

## Lecture 0: Introduction

Yi, Yung (이용)

EE210: Probability and Introductory Random Processes  
KAIST EE

MONTH DAY, 2021

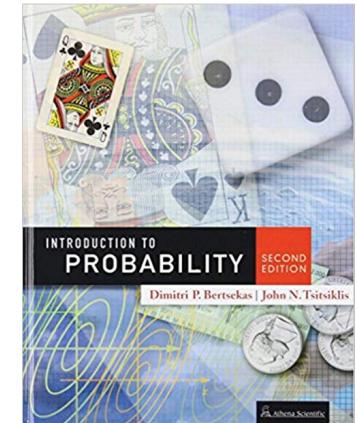
- Course logistics
- Why necessary to take the course of probability and random process?

Instructor

- Yi, Yung (이용)
- Office: N1, 810
- <http://lanada.kaist.ac.kr>, [yiyung@kaist.edu](mailto: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](mailto: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



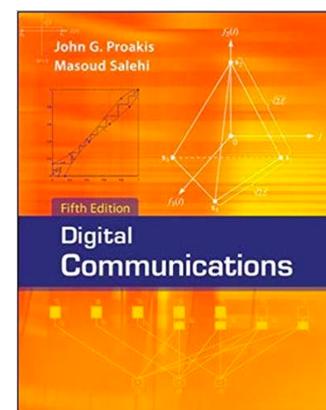
- Three Parts
  - Part I: Fundamentals of Probability
  - Part II: Inference and Limit Theorems
  - Part III: Random Processes
- On-line lectures at MIT and EdX
  - MIT: <http://bit.ly/2PkvYdr>
  - EdX: <http://bit.ly/3pHmZRd>
  - You can find older urls (2006, 2010, 2013) for this lecture, where there 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.

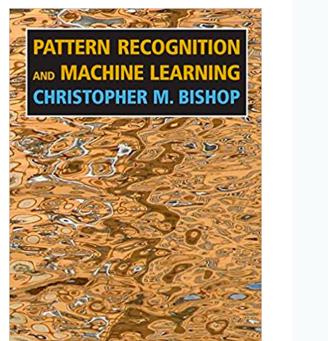
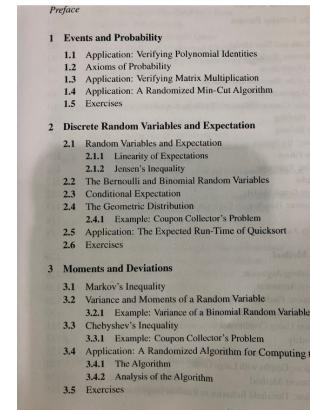
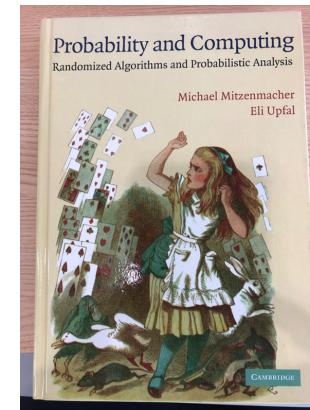
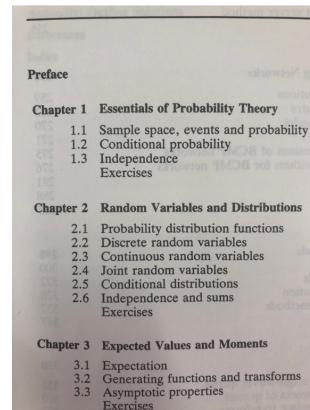
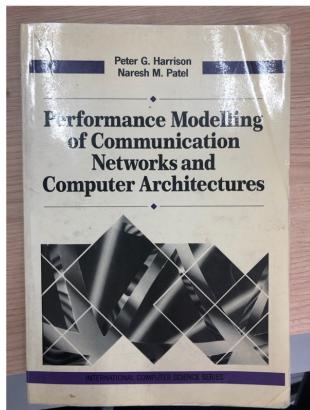
- Most should be via KLMS
  - Technical questions about lectures, homework, and etc
- Please DO NOT individually send emails to Prof. Yung Yi and TAs (or making calls or sending KakaoTalk msgs) about the technical questions (course contents, homework, etc)
  - All the questions need to be shared among 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"
- 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<br>1-1 Overview of the Book<br>1-2 Bibliographical Notes and References<br><br><b>2 Probability and Stochastic Processes</b><br>2-1 Probability<br>2-1-1 Random Variables, Probability Distributions, and Probability Densities<br>2-1-2 Functions of Random Variables<br>2-1-3 Statistical Averages of Random Variables<br>2-1-4 Some Useful Probability Distributions<br>2-1-5 Upper Bounds on the Tail Probability<br>2-1-6 Sums of Random Variables and the Central Limit Theorem<br><br>2-2 Stochastic Processes<br>2-2-1 Statistical Averages<br>2-2-2 Power Density Spectrum<br>2-2-3 Response of a Linear Time-Invariant System to a Random Input Signal<br>2-2-4 Sampling Theorem for Band-Limited Stochastic Processes<br>2-2-5 Discrete-Time Stochastic Signals and Systems<br>2-2-6 Cyclostationary Processes<br><br>2-3 Bibliographical Notes and References<br>Problems | 13<br>16<br>16<br><br>17<br>17<br>22<br>28<br>33<br>37<br>53<br><br>58<br>62<br>64<br>67<br><br>68<br>72<br>74<br>75<br>77<br>77 |
|--|--|



|  |     |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
|--|-----|----------|----------------------|-----------------------------|----|----------------------|----|-----------------------------|----|---------------------------|----|----------------------------------|----|-----------------------------|----|--|----|---------------------------------------|----|---|----|---|----|----------------------------|----|---|----|--------------------------------|-----|------------------------------|-----|----------------------------|-----|----------------------------|-----|--|-----|------------------------|-----|-----------------------------|-----|---------------------------|-----|---------------------------------|-----|---------------------------------|-----|-----------|-----|--------------------------------|--|----------------------------|-----|--|-----|---------------------------------|-----|---------------------------|-----|---------------------------------|-----|
| Copyrighted Material   | xiv | CONTENTS | Copyrighted Material |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
|  |     |          | xiii                 |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| <table border="0"> <tr> <td style="vertical-align: top; padding-right: 10px;">2 Probability Distributions</td> <td>67</td> </tr> <tr> <td>2.1 Binary Variables</td> <td>68</td> </tr> <tr> <td>2.1.1 The beta distribution</td> <td>71</td> </tr> <tr> <td>2.2 Multinomial Variables</td> <td>74</td> </tr> <tr> <td>2.2.1 The Dirichlet distribution</td> <td>76</td> </tr> <tr> <td>2.3 The Normal Distribution</td> <td>78</td> </tr> <tr> <td>2.3.1 Conditional Gaussian distributions</td> <td>85</td> </tr> <tr> <td>2.3.2 Marginal Gaussian distributions</td> <td>88</td> </tr> <tr> <td>2.3.3 Bayes' theorem for Gaussian variables</td> <td>90</td> </tr> <tr> <td>2.3.4 Maximum likelihood for the Gaussian</td> <td>93</td> </tr> <tr> <td>2.3.5 Sequential estimator</td> <td>94</td> </tr> <tr> <td>2.3.6 Bayesian inference for the Gaussian</td> <td>97</td> </tr> <tr> <td>2.3.7 Student's t-distribution</td> <td>102</td> </tr> <tr> <td>2.3.8 Periodic distributions</td> <td>105</td> </tr> <tr> <td>2.3.9 Mixture of Gaussians</td> <td>110</td> </tr> <tr> <td>2.4 The Exponential Family</td> <td>113</td> </tr> <tr> <td>2.4.1 Maximum likelihood and sufficient statistics</td> <td>116</td> </tr> <tr> <td>2.4.2 Conjugate priors</td> <td>117</td> </tr> <tr> <td>2.4.3 Noninformative priors</td> <td>117</td> </tr> <tr> <td>2.5 Nonparametric Methods</td> <td>120</td> </tr> <tr> <td>2.5.1 Kernel density estimators</td> <td>122</td> </tr> <tr> <td>2.5.2 Nearest-neighbour methods</td> <td>124</td> </tr> <tr> <td>Exercises</td> <td>127</td> </tr> <tr> <td colspan="2">3 Linear Models for Regression</td> </tr> <tr> <td>3.1 Linear Function Models</td> <td>137</td> </tr> <tr> <td>3.1.1 Maximum likelihood and least squares</td> <td>138</td> </tr> <tr> <td>3.1.2 Geometry of least squares</td> <td>140</td> </tr> <tr> <td>3.1.3 Sequential learning</td> <td>143</td> </tr> <tr> <td>3.1.4 Regularized least squares</td> <td>144</td> </tr> </table> |     |          |                      | 2 Probability Distributions | 67 | 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 Normal 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 estimator | 94 | 2.3.6 Bayesian inference for the Gaussian | 97 | 2.3.7 Student's t-distribution | 102 | 2.3.8 Periodic distributions | 105 | 2.3.9 Mixture 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 |  | 3.1 Linear Function Models | 137 | 3.1.1 Maximum likelihood and least squares | 138 | 3.1.2 Geometry of least squares | 140 | 3.1.3 Sequential learning | 143 | 3.1.4 Regularized least squares | 144 |
| 2 Probability Distributions  | 67  |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 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 Normal 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 estimator   | 94  |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 2.3.6 Bayesian inference for the Gaussian  | 97  |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 2.3.7 Student's t-distribution   | 102 |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 2.3.8 Periodic distributions   | 105 |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 2.3.9 Mixture 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   |     |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 3.1 Linear Function Models   | 137 |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 3.1.1 Maximum likelihood and least squares   | 138 |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 3.1.2 Geometry of least squares  | 140 |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 3.1.3 Sequential learning  | 143 |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |
| 3.1.4 Regularized least squares  | 144 |          |                      |                             |    |                      |    |                             |    |                           |    |                                  |    |                             |    |  |    |                                       |    |   |    |   |    |                            |    |   |    |                                |     |                              |     |                            |     |                            |     |  |     |                        |     |                             |     |                           |     |                                 |     |                                 |     |           |     |                                |  |                            |     |  |     |                                 |     |                           |     |                                 |     |

- 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!

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

Questions?

