Probability distributions: Course manual

EBP038A05, 2021-2022

Nicky D. van Foreest

2022-10-18

1 Schedule

The schedule appears on top to facilitate look ups.

Table 1: schedule

Week	Lecture	Section	Video	Exercises of BH	Assignment
1	1	7.1	19	**7.1**, **7.9**, 7.11	1
		7.2	-	**7.13**, 7.15, **7.24**	
2	2	7.3	21	**7.29**, **7.38**, 7.53	2
		7.4, 7.5	20	7.58, **7.59**, 7.71	
3	3	8.1	22	**7.86**, **7.10**, **8.11**	3
		8.2	-	**8.15**, **8.18**, **8.23**	
4	4	8.3	23	8.31, 8.36, 8.40	4
		8.4, 8.5	24, 25	**8.52**, **8.54**	
5	5	9.1, 9.2	26	**9.1**, 9.25, 9.28, **9.32**	5
		9.2, 9.3	27 (until 0:30)	**9.37**, 9.39	
6	6	9.5, 9.6	27 (rest)	**9.50**, 9.52, 9.55, **9.57**	6
		10.1	28	**9.58**, **10.2**, 10.3, **10.6**	
7	7	10.2, 10.3	29	**10.9**, **10.23**, **10.26**	7
		10.3, 10.4	30	**10.28**, **10.30**, **10.36**, 10.39	

- The video column refers to the video's of Blitzstein.
- The **boldface** exercises will be explained below in the Section 'Obligatory parts and exercises of BH'.

2 Course goals

See Ocasys.

3 Material

1. We use Introduction to Probability by Blitzstein and Hwang (BH) as the main material. The book's homepage provides an electronic copy of the book, video lectures by Blitzstein, and

a solution manual for part of the problems. I advise you to buy the book as it makes studying easier, and a physical book is much more practical than an ebook when working on the material and solving problems.

2. `pdf-files/lectures-questions.pdf`: We will illustrate the material during the lectures by solving problems. Before our lectures, we expect you to have watched the relevant lectures of Blitzstein for overview and general background of the book. `lectures.pdf` contains the problems per lecture. We expect that you have read, and thought about, them. However, you don't have to solve the problems before the lecture.

3. `pdf-files/lectures-solutions.pdf`: We will make the solutions progressively available after the lecture. Our lectures will be given online. As all lecture material (plus solutions) are available, we will not record the lectures.

4. `pdf-files/study-guide.pdf`: many questions, hints and solutions. See its intro.

5. `pdf-files/assignments.pdf`: the material for the assignments. See its intro.

6. `pdf-files/course-overview.pdf`: this document.

All material, including the LaTeX\ sources, is available on Github.

4 Assignments

- See the intro of `assignments.pdf`
- Each group consists of two members. Each group turns in an LaTeX\ document with the answers for the assignment. You can choose your group member when we open the groups. We form random groups for the remaining students.
- The due dates are simple: turn in the assignment of week n on Wednesday 18.00h of week $n+1$, for $n = 1, \ldots, 7$.
- We provide a document `sources/assignment-answer-template.tex`, see Github, that you should use as starting point for your assignments.

5 Obligatory parts and exercises of BH

At the exam we will ask simple variations of the following examples and exercises of BH

- BH 7.3.8, 7.5.9
- BH 8.1.3 (proof), 8.1.4, 8.1.9, 8.2.4, 8.3.3
- BH 9.6.1, 9.6.2
- BH 10.1.10 (proof), 10.3.2 (proof), 10.3.4, 10.3.5, 10.3.6
- The exercises marked in **boldface** in Table schedule.

If you learn the solutions for these examples and exercises really well, then your changes of passing the exam will be much higher. Moreover, learning this material by heart will save you lots of time during your student career since you will encounter these topics and techniques many times after this course. Since you can prepare this, we will be very critical on the exam on your answers.

6 Exam

The exam will have the following format:

1. Two or three questions selected at random from (small variations of) the obligatory parts and exercises of BH. Weight: 40 %
2. A question based on the code of the assignments in which you have to explain how it works or you have to find small bugs and explain why these are bugs and how to correct them. Weight: 20 %
3. One question selected at random from the simple questions of the study guide. Weight: 10 %
4. One or two questions that are (larger) variations of the exercises of BH, the material of the study guide, and the lecture slides. Weight: 30 %

Like this, if you study hard on the code of the assignments and the obligatory parts and exercises of BH you can pass the course.

Some remarks:

- The exam is closed book and has a duration of 2h.
- We will plan a perusal after the grades of the exam (resit) have been communicated. This will be the only moment in which you can check your exam.
- For the following distributions, you have to know by heart the form and the parameters, and either learn (or be able to derive at the exam) the mean and variance: Bernoulli, Binomial, First success, Geometric, Poisson, Uniform (discrete and continuous) and Exponential.
- For the hypergeometric distribution you have to know the pmf and the parameters, but not the mean and variance. If necessary, we will provide the mean and variance at the exam.
- We will not ask any question that involves calculus (e.g., integration) with (the cdf or pmf of): Negative hyper geometric, Weibull, Log normal, Chi-square or Student-t.
- The exam contains a list only with distributions that you don't have to learn by heart. See the mock exam for the list.

7 Grading

Each of the seven assignments will be graded as 1, 4, 7, 9, or 10. All assignments have the same weight. If you skip the challenges, the highest grade you can get is a 9. If you don't turn in an

assignment, the grade will default to 1. For the assignment grade, if a_i is the grade of your ith assignment, then a is the number obtained after rounding $\sum_{i=1}^{7} a_i/7$ to one decimal.

For the exam (resit), we add up all the points you earned and divide that by the total amount of points of the exam. Round this to one decimal, and let this be e.

Then we compute your final grade g for the course like this (supposing for the moment that the number of points of the exam is 25): []python from sigfig import round

tot = 24

def grade(a, e): ga = round(sum(a) / len(a), sigfigs=2) ge = round(10 * e / tot, sigfigs=2) if ge < 5: g = max(ge, 1) elif ga >= 6: g = max(0.75 * ge + 0.25 * ga, ge) else: g = 0.75 * ge + 0.25 * ga final = round(g, sigfigs=1) print(f"a=, ga=, e=, ge=, final=") return final

some tests grade(a=[10, 10, 10, 10, 10, 10, 10], e=5) grade(a=[10, 10, 10, 10, 10, 10, 10], e=12) grade(a=[10, 10, 10, 10, 10, 10, 10], e=13) grade(a=[7, 7, 7, 7, 10, 9, 9], e=11) grade(a=[7, 7, 7, 7, 10, 1, 9], e=12) grade(a=[4, 4, 4, 4, 4, 1, 1], e=21) grade(a=[1, 1, 1, 1, 10, 1, 10], e=24) grade(a=[7, 7, 7, 8, 9, 10, 4], e=24) grade(a=[7, 7, 7, 7, 7, 7, 7], e=12)

Results of (the) previous year(s) do not carry over to this year. For instance, you have to the assignments again (if you want to obtain an assignment grade higher than 1).

8 Estimated work load

The estimated (!) work load is as in the table. The lectures of Blitzstein vary in length, hence the amount of hours spent on his lectures is just a rough estimate.

Activity	Load	Hours
Lectures by Blitzstein	14×1.5	21
Lectures by us	14×1	14
Tutorial	7×2	14
Study book	14×2	28
Assignments	7×3	21
Exercises	14×2.5	35
Exam		3
Total		136

9 Contact info

- Nicky van Foreest (coordinator, lecturer)
- Arpan Rijal (lecturer)
- Joost Doornbos (TA)
- Wietze Koops (TA)
- Machiel Kroon (TA)
- Mikael Makonnen (TA)

In case some parts of this course guide are unclear, please mail `n.d.van.foreest@rug.nl` or `a.rijal@rug.nl`.

10 Interesting other literature for now and later

There are a number of books that you might like too. (From experience I can tell that reading different types of explanation can be very helpful.)

1. R.B. Ash: Basic probability theory, free online
2. C.M. Grinsted and J. Laurie Snell: Introduction to probability, also free online
3. F.M. Dekking, et al.: A Modern Introduction to Probability and Statistics, Understanding Why and How.

After the course you might be interested in the following books that I liked a lot.

1. D.V. Lindley, Understanding Uncertainty. This book explains why probability theory is the way it is. There are three rules that any coherent system of probabilities has to satisfy.
   (a) For any event E, $\mathsf{P}[E] \in [0,1]$;
   (b) $\mathsf{P}[E \text{ or } F] = \mathsf{P}[E] + \mathsf{P}[F] - \mathsf{P}[EF]$;
   (c) $\mathsf{P}[E \cap F] = \mathsf{P}[F \mid E]\,\mathsf{P}[E]$.

   If you want to understand probability in terms of betting, any sensible strategy you can imagine should satisfy these rules, for otherwise people can use arbitrage (an essential idea in the financial theory and asset and option pricing) to consistently make money from you.
2. E.T. Jaynes, Probability Theory: The Logic of Science. It is hard at times, but very interesting. it discusses applications and ideas behind probability and statistics, and how to think about these topics as a sensible person (not just as a theoretician).
3. C. Bishop, Pattern Recognition and Machine Learning. This is a really nice book on data analysis and machine learning. After the course, you can read the first two chapters. At the end of the master, you can read most of the book.