

# Lecture 0: Introduction

Yi, Yung (이융)

EE210: Probability and Introductory Random Processes  
KAIST EE

August 27, 2022

- Course logistics
- Why this course?

- Yi, Yung (이웅)
- Office: N1, 810
- Homepage: <https://yung-web.github.io/home/>
- E-mail: [yiyung@kaist.edu](mailto:yiyung@kaist.edu)
- Computer Division
- In KAIST EE since 2008

- All lecture videos have already been pre-recorded. Available in **YouTube**.
- **non-real-time online ( $\leq 50\%$ ) + real-time offline/online ( $\geq 50\%$ )**
- **non-real-time online:** Watch and study anytime and anywhere you like.
- **realtime offline/online:** Watch lecture videos in the classroom or in the zoom, with asking and answering questions.
- No attendance check!

# Accessing Lecture Videos and Slides



- Method 1:

<https://yung-web.github.io/home/courses/probability.html>

- Method 2: (a) Type Yung Yi in the google, (b) visit his GitHub homepage, (c) find the links on Course.

Google search results for "yung yi":

- Yung Yi - Google Scholar: Yung Yi, Professor of Electrical Engineering, KAIST. Verified email at kaist.edu - Homepage · Applied machine learningcomputer networkingperformance ...
- Yi, Yung – KAIST ELECTRICAL ENGINEERING: Yi, Yung Yi, Yung · Research Group. Computer · Research. Machine-learning based computer networking and communication systems, modeling, analysis, and developing ...
- Yung Yi, KAIST - GitHub Pages**: 2021.6.15. Short Bio: Yung Yi received his B.S. and the M.S. in the School of Computer Science and Engineering from Seoul National University, ...
- Yung Yi | OpenReview: Korea Advanced Institute of Science and Technology · Names · Emails · Personal Links · Education & Career History · Advisors, Relations & Conflicts · Expertise.
- Yung Yi - Home - ACM Digital Library: Mobile networks Wireless access networks Machine learning Network protocols Sequential decision making Design and analysis of algorithms Local area networks ...

Profile page for Yung Yi:



Short Bio: Yung Yi received his B.S. and the M.S. in the School of Computer Science and Engineering from Seoul National U in 1997 and 1999, respectively, and his Ph.D. in the Department of Electrical and Computer Engineering at the University o in 2006. From 2006 to 2008, he was a post-doctoral research associate in the Department of Electrical Engineering at Prince he is a KAIST Chair Full professor at the Department of Electrical Engineering at KAIST, South Korea. His current research i applied machine learning, design and analysis of wired/wireless networking systems. He was the recipient of two best paper SECON 2013 and ACM MobiHoc 2013. He was the co-recipient of IEEE William R. Bennett Award, 2016.

**LANADA (LAboratory of Network Architecture, Design, and Analysis)**  
LANADA is a research group which I currently lead. Currently, we do not hire new graduate students.

**Education**

- Ph.D: Dept. of Electrical and Computer Engineering, University of Texas at Austin, 2006
- M.S: Dept. of Computer Science and Engineering, Seoul National University, 1999
- B.S: Dept. of Computer Science and Engineering, Seoul National University, 1997

**Position**

- KAIST Chair Professor (KAIST 지정석교수): Dept. of Electrical Engineering, KAIST, 2021 - Current
- Full Professor: Dept. of Electrical Engineering, KAIST, 2018.2 - Current
- Associate Professor: Dept. of Electrical Engineering, KAIST, 2011.8 - 2018.2
- Assistant Professor: Dept. of Electrical Engineering, KAIST, 2008.8 - 2011.8
- Postdoctoral Research Associate: Dept. of Electrical Engineering Princeton University, 2006.8 - 2008.0

**Courses**

- Probability and Introductory Random Process (video included), Undergraduate
- Data Structures for Electrical Engineers, Undergraduate
- Mathematics for Machine Learning, Undergraduate
- Computer Network, Undergraduate
- Complex Network Analysis: Epidemics and Rumours (video included), Graduate

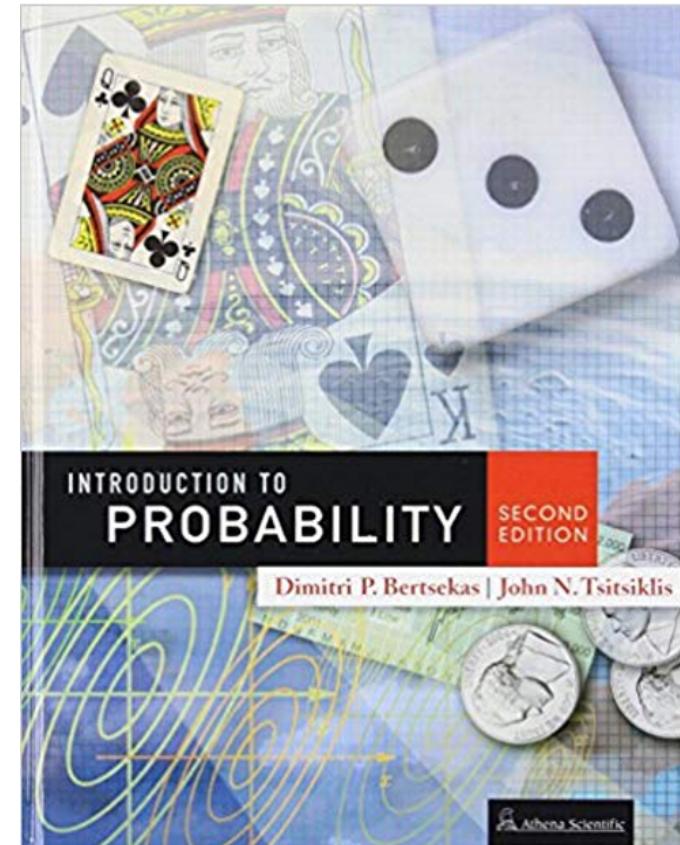
**Contact**  
291 Daehak-  
Electrical Eng.  
N1 Building,  
Daejon, South  
Phone: +82 42  
Fax: +82 42  
Email: lastna  
edu  
Office Hours

**Memor**  
  
2020 Spring

- **KLMS**: All notifications and announcements (also sent to you via email)
- **KLMS**: Questions about course logistics
- **KLMS**: Homework upload
- **KLMS**: Score upload and all the grade-related things
- **Campuswire**: Questions about course contents (should be in English)
- NOT individual emails to the instructor or the TAs
- Emails to the instructor, Prof. Yung Yi, are allowed for handling private situations.



- Introduction to Probability  
(2nd edition)
  - MIT course textbook
  - Dimitri P. Bertsekas and John N. Tsitsiklis
- You can order it from Yes24, Aladin, Kyobo
  - Yes24: <http://www.yes24.com/Product/Goods/3995311>
  - Aladin: <https://www.aladin.co.kr/shop/wproduct.aspx?ItemId=12945615>
  - Kyobo: <http://www.kyobobook.co.kr/product/detailViewEng.laf?ejkGb=ENG&mallGb=ENG&barcode=9781886529380&orderClick=LAG&Kc=>



- <http://athenasc.com/probbook.html>
- **Solutions for all problems** (so you have all solutions for your homework)
- Links to the old MIT courses
- You can find the urls (2006, 2010, 2013) for the MIT lectures based on the same textbook, where there are many useful resources (recitation problems, homework problems, old exam problems, etc)
- Some of my lecture slides are based on theirs, but my slides are largely modified/reorganized/edited in many places for our purpose.

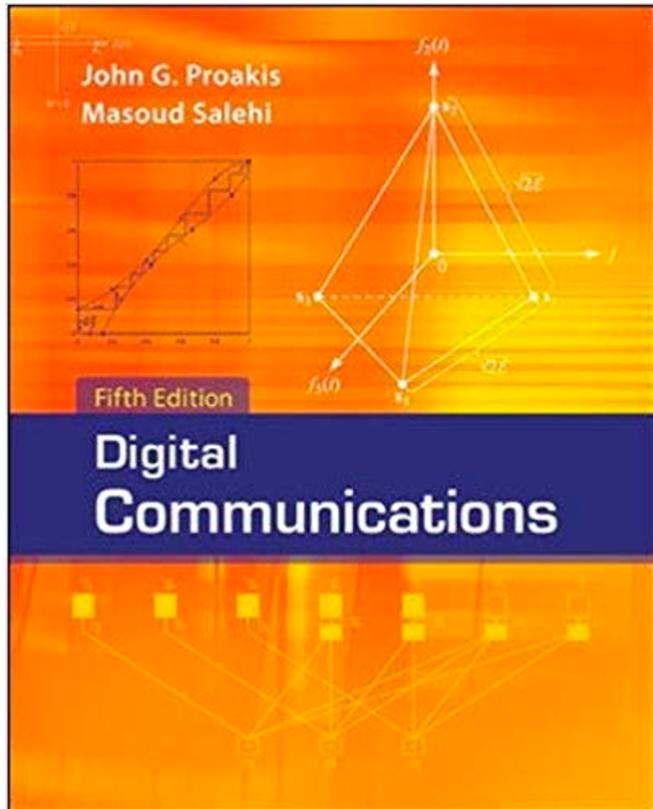
1. Probabilistic model (0.5 week)
2. Conditioning and Independence (0.5 week)
3. Random Variable, Part I (Discrete Random Variable) (1.5 week)
4. Random Variable, Part II (Continuous Random Variable) (1.5 week)
5. Random Variable, Part III (Advanced Topic on Random Variable) (1.5 week)
6. Limit of Scaled Sum of Random Variables: Central Limit Theorem and Weak Law of Large Numbers (1.5 week)
7. Random Process: Bernoulli and Poisson Processes (2 week)
8. Random Process: Markov Chain (2 week)
9. Introduction to Statistical Inference (2 week)

- 2 Exams (mid-term and final)
- Homeworks
  - All problems are from exercise problems in the textbook.
  - We do NOT check whether you copy your solution from the problem solutions or not.
- 9 Homeworks for each of 9 chapters.

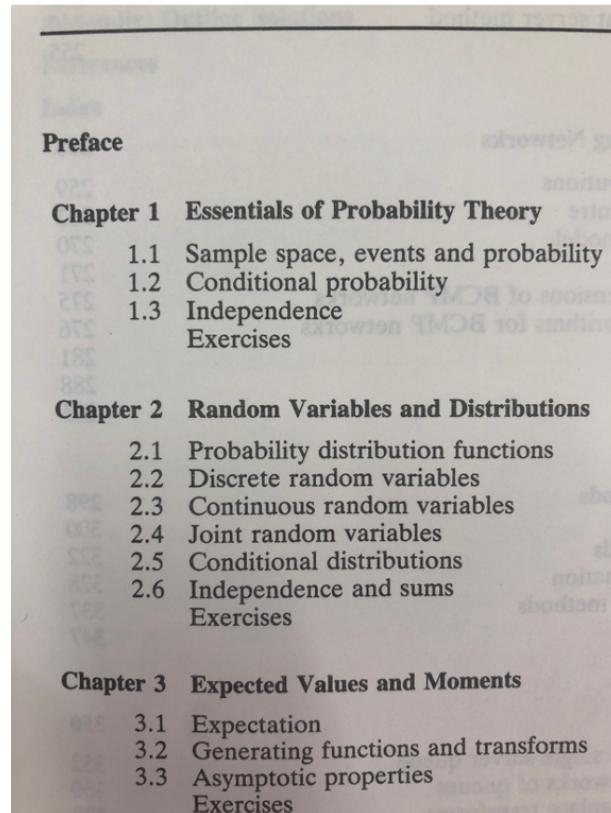
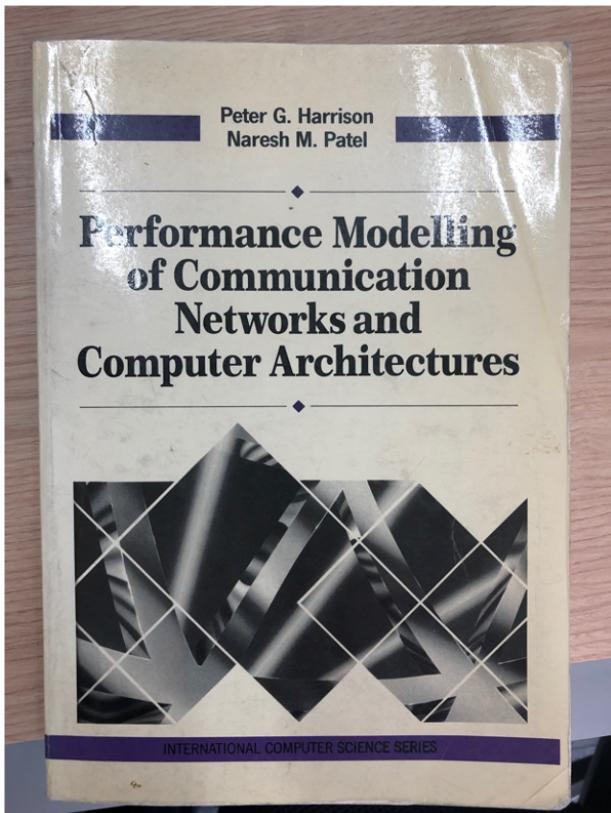
- Read **ALL** the emails.
- Need to buy the textbook?
  - Strongly recommend it. Taking a course is NOT just solving mid-term and final exam problems and getting a good grade.
- OK not to be present in the classroom? Yes.
- OK that my homework solutions is same as those in the solutions book? Yes.
- Can I ask for a personal meeting to ask questions or get other general advices?  
Sure. Send me an email.

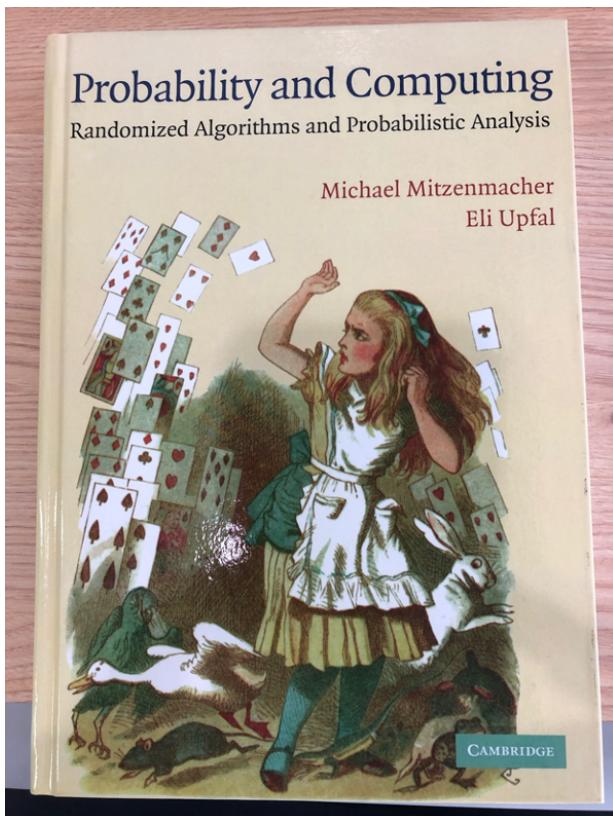
Questions?

- 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



<b>Communications</b>	13
1-5 Overview of the Book	16
1-6 Bibliographical Notes and References	16
<b>2 Probability and Stochastic Processes</b>	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*

**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

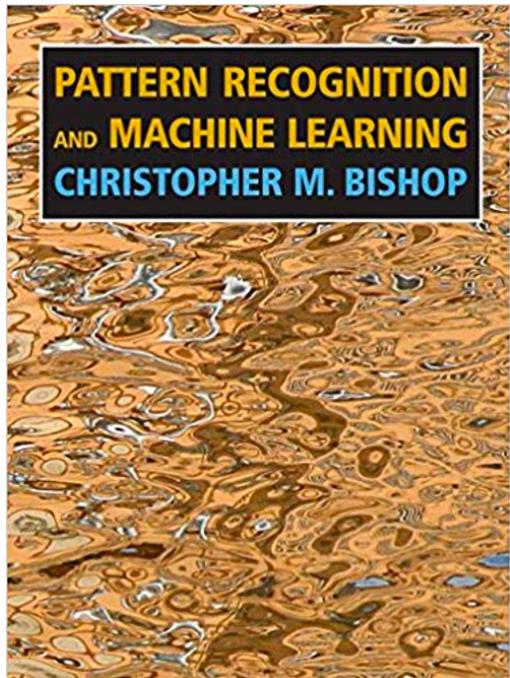
**2 Discrete Random Variables and Expectation**

- 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><b>CONTENTS</b></p> <p><b>2 Probability Distributions</b> <span style="float: right;">67</span></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><b>3 Linear Models for Regression</b> <span style="float: right;">137</span></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><b>Copyrighted Material</b></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	 		<b>3 Linear Models for Regression</b> <span style="float: right;">137</span>		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																																																												
<b>3 Linear Models for Regression</b> <span style="float: right;">137</span>																																																													
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?
- In the year of 2022, 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