PROBABILITY SYLLABUS, Spring Semester 2022
Budapest Semesters in Mathematics
Mo 10:15am - noon, Room 106
Tu 8:15am - 10:00, Room 106

Péter Bálint
E-mail: pet@math.bme.hu
Office Hours: Mo 9:15am - 10:00, Room 106
(or online, by appointment)
Text is: *A First Course in Probability* by S. Ross
(see also lecture notes available from the course webpage)
Course homepage: https://math.bme.hu/~pet/pro/

Course Description: This is a first course on the mathematical phenomenon of uncertainty and techniques used to handle them. Not only being challenging itself, this field is of increasing interest in many areas of engineering, economical, physical, biological and sociological sciences as well. In this course we cover the basic notions and methods of probability theory, also giving emphasize on examples, applications and problem solving. Briefly, the topics include probability in discrete sample spaces, methods of enumeration (combinatorics), conditional probability and independence, random variables, properties of expectations, the Weak Law of Large Numbers, and the Central Limit Theorem.

Probability is a conceptually difficult field, although it might seem easy and straightforward at first. One has to distinguish between very different mathematical objects, and find their connection to real-life situations within the same problem. Therefore it is very important to follow classes and deeply understand the material during the semester.

Grading and assignments: There will be two in-class exams, weekly homeworks to be handed in during the semester, and a final exam.

- **Two in-class exams** are to be scheduled in due time. The first exam will be on Chapters 1, 2 and 3, the second one will be on Chapters 4, 5 and 6. **Each** worth **160 points** (each 20% of the total possible points).
- **13 homework sets** are to be handed in during the semester. **Each** worth **20 points**, the worst of all homeworks will be dropped. This way, a **total** of **240 points** (30% of the total possible points) can be earned from these assignments. Solving the homework problems by no means guarantees that you have the necessary level of practice. Please do other exercises (and check the answers in the back of the book *after* solving them) until you feel safe with problems on the topics in question. It is a good idea to simulate exam-like situations: solve exercises in limited time, without the use of the book or your notes (or your classmates).
- **The final exam** is to be scheduled in due time. Half of it will cover Chapters 1 to 6, the other half is on Chapters 7 and 8 of the book. It is worth **240 points** (30 % of the total possible points).
- **Bonus questions** are also to be found in the homework sets. While a total of 800 points can be earned by the exams and homeworks, an additional **4 points** can be given for a solution of each bonus problem.

Grades will be based on the total of 800 points approximating the following standards:

Grade	**Points**
A^+	≥ 775
A	$\in [745, 775)$
A^-	$\in [720, 745)$
B^+	$\in [695, 720)$
B	$\in [665, 695)$
B^-	$\in [640, 665)$
C^+	$\in [615, 640)$
C	$\in [585, 615)$
C^-	$\in [560, 585)$
D	$\in [480, 560)$
F	< 480

Because of this standard, you are not in competition with your classmates nor does their performance influence positively or negatively your performance. You are encouraged to form study/problem groups with your classmates; things not clear to you may become obvious when you try to explain them to others or when you hear other points of view. Sometimes just verbalizing your mathematical thoughts can deepen your understanding. However, if you discuss with others the exercises, each person should write up her/his own version of the solution. Please note that much less can be learned by just understanding and writing up someone else's solution than by coming up (or even just trying to come up) with original ideas and solving the problem.

Please feel free to contact me any time outside class via e-mail if you have questions or suggestions about this course.

A tentative course schedule for the first couple of weeks with the already assigned homework problem sets follows on the next pages. The content of the syllabus is continuously updated as more information are available.

Tentative Schedule for BSM Probability, Spring 2022

Date	Chapter of the book	Homework due on the date to the left
Feb 7 Mo	1.1, 1.2, 1.3, 1.4, 1.5, 1.6	-
Feb 8 Tu	2.2, 2.3, 2.4, 2.5	-
Feb 14 Mo	2.5, 2.6	-
Feb 15 Tu	3.2, 3.3	Homework #1
Feb 21 Mo	3.3, 3.4	-
Feb 22 Tu	3.4, 3.5	Homework #2
Feb 28 Mo	3.5, 4.1, 4.2	-
Mar 1 Tu	4.3, 4.4, 4.5, 4.6	Homework #3
Mar 7 Mo	4.6, 4.7	-
Mar 8 Tu	4.7, 4.8	Homework #4
Mar 14 Mo	*National Holiday*	-
Mar 15 Tu	*National Holiday*	Homework #5
Mar 21 Mo	4.8, 4.9, 4.10	-
Mar 21 Mo, 5pm	**Consultation**, Room 104	-
Mar 22 Tu	**Midterm 1**	-
Mar 28 Mo	4.10, 5.1, 5.2	-
Mar 29 Tu	5.3, 5.4	Homework #6
Mar 31 Th, **MUC**	5.4, 5.5	-
Apr 4 Mo	5.5, 5.7	-
Apr 5 Tu	6.1, 6.2	Homework #7
Apr 11 Mo	*Spring break*	-
Apr 12 Tu	*Spring break*	-
Apr 19 Tu	6.2, 6.3	-
Apr 21 Th, **MUC**	6.3, 6.4	Homework #8, Homework #9
Apr 25 Mo	*Class canceled*	-
Apr 26 Tu	6.4, 6.5, 7.1, 7.2	-
Apr 28 Th, **MUC**	**Midterm 2**	-
May 2 Mo	7.2, 7.3, 7.4	-
May 3 Tu	7.4, 7.5	Homework #10
May 9 Mo	7.6, 7.7	-
May 10 Tu	7.8, 8.2	Homework #11
May 16 Mo	8.3	-
May 17 Tu	8.4	Homework #12
May 20 Fr, 4pm	**Consultation**	Homework #13
May 23 Mo	**Final Exam**	-

Homework problem sets for BSM Probability, Spring 2022

Group work is encouraged, but write up your own solution. Please *show your work leading to the result, not only the result.* Problems are either from the book or written here explicitly. Please make sure you solve the problem indicated here, and not another one (the one below or above it, or the problem with the same number but from another chapter or from the other edition, etc.). *Numbers refer to the 8th edition of the book. If you have another edition, let me know to check what the relevant numbers are.* Each problem is for the number of •'s you see right next to it. Hence, for example, Problem 4 of Chapter 1 is for two points, while Theoretical Exercise 10 of Chapter 1 is for three points. Bonus problems are also to be found here, for 4 points each. Note that they are also due on the due date of the corresponding homework.

Homework #1 (Due on February 15)

Chapter 1, Problem $4^{\bullet\bullet}$, $7^{\bullet\bullet\bullet}$, $15^{\bullet\bullet\bullet}$, $19^{\bullet\bullet\bullet}$, $21^{\bullet\bullet\bullet}$, $31^{\bullet\bullet\bullet}$,
Chapter 1, Theoretical Exercise $10^{\bullet\bullet\bullet}$

Homework #2 (Due on February 22)

Chapter 2, Problem $13^{\bullet\bullet}$, $17^{\bullet\bullet}$, $21^{\bullet\bullet}$, $23^{\bullet\bullet}$, $27^{\bullet\bullet}$, $33^{\bullet\bullet}$, $44^{\bullet\bullet}$,
Problem #2A$^{\bullet\bullet\bullet\bullet}$: We roll a die ten times. What is the probability that each of the results $1, 2, \dots, 6$ shows up at least once? HINT: Define the events $A_i := \{\text{number } i \text{ doesn't show up at all during the ten rolls}\}$, $i = 1 \dots 6$. Note that these events are *not* mutually exclusive.
Problem #2B$^{\bullet\bullet}$: For the events A and B, we know that $P(A) \geq 0.8$ and $P(B) \geq 0.5$. Show that $P(A \cap B) \geq 0.3$.
Bonus problem #2C: A closet contains n pairs of shoes. If $2r$ shoes are randomly selected ($2r \leq n$), what is the probability that there will be (a) no complete pair, (b) exactly one complete pair, (c) exactly two complete pairs?

Homework #3 (Due on March 1)

Chapter 3, Problem $7^{\bullet\bullet\bullet}$ (HINT: it is not $\frac{1}{2}$.), $12^{\bullet\bullet\bullet\bullet}$, $44^{\bullet\bullet\bullet}$, $50^{\bullet\bullet\bullet}$,
Problem #3A$^{\bullet\bullet\bullet}$: We repeatedly roll two dice at the same time, and only stop when at least one of them shows a six. What is the probability that the other also shows a six? (HINT: it is not $\frac{1}{6}$).
Problem #3B$^{\bullet\bullet\bullet\bullet}$: Eggs are sold in boxes containing 10 eggs at the farmer's market. In 60% of these boxes all the eggs are unbroken, in 30% of the boxes there is exactly one broken egg, in 10% of the boxes there are exactly two broken eggs. I buy a box of eggs at the farmer's market, at home I take an egg out of it and I realize that it is broken. What is the chance that there is another broken egg in this box?
Bonus problem #3C: Consider a group of n people, and let A_{ij} denote the event that person $\#i$ and person $\#j$ have a common birthday. (a) Are these $\binom{n}{2}$ events *pairwise* independent? (b) Are these $\binom{n}{2}$ events independent as a collection?

Homework #4 (Due on March 8)

Chapter 3, Problem $74^{\bullet\bullet\bullet}$; Theoretical Exercise $22^{\bullet\bullet\bullet}$;
Problem #4A$^{\bullet\bullet\bullet}$: Die α has four red and two white faces, while die β has two red and four white faces. We flip a fair coin. If it comes head then we use die α, if it comes tail, then we use die β. We then roll the die selected this way n times. What is the probability that the k-th roll will be red, given that all previous rolls were red ($k = 1, 2, \dots, n$)?
Problem #4B$^{\bullet\bullet}$: Andrew and Bob play the following game. There are 5 red balls and 5 blue balls in a urn, out of which two balls are drawn. If the two balls drawn are of the same color, Andrew pays Bob 10\$, if the two balls are of two different colors, Bob plays Andrew x\$. How much is x if the game is fair?
Problem #4C$^{\bullet\bullet\bullet}$: Numbers 1, ..., 5 are randomly distributed among the five players A, B, C, D and E, who play the following game. The first match is between A and B, the one who has the greater number proceeds and plays the second match with C, then the winner of the second match plays against D, and so on. Let X denote the number of matches won by A. Determine the probability mass function of the random variable X.
Problem #4D$^{\bullet\bullet\bullet}$: Let $S = \{1, 2, \dots, n\}$, and suppose that A and B are, independently, equally likely to be any of the 2^n subsets (including the null set and S itself) of S.

(a) Show that $P\{A \subset B\} = \left(\frac{3}{4}\right)^n$.

(b) Show that $P\{A \cap B = \emptyset\} = \left(\frac{3}{4}\right)^n$.

Problem #4E$^{\bullet\bullet\bullet}$: Cities A, B, C, D are located (in this order) on the four corners of a square. Between them, we have the following roads: $A \leftrightarrow B$, $B \leftrightarrow C$, $C \leftrightarrow D$, $D \leftrightarrow A$, $B \leftrightarrow D$. One night each of these roads gets blocked by the snow independently with probability 1/2. Show that the next morning city C is accessible from city A with probability 1/2.
Bonus Problem #4F: Recall Pólya's urn model: initially, there are two balls in the urn, one blue ball and one red ball. At each step, a ball is drawn, its color is observed, and, in addition to the ball drawn, another ball of the same color is put into the urn. Determine the distribution of the number of red balls after n steps, that is, the probability that there are exactly k red balls in the urn ($k = 1, 2, \ldots, n, n+1$).

Homework #5 (Due on March 15)

Chapter 4, Problem $4^{\bullet\bullet\bullet}$, $21^{\bullet\bullet}$, $26^{\bullet\bullet}$, $30^{\bullet\bullet}$, $37^{\bullet\bullet}$, $38^{\bullet\bullet\bullet}$,
Problem #5A$^{\bullet\bullet\bullet}$: Let N be a random variable that takes non-negative integer values. Show that

$$\mathbb{E}(N) = \sum_{i=1}^{\infty} \mathbb{P}(N \geq i).$$

(*Hint:* $\sum_{i=1}^{\infty} \mathbb{P}(N \geq i) = \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} \mathbb{P}(N = k)$. Now interchange the order of the summation.)
Problem #5B$^{\bullet\bullet\bullet}$: A man has n keys, out of which there is just one that opens a specific lock. The man keeps trying the keys until he finds the right one. Compute the expected number of trials in both of the following cases: (a) he puts away the keys he has tried and turned out to be wrong, and at each trial he picks uniformly from the remaining ones (sampling without replacement); (b) at each trial, he picks one uniformly from all the n keys (sampling with replacement).

Homework #6 (Due on March 29)

Chapter 4, Problem $46^{\bullet\bullet\bullet}$, $50^{\bullet\bullet\bullet}$, $51^{\bullet\bullet}$, $56^{\bullet\bullet\bullet}$,
Chapter 4, Theoretical Exercise $16^{\bullet\bullet\bullet}$, $25^{\bullet\bullet\bullet}$
Problem #6A$^{\bullet\bullet\bullet}$: A book was printed in poor quality, it has 100 pages with exactly one typo, and 60 pages with exactly two typos. Provide an estimate on the total number of pages in the book.
Bonus problem #6B: The average density of a forest is 16 trees on every 100 square yards. The tree trunks can be considered as cylinders of a diameter of 0.2 yards. We are standing inside the forest, 120 yards from its edge. If we shoot a gun bullet out of the forest without aiming, what is the probability that it will hit a tree trunk? (Ignore the marginal fact that centers of the tree trunks cannot be closer than 0.2 yards to each other.)

Homework #7 (Due on April 5)

Chapter 4, Problem $19^{\bullet\bullet}$, $75^{\bullet\bullet\bullet}$, $78^{\bullet\bullet\bullet}$,
Chapter 5, Problem $2^{\bullet\bullet}$, $3^{\bullet\bullet}$, $5^{\bullet\bullet}$, $6^{\bullet\bullet\bullet}$,
Problem #7A$^{\bullet\bullet\bullet}$: For which values of α and c will the function $F(x) = \exp(-ce^{-\alpha x})$ be a distribution function? For such values, what is the corresponding density?
Bonus Problem #7B: Two players target shoot, player one begins. They shoot by turns, and hit the target with probability p_1 and p_2, respectively. The winner is the one who hits the target first. What is the probability that the first player wins? What is the expected number of shots in this game? (The expected number of shots is easy to compute in case $p_1 = p_2 = p$ (why?). Compute it, and verify if it agrees with your result when plugging in $p_1 = p,\ p_2 = p$.)

Homework #8 (Due on April 21)

Chapter 5, Problem $10^{\bullet\bullet}$, $12^{\bullet\bullet\bullet}$ (If you are interested, try to find the optimal solution (no points for this, but it's fun).), $18^{\bullet\bullet}$, $21^{\bullet\bullet}$, $26^{\bullet\bullet\bullet}$, $29^{\bullet\bullet\bullet}$,
Problem #8A$^{\bullet\bullet}$: A stick of length l is broken randomly. What is the distribution function of the length of the shorter piece?
Problem #8B$^{\bullet\bullet\bullet}$: A fraction p of citizens in a city smoke. We are to determine the value of p by making a survey involving n citizens who we select randomly. If k of these n people smoke, then $p' = k/n$ will be our result. How large should we choose n if we want our result p' to be closer to the real value p than 0.02 with probability at least 0.97? In other words: determine the smallest number n_0 such that $P(|p' - p| \le 0.02) \ge 0.97$ for any $p \in (0, 1)$ and $n \ge n_0$.
Bonus Problem #8C: Compute the "absolute moments"

$$\mathbb{E}(|Y|^k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |y|^k e^{-y^2/2} dy$$

of the standard normal distribution.
HINT: for even k's, compute and use

$$\frac{d^l}{d\lambda^l} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\lambda y^2/2} dy \bigg|_{\lambda=1}.$$

For odd k's, introduce the new variable $z = y^2$ in the integral.

Homework #9 (Due on April 21)

Chapter 5, Problem $41^{\bullet\bullet}$,
Chapter 6, Problem $1^{\bullet\bullet\bullet}$, $16^{\bullet\bullet\bullet}$,
Chapter 5, Theoretical Exercise $13^{\bullet\bullet}$, $15^{\bullet\bullet}$, $29^{\bullet\bullet}$, $31^{\bullet\bullet}$ and for the last one also$^{\bullet}$: show here (without much computation) that CY^α has also lognormal distribution with parameters $\mu' = \alpha\mu + \log C$ and $\sigma'^2 = \alpha^2\sigma^2$. HINT: As we have $Y = e^X$ with X normal, write CY^α as e^Z, and figure out the connection of Z to X. What kind of variable is Z? With which parameters?
Problem #9A$^{\bullet\bullet\bullet}$: We randomly select a point on the [0, 1] interval of the x axis. Let ξ denote the distance between this point and the point at coordinate (0, 1) of the plane. Determine the density function of the distribution of the random variable ξ.

Homework #10 (Due on May 3)

Chapter 6, Problem $12^{\bullet\bullet}$, $13^{\bullet\bullet\bullet}$, $20^{\bullet\bullet}$, $52^{\bullet\bullet\bullet}$ HINT: Write up the definition of the joint distribution function of R and Θ, which is in fact the probability of some set of points. Make a good picture of that set to determine its probability, i.e. the joint distribution function. Then find the density.
Chapter 6, Theoretical Exercise $9^{\bullet\bullet}$, $14^{\bullet\bullet\bullet}$
Problem #10A$^{\bullet\bullet}$: Choose a random point in the unit square of the plane. Let ξ be its distance from the closest edge of the square. Determine the distribution function of ξ. HINT: Draw!
Problem #10B$^{\bullet\bullet\bullet}$: A man takes the train and then transfers to the bus when commuting to work each day. It takes two minutes for him to walk from the train to the bus at the transfer station. *In principle* the train arrives at 7:30am, and the bus leaves at 7:37am. The fact is that the train arrives at a normally distributed random time, having mean 7:30am and standard deviation 4 minutes. Independently, the bus leaves at a normally distributed random time with mean 7:37am and standard deviation 3 minutes. What is the probability that our man misses his bus at most once on the five working days of the week? HINT: Use what we know about the

sum of normal variables.
Bonus Problem #10C: Let X and Y be i.i.d. geometric random variables, both with parameter p. Define $U := \min\{X, Y\}$ and $V := X - Y$. Show that U and V are independent.

Homework #11 (Due on May 10)

Chapter 6, Problem 39$^{\bullet\bullet}$, 42$^{\bullet\bullet}$
Chapter 7, Problem 5$^{\bullet\bullet}$, 8$^{\bullet\bullet\bullet}$, 12$^{\bullet\bullet\bullet}$, 18$^{\bullet\bullet}$, 21$^{\bullet\bullet\bullet}$ (Number of distinct birthdays = number of days with at least one birthday.),
Chapter 7, Theoretical Exercise 9$^{\bullet\bullet\bullet}$

Homework #12 (Due on May 17)

Chapter 7, Problem 26$^{\bullet\bullet\bullet\bullet}$ HINT: Find the distribution function of max or $\min(X_1, \ldots, X_n)$. 30$^{\bullet\bullet\bullet}$, 40$^{\bullet\bullet}$, 50$^{\bullet\bullet}$, 47$^{\bullet\bullet\bullet\bullet}$ (This is the so-called *Erdős-Rényi* random graph.) HINT: Use indicator variables for the edges.
Chapter 7, Theoretical Exercise 13$^{\bullet\bullet\bullet}$ HINT: Define the indicator variable I_i to be one if X_i happens to be a record value, and use the fact that each of X_1, X_2, …, X_i is equally likely to be the largest among themselves.
Problem #12A$^{\bullet\bullet}$: In my purse the number of pennies, nickels, dimes and quarters are i.i.d. Poisson random variables with parameter λ. Use linearity of expectations to compute the expectation and variance of the joint value of my coins.

Homework #13 (Due on May 20)

Chapter 7, Problem 68$^{\bullet\bullet}$, 73$^{\bullet\bullet\bullet}$ (Example 6b is the transmitted value, received value problem $(S \sim \mathcal{N}(\mu, \sigma^2); R|S \sim \mathcal{N}(S, 1))$ from class), 76$^{\bullet\bullet}$, 78$^{\bullet\bullet\bullet}$
Chapter 7, Theoretical Exercise 29$^{\bullet\bullet}$ HINT: Use symmetry and the fact that (finite) summation and conditional expectation (with the same condition) are interchangeable. 41$^{\bullet\bullet\bullet}$,
Problem #13A:$^{\bullet\bullet\bullet}$ The random variables X and Y have a jointly absolutely continuous distribution, the (marginal) density of X is $f_X(x) = \lambda^2 x e^{-\lambda x}$ if $x \geq 0$, and 0 for $x < 0$, where $\lambda = 1/2$; while the conditional distribution of Y given X is uniform on the interval $[0, X]$. (Interpretation: X is the lifetime of a light bulb measured in years, while Y is the time when it starts to twinkle.) Determine: (a) the joint density (do not forget to specify the domain where it is nonzero), (b) the marginal density of Y and $\mathbb{E}(Y)$, (c) $\mathbb{E}(Y|X = x)$ and $\mathbb{E}(X|Y = y)$, and (d) $\mathbb{E}(X)$ (for (d) it simplifies the computation if you rely on (c)). (No points for this, but it is worth a thought: how does this problem relate to convolutions? And to the Poisson process?)
Problem #13B:$^{\bullet\bullet}$ According to statistical data, for Hungarian married couples the heights of the two individuals are *jointly normally distributed*: the height of the wife has mean 5 ft 4.5 in and standard deviation 4 in, while the height of the husband has mean 5 ft 9.5 in and standard deviation 5 in. The two quantities are *not independent*, in fact, their correlation coefficient is 0.5. Pick a Hungarian married couple at random, what is the chance that the woman is taller than the man? (FYI: 1 ft = 12 in.)

Relevant pages of the book are available here:
https://math.bme.hu/~pet/pro/HW9.pdf
https://math.bme.hu/~pet/pro/HW11.pdf