr><tr><td> 2.4.1 Maximum likelihood and sufficient statistics</td><td style="text-align: right;">116</td></tr><tr><td> 2.4.2 Conjugate priors</td><td style="text-align: right;">117</td></tr><tr><td> 2.4.3 Noninformative priors</td><td style="text-align: right;">117</td></tr><tr><td>2.5 Nonparametric Methods</td><td style="text-align: right;">120</td></tr><tr><td> 2.5.1 Kernel density estimators</td><td style="text-align: right;">122</td></tr><tr><td> 2.5.2 Nearest-neighbour methods</td><td style="text-align: right;">124</td></tr><tr><td>Exercises</td><td style="text-align: right;">127</td></tr><tr><td colspan="2"> </td></tr><tr><td>3 Linear Models for Regression 137</td><td></td></tr><tr><td>3.1 Linear Basis Function Models</td><td style="text-align: right;">138</td></tr><tr><td> 3.1.1 Maximum likelihood and least squares</td><td style="text-align: right;">140</td></tr><tr><td> 3.1.2 Geometry of least squares</td><td style="text-align: right;">143</td></tr><tr><td> 3.1.3 Sequential learning</td><td style="text-align: right;">143</td></tr><tr><td> 3.1.4 Regularized least squares</td><td style="text-align: right;">144</td></tr><tr><td> 3.1.5</td><td style="text-align: right;">144</td></tr></table> <p>Copyrighted Material</p> | 2.1 Binary Variables | 68 | 2.1.1 The beta distribution | 71 | 2.2 Multinomial Variables | 74 | 2.2.1 The Dirichlet distribution | 76 | 2.3 The Gaussian Distribution | 78 | 2.3.1 Conditional Gaussian distributions | 85 | 2.3.2 Marginal Gaussian distributions | 88 | 2.3.3 Bayes' theorem for Gaussian variables | 90 | 2.3.4 Maximum likelihood for the Gaussian | 93 | 2.3.5 Sequential estimation | 94 | 2.3.6 Bayesian inference for the Gaussian | 97 | 2.3.7 Student's t-distribution | 102 | 2.3.8 Periodic variables | 105 | 2.3.9 Mixtures of Gaussians | 110 | 2.4 The Exponential Family | 113 | 2.4.1 Maximum likelihood and sufficient statistics | 116 | 2.4.2 Conjugate priors | 117 | 2.4.3 Noninformative priors | 117 | 2.5 Nonparametric Methods | 120 | 2.5.1 Kernel density estimators | 122 | 2.5.2 Nearest-neighbour methods | 124 | Exercises | 127 | | | 3 Linear Models for Regression 137 | | 3.1 Linear Basis Function Models | 138 | 3.1.1 Maximum likelihood and least squares | 140 | 3.1.2 Geometry of least squares | 143 | 3.1.3 Sequential learning | 143 | 3.1.4 Regularized least squares | 144 | 3.1.5 | 144 |
| 2.1 Binary Variables | 68 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.1.1 The beta distribution | 71 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.2 Multinomial Variables | 74 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.2.1 The Dirichlet distribution | 76 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3 The Gaussian Distribution | 78 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.1 Conditional Gaussian distributions | 85 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.2 Marginal Gaussian distributions | 88 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.3 Bayes' theorem for Gaussian variables | 90 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.4 Maximum likelihood for the Gaussian | 93 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.5 Sequential estimation | 94 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.6 Bayesian inference for the Gaussian | 97 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.7 Student's t-distribution | 102 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.8 Periodic variables | 105 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.3.9 Mixtures of Gaussians | 110 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.4 The Exponential Family | 113 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.4.1 Maximum likelihood and sufficient statistics | 116 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.4.2 Conjugate priors | 117 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.4.3 Noninformative priors | 117 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.5 Nonparametric Methods | 120 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.5.1 Kernel density estimators | 122 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2.5.2 Nearest-neighbour methods | 124 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Exercises | 127 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 Linear Models for Regression 137 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3.1 Linear Basis Function Models | 138 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3.1.1 Maximum likelihood and least squares | 140 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3.1.2 Geometry of least squares | 143 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3.1.3 Sequential learning | 143 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3.1.4 Regularized least squares | 144 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3.1.5 | 144 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

These days, every area in CS and EE is directly or indirectly related to machine learning!

- Designer's perspective?

- Designer's perspective?
- In the year of 2021, suppose that unfortunately there is no theory of mathematically studying the *uncertainty* of some phenomena, events, etc.

- Designer's perspective?
- In the year of 2021, suppose that unfortunately there is no theory of mathematically studying the *uncertainty* of some phenomena, events, etc.
- You have to design such a theory called "probability". How are you going to do it?
Where are you going to start?
- You just have other basic mathematical theories such as set theory.

- Designer's perspective?
- In the year of 2021, suppose that unfortunately there is no theory of mathematically studying the *uncertainty* of some phenomena, events, etc.
- You have to design such a theory called "probability". How are you going to do it?
Where are you going to start?
- You just have other basic mathematical theories such as set theory.
- You need to get used to the *English terms* on probability (e.g., sample space = 표본공간, probability density function = 확률밀도함수).

- Designer's perspective?
- In the year of 2021, suppose that unfortunately there is no theory of mathematically studying the *uncertainty* of some phenomena, events, etc.
- You have to design such a theory called "probability". How are you going to do it?
Where are you going to start?
- You just have other basic mathematical theories such as set theory.
- You need to get used to the *English terms* on probability (e.g., sample space = 표본공간, probability density function = 확률밀도함수).
- We will take this exciting journey from the next lecture!

Questions?



@imagelife.xyz