

*INTRODUCTION TO PROBABILITY AND STATISTICS FOR MACHINE LEARNING*

CREDIT	3	INSTRUCTOR	Taeyoung Park
OFFICE	413 Daewoo Hall	OFFICE HOURS	
TIME	TBA	CLASSROOM LOCATION	TBA
E-MAIL	tpark@yonsei.ac.kr		

**[COURSE INFORMATION]**

<b>COURSE DESCRIPTION &amp; GOALS</b>	<p>Have you ever heard of Big Data or AI? What about machine learning?</p> <p>Big Data refers to large-scale data that are quickly generated in our daily activities with various types. Although we might not realize how much Big Data affects our everyday lives, the use of Big Data has become more common and essential for better decision making and prediction. Tools for analyzing Big Data are called machine learning (or statistical learning), which is the workhorse of artificial intelligence (AI). The most machine learning algorithms are constructed on the basis of probability and statistics, so without the knowledge of probability and statistics, we cannot fully understand how the machine learning algorithms such as Deep Learning are working.</p> <p>This course is designed to provide basic probability concepts and statistical inference, which are fundamental statistical principles behind machine learning. Topics include conditional probability, Bayes' theorem, random variables, probability distributions, expectations, law of large numbers, convergence, limiting distributions, central limit theorem, point and interval estimation, and hypothesis testing, as well as computer simulations to illustrate probability and statistics concepts. This course is also known as a mathematical statistics course.</p>
<b>PREREQUISITE</b>	Basic calculus and statistics courses are highly recommended.
<b>COURSE REQUIREMENTS</b>	PC for computer simulations and illustrations
<b>GRADING POLICY</b>	<ol style="list-style-type: none"> <li>1. Midterm and Final Exams (30% each) <ul style="list-style-type: none"> <li>- On-line midterm and final exams are to be held, and practice exam questions with solution keys are given for your preparation.</li> <li>- Exams consist of multiple-choice, matching, and short-answer questions.</li> </ul> </li> </ol>

	<p>2. Pop-up Quizzes (20%)</p> <ul style="list-style-type: none"> <li>- There are more than 10 pop-up quizzes each of which amounts to 2% toward your final letter grade.</li> <li>- Pop-up quizzes will be from lecture videos, and you should be able to answer them based on what you learned by simply watching videos</li> <li>- The purpose of the pop-up quiz is to check whether you understand materials in lecture videos.</li> <li>- Percentage points above 20% will be given as your bonus points. For example, if there are 13 pop-up quizzes and you got the correct answers for all of them, you will receive 6% bonus points.</li> </ul> <p>3. Attendance (20%)</p> <ul style="list-style-type: none"> <li>- Lecture videos will be posted on YSCEC every Monday, and you are supposed to watch videos within a one-week timeframe after posting (Monday through Sunday).</li> <li>- View progress is automatically tracked by the system, and you can check your view progress rate as well.</li> <li>- Percentage points for attendance will be given by the overall view progress rate.</li> </ul> <p>For example, if your overall view progress rate is 80%, the percentage points for attendance will be <math>20 \times 0.8 = 16\%</math> points.</p>
<p><b>TEXTS &amp; NOTES</b></p>	<ul style="list-style-type: none"> <li>- Course materials including lecture notes, previous exams with solution keys, and problem sets with solutions are given.</li> <li>- You are not required to submit your answers to problem sets.</li> </ul>
<p><b>INSTRUCTOR'S PROFILE</b></p>	<p>Taeyoung Park, PhD</p> <ul style="list-style-type: none"> <li>- Positions <ul style="list-style-type: none"> <li>• Professor, Department of Applied Statistics, Yonsei University</li> <li>• Professor, Department of Statistics and Data Science, Yonsei University</li> <li>• Associate Dean, Office of University Planning and Development, Yonsei University</li> <li>• Consulting Professor, AmorePacific Corporation</li> <li>• Consulting Professor, Bank of Korea</li> <li>• Advisory Professor, Data Science Lab at Yonsei, Yonsei University</li> </ul> </li> <li>- Academic Experience <ul style="list-style-type: none"> <li>• Assistant Professor, Department of Statistics, University of Pittsburgh, USA</li> </ul> </li> <li>- Education <ul style="list-style-type: none"> <li>• B.A. in Applied Statistics, Yonsei University</li> </ul> </li> </ul>

	<ul style="list-style-type: none"> <li>• A.M. and Ph.D. in Statistics, Harvard University, USA</li> <li>- Awards</li> <li>• Outstanding Achievement Award in Teaching, Yonsei University (2018, 2019)</li> <li>• Best Teaching Award, Yonsei University (2011, 2012, 2013, 2017, 2018)</li> <li>• Cho-Heon Research Excellence Award (2013)</li> <li>• Certificate of Distinction in Teaching, Harvard University (2002, 2003, 2004, 2005)</li> </ul>
--	---

**[WEEKLY SCHEDULE]**

WEEK (PERIOD)	WEEKLY TOPIC & CONTENTS	COURSE MATERIAL & ASSIGNMENTS	NOTES
1	<ul style="list-style-type: none"> <li>• 1.1 Introduction</li> <li>• 1.2 Sets</li> <li>• 1.3 Probability set function</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• Computer simulation</li> <li>• 1.3.1 Counting rules</li> <li>• 1.4 Conditional probability</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• 1.5 Discrete random variables</li> <li>• 1.6 Continuous random variables</li> <li>• 1.7 Expectation of random variables</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• FAQ</li> <li>• 1.8 Some special expectations</li> <li>• 1.9 Important inequalities</li> <li>• 2.1 Several random variables</li> </ul>	Readings, problem sets, pop-up quiz	
2	<ul style="list-style-type: none"> <li>• 2.2 Transformations</li> <li>• FAQ</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• 2.3 Conditional distributions</li> <li>• 2.4 Correlation coefficients</li> <li>• 2.5 Independent random variables</li> <li>• FAQ</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• 2.6 Linear combinations</li> <li>• 3.1.1 Binomial distribution</li> </ul>	Readings, problem sets, pop-up quiz	

WEEK (PERIOD)	WEEKLY TOPIC & CONTENTS	COURSE MATERIAL & ASSIGNMENTS	NOTES
	Midterm exam		
3	<ul style="list-style-type: none"> <li>• 3.1.2 Poisson distribution</li> <li>• 3.2.1 Normal distribution</li> <li>• 3.2.2 Multivariate normal distribution</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• 3.2.3-1 Gamma distribution</li> <li>• 3.2.3-2 Chi-squared distribution</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• 3.2.3-3 Beta distribution</li> <li>• 3.2.4-1 t distribution</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• 3.2.4-2 F distribution</li> <li>• 3.3 Mixture distributions</li> <li>• 3.4 Distributions of order statistics</li> </ul>	Readings, problem sets, pop-up quiz	
4	<ul style="list-style-type: none"> <li>• 4.1 Sampling and Statistics</li> <li>• 4.2 Confidence intervals</li> <li>• 4.3 Hypothesis testing</li> </ul>	Readings, problem sets, pop-up quiz	
	<ul style="list-style-type: none"> <li>• 5.1 Convergence in probability</li> <li>• 5.2 Convergence in distribution</li> <li>• 5.3 Central limit theorem</li> </ul>	Readings, problem sets, pop-up quiz	
	Final exam		