BIO STATISTICS COLLABORATION OF AUSTRALIA

Study Guide

Probability and Distribution Theory (PDT)

Semester 1, 2022

Prepared by:

Georgy Sofronov, Houying Zhu
School of Mathematical and Physical Sciences
Faculty of Science and Engineering
Macquarie University

MACQUARIE University
SYDNEY·AUSTRALIA

Contents

Probability and Distribution Theory (PDT)
Semester 1, 2022

Instructor contact details

Georgy Sofronov	**Houying Zhu**
Room 526 12 Wally's Walk School of Mathematical and Physical Sciences Macquarie University	Room 705 12 Wally's Walk School of Mathematical and Physical Sciences Macquarie University
phone: (02) 9850 8544	phone: (02) 9850 4749
e-mail: georgy.sofronov@mq.edu.au	e-mail: houying.zhu@mq.edu.au

Georgy Sofronov will be the lead unit convenor and the lecturer in the second half of the semester, Houying Zhu will be the second unit convenor and the lecturer in the first half of the semester.

Background

To obtain a sound understanding of the statistical methods used in the design and analysis of medical and health studies, it is essential to have a thorough knowledge of the theoretical basis for these techniques. This unit will focus on applying the calculus-based techniques learned in Mathematical Background for Biostatistics (MBB) to the study of probability and statistical distributions. These two units, together with Principles of Statistical Inference (PSI), will provide the core prerequisite mathematical statistics background required for the study of later subjects in the BCA program.

Unit summary

The emphasis of this unit is on the understanding of probability concepts and their application. Examples are taken from areas as diverse as biology, medicine, finance, sport, and the social and physical sciences. Topics include: the foundations of probability; probability models and their properties; some commonly used statistical distributions; relationships and association between variables; distribution of functions of random variables and sample statistics; approximations including the central limit theorem; and an introduction to the behaviour of random processes. Simulation is used to demonstrate many of these concepts.

Workload requirements

The expected workload for this unit is 10-12 hours per week on average, consisting of guided readings, discussion posts, independent study and completion of assessment tasks.

Prerequisites

Mathematical Background for Biostatistics (MBB)

Co-requisites

None

Learning Outcomes

At the completion of this unit students should be able to:

1. Analyse probability and conditional probability of an event by applying a probabilistic model for an experiment.
2. Apply a range of strategies to find and interpret the moments of discrete and continuous random variables including their expected values and variances.
3. Apply the Law of Large Numbers (LLN) and the Central Limit Theorem (CLT) to find asymptotic distribution of a sample mean
4. Analyse a bivariate probability distribution to find and interpret corresponding covariances, correlations, marginal and conditional probability distributions.
5. Apply Markov Chain (MC) theory to practical problems and tasks.

Unit content

The unit is divided into 12 topics/lectures, summarised in more detail below. Each topic will involve 1 week of study and generally includes the following material:

1. Lecture notes describing concepts and methods, and including examples of a more theoretical nature.
2. Selected readings from the recommended textbook.
3. One or more exercises illustrating the concepts/methods introduced in the notes and including more practically oriented exercises.

Study materials for all topics are downloadable from the eLearning unit site. Assignments and supplementary material, such as datasets will be posted to the unit site. Please note that we may not be able to post copies of copyright material (for example journal articles and book extracts)—for these you will have to rely on resources from your home university's library.

Recommended approaches to study

Students should work through each module systematically, following the module notes and any readings referred to, and working through the accompanying exercises. *You will learn a lot more efficiently if you tackle the exercises systematically as you work through the notes.* You are encouraged to post any content-related questions to eLearning, whether they relate directly to a given exercise, or are a request for clarification or further explanation of an area in the notes. You should also work through all of the computational examples in the notes for yourself on your own computer.

Solutions to the exercises in each module (except those to be submitted for assessment, as described below) will be posted online.

Method of communication with coordinator(s)

Questions about administrative aspects or course content can be emailed to the coordinator, and when doing so please use "PDT:" in the Subject line of your email to assist in keeping track of our email messages. Coordinator(s) will be available to answer questions related to the module notes and practical exercises, and to address any other issues that require clarification. However, please note that instructors are not necessarily available every day of the week and you should expect that it may take a day or so to respond to questions (possibly longer over weekends and during breaks).

We strongly recommend that you post content-related questions to the Discussions tool in the eLearning site. In 2022 we are using the Learning Management system hosted by the University of Sydney and Macquarie University. You should already be familiar with the university student learning system from previous units.

Topic descriptions

Below is an outline of the study topics, followed by a timetable and assessment description table

Each topic is scheduled to begin on a Monday and conclude on the Sunday of the following week. **The due date for submission of the assessments tasks is as shown in the table. Further details will be provided on the university learning management system.**

Topic	Material Covered
1	Experiments, sample spaces, Probability Rules, Permutations and Combinations.
2	Conditional Probability. Independence, Bayes' Theorem.
3	Random Variables. Probability Functions, Discrete Probability Distributions, Cumulative Distribution functions, Expected value and Variance. Moments.
4	Important Discrete Distributions: Bernoulli, Binomial, Geometric and Poisson.
5	Moment generating functions. Discrete Distributions: Negative Binomial and Hypergeometric.
6	Introduction to Continuous random variables. Cumulative distribution function.
7	Continuous Distributions: Uniform, Exponential.
8	Normal distribution.
9	Continuous Distributions: Gamma and Beta Distributions. Chebyshev's Theorem.
10	Sampling Distributions.
11	Joint Distributions: Discrete and Continuous cases.
12	Introduction to stochastic processes. Markov Chains.

Unit schedule

Semester 1, 2022 starts on Monday 21st Feb

Week	Week commencing	Topic	Notes	Assessment
1	21 Feb	1		
2	28 Feb	2		
3	7 Mar	3		
4	14 Mar	4		
5	21 Mar	5		
6	28 Apr	6		
7	4 Apr	7		Assignment 1
	11 - 22 Apr		Mid semester break	
8	25 Apr	8		
9	2 May	9		
10	9 May	10		
11	16 May	11		
12	23 May	12		
13	30 May			Assignment 2

Assessment

Assessment will be 2 assessments tasks as described below. These assessments will be posted on the eLearning site together with an online Announcement broadcasting their availability. Further details will be provided on the unit's eLearning site.

Assessment name	Assessment type	Coverage	Learning objectives	Weight
Assignment 1	Assignment	Topics 1-6	1,2,3	40%
Assignment 2	Assignment	Topics 1-12	1,2,3,4,5	60%

In general, you are required to submit your work typed in Word or similar (e.g. using Microsoft's Equation Editor for algebraic work) and we strongly recommend that you become familiar with equation typesetting software. If extensive algebraic work is involved you may submit neatly handwritten work, however please note that marks will potentially be lost if the solution cannot be understood by the markers due to unclear or illegible writing. This handwritten work should be scanned and collated into a single pdf file and submitted via the eLearning site. See the BCA Assessment Guide document for specific guidelines on acceptable standards for assessable work.

The instructors will generally avoid answering questions relating directly to the assessable material until after it has been submitted, but we encourage students to discuss the relevant parts of the notes among themselves, via eLearning. However **explicit solutions to assessable exercises should not be posted for others to use**, and each student's submitted work **must be** clearly their own, with anything derived from other students' discussion contributions clearly attributed to the source.

Submission of assessments and academic honesty policy

You should submit all your assessment material via eLearning unless otherwise advised. The use of Turnitin for submitting assessment items has been instigated within unit sites. For more detail please see pages 3-5 the BCA Student Assessment Guide.

The BCA pays great attention to academic honesty procedures. Please be sure to familiarise yourself with these procedures and policies at your university of enrolment. Links to these are available in the BCA Student Assessment Guide. Please also read carefully the Academic Honesty document in web page of this unit. When submitting assessments using Turnitin you will need to indicate your compliance with the plagiarism guidelines and policy at your university of enrolment before making the submission.

A special note regarding "contract cheating" sites: Unfortunately there have been instances in the past of students using such websites to post assignment questions and receive solutions (usually for a fee). We have arrangements with these sites to identify the students posting questions or accessing the solutions, and such students will be referred to and face disciplinary processes at their home university.

Late submission of assessments and extension procedure

The standard BCA policy for late penalties for submitted work is a 5% deduction from the earned mark for each day the assessment is late, up to a maximum of 10 days (including weekends and public holidays). Extensions are possible, but these need to be applied for (by email) as early as possible. The Unit Coordinator is not able to approve extensions beyond three days; for extensions beyond three days you need to apply to your home university, using their standard procedures.

Learning resources

There is no required textbook for this unit. Students may benefit from having access to the following background reference for additional reading and problems:

- Wackerly, D. D., Mendenhall, W., Scheaffer, R. L. Mathematical Statistics with Applications (4th,5th, 6th or 7th Editions)

The following books may also be useful background references:

- Ross, S. A First Course in Probability, Pearson (5th, 6th, 7th, 9th or 9th Editions)
- Ward, M. D. and Gundlach, E. (2016) Introduction to Probability, W. H. Freeman and Company
- Kinney, J.J. (1997) Probability - An Introduction with Statistical Applications, John Wiley and Sons
- Scheaffer R.L. (1994) Introduction to Probability and Its Applications, (2nd Edition) Duxbury Press
- Sincich,T., Levine, D.M., Stephan, D. (1999) Practical Statistics by Example using Microsoft Excel

Software

R, Excel and Wolfram Alpha will be used in the unit.

Feedback

Our feedback to you:

The types of feedback you can expect to receive in this unit are:

- Formal individual feedback on submitted assignments
- Responses to questions posted on iLearn

Your feedback to us:

One of the formal ways students have to provide feedback on teaching and their learning experience is through the BCA student evaluations at the end of each unit. The feedback is anonymous and provides the BCA with evidence of aspects that students are satisfied with and areas for improvement.

Required mathematical background

The mathematical techniques covered in MBB, in particular, differentiation and integration.

Changes to PDT since last delivery, including changes in response to student evaluation

PDT was last delivered by Monash University in Semester 2 2021. The current offering was developed by Georgy Sofronov and Houying Zhu from Macquarie University.