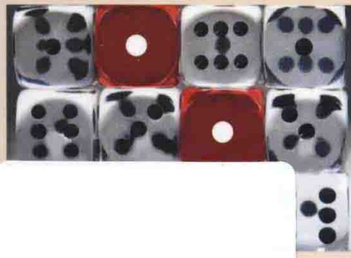

World Scientific Series on
Probability Theory and Its Applications

Volume 1

Random Matrices and Random Partitions

Normal Convergence

Zhonggen Su

World Scientific

World Scientific Series on
Probability Theory and Its Applications

Volume 1

Random Matrices and Random Partitions

Normal Convergence

Zhonggen Su
Zhejiang University, China

World Scientific

NEW JERSEY · LONDON · SINGAPORE · BEIJING · SHANGHAI · HONG KONG · TAIPEI · CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Su, Zhonggen.
Random matrices and random partitions normal convergence / by Zhonggen Su (Zhejiang University, China).
pages cm. -- (World scientific series on probability theory and its applications ; volume 1)
Includes bibliographical references and index.
ISBN 978-9814612227 (hardcover : alk. paper)
1. Random matrices. 2. Probabilities. I. Title.
QA273.43.S89 2015
519.2'3--dc23
2015004842

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Printed by FuIsland Offset Printing (S) Pte Ltd Singapore

Random Matrices and Random Partitions

Normal Convergence

World Scientific Series on Probability Theory and Its Applications

Series Editors: Zhenghu Li *(Beijing Normal University, China)*
Yimin Xiao *(Michigan State University, USA)*

Vol. 1 Random Matrices and Random Partitions: Normal Convergence
by Zhonggen Su (Zhejiang University, China)

To Yanping and Wanning

Preface

This book is intended to provide an introduction to remarkable probability limit theorems in random matrices and random partitions, which look rather different at a glance but have many surprising similarities from a probabilistic viewpoint.

Both random matrices and random partitions play a ubiquitous role in mathematics and its applications. There have been a great deal of research activities around them, and an enormous exciting advancement had been seen in the last three decades. A couple of excellent and big books have come out in recent years. However, the work on these two objects are so rich and colourful in theoretic results, practical applications and research techniques. No one book is able to cover all existing materials. Needless to say, these are rapidly developing and ever-green research fields. Only recently, a number of new interesting works emerged in literature. For instance, based on Johansson's work on deformed Gaussian unitary ensembles, two groups led respectively by Erdös-Yau and Tao-Vu successfully solved, around 2010, the long-standing conjecture of Dyson-Gaudin-Mehta-Wigner's bulk universality in random matrices by developing new techniques like the comparison principles and rigidity properties. Another example is that with the help of concepts of determinantal point processes coined by Borodin and Olshanski, around 2000, in the study of symmetric groups and random partitions, a big breakthrough has been made in understanding universality properties of random growth processes. Each of them is worthy of a new book.

This book is mainly concerned with normal convergence, namely central limit theorems, of various statistics from random matrices and random partitions as the model size tends to infinity. For the sake of writing and learning, we shall only focus on the simplest models among which are circular unitary ensemble, Gaussian unitary ensemble, random uniform partitions and random Plancherel partitions. As a matter of fact, many of the results addressed in this book are found valid for more general models. This book consists of three parts as follows.

We shall first give a brief survey on normal convergence in Chapter 1. It includes the well-known laws of large numbers and central limit theorems for independent identically distributed random variables and a few methods widely used in dealing with normal convergence. In fact, the central limit theorems are arguably regarded as one of the most important universality principles in describing laws of random phenomena. Most of the materials can be found in any standard probability theory at graduate level. Because neither the eigenvalues of a random matrix with all entries independent nor the parts of random partitions are independent of each other, we need new tools to treat statistics of dependent random variables. Taking this into account, we shall simply review the central limit theorems for martingale difference sequences and Markov chains. Besides, we shall review some basic concepts and properties of convergence of random processes. The statistic of interest is sometimes a functional of certain random process in the study of random matrices and random partitions. We will be able to make use of functional central limit theorems if the random processes under consideration is weakly convergent. Even under the stochastic equicontinuity condition, a slightly weaker condition than uniform tightness, the Gikhmann-Skorohod theorem can be used to guarantee convergence in distribution for a wide class of integral functionals.

In Chapters 2 and 3 we shall treat circular unitary ensemble and Gaussian unitary ensemble respectively. A common feature is that there exists an explicit joint probability density function for eigenvalues of each matrix model. This is a classic result due to Weyl as early as the 1930s. Such an explicit formula is our starting point and this makes delicate analysis possible. Our focus is upon the second-order fluctuation, namely asymptotic distribution of a certain class of linear functional statistics of eigenvalues. Under some smooth conditions, a linear eigenvalue statistic satisfies the central limit theorem without normalizing constant $\sqrt{n}$, which appears in classic Lévy-Feller central limit theorem for independent identically distributed random variables. On the other hand, either indicator function or logarithm function does not satisfy the so-called smooth condition. It turns out that the number of eigenvalues in an interval and the logarithm of characteristic polynomials do still satisfy the central limit theorem after suitably normalized by $\sqrt{\log n}$. The $\log n$-phenomena is worthy of more attention since it will also appear in the study of other similar models. In addition to circular and Gaussian unitary ensembles, we shall consider their extensions like circular β matrices and Hermite β matrices where $\beta > 0$ is a model parameter. These models were introduced and studied at length by Dyson in the early 1960s to investigate energy level behaviors in complex dynamic systems. A remarkable contribution at this direction is that there is a five (resp. three) diagonal sparse matrix model representing circular β ensemble (resp. Hermite β ensemble).

In Chapters 4 and 5 we shall deal with random uniform partitions and random Plancherel partitions. The study of integer partitions dates back to Euler as early as the 1750s, who laid the foundation of partition theory by determining the number of all distinct partitions of a natural number. We will naturally produce a probability space by assigning a probability to each partition of a natural number. Uniform measure and Plancherel measure are two best-studied objects. Young diagram and Young tableau are effective geometric representation in analyzing algebraic, combinatorial and probabilistic properties of a partition. Particularly interesting, there exists a nonrandom limit shape (curve) for suitably scaled Young diagrams under both uniform and Plancherel measure. This is a kind of weak law of large numbers from the probabilistic viewpoint. To proceed, we shall further investigate the second-order fluctuation of a random Young diagram around its limit shape. We need to treat separately three different cases: at the edge, in the bulk and integrated. It is remarkable that Gumbel law, normal law and Tracy-Widom law can be simultaneously found in the study of random integer partitions. A basic strategy of analysis is to construct a larger probability space (grand ensemble) and to use the conditioning argument. Through enlarging probability space, we luckily produce a family of independent geometric random variables and a family of determinantal point processes respectively. Then a lot of well-known techniques and results are applicable.

Random matrices and random partitions are at the interface of many science branches and they are fast-growing research fields. It is a formidable and confusing task for a new learner to access the research literature, to acquaint with terminologies, to understand theorems and techniques. Throughout the book, I try to state and prove each theorem using language and ways of reasoning from standard probability theory. I hope it will be found suitable for graduate students in mathematics or related sciences who master probability theory at graduate level and those with interest in these fields. The choice of results and references is to a large extent subjective and determined by my personal point of view and taste of research. The references at the end of the book are far from exhaustive and in fact are rather limited. There is no claim for completeness.

This book started as a lecture note used in seminars on random matrices and random partitions for graduate students in the Zhejiang University over these years. I would like to thank all participants for their attendance and comments. This book is a by-product of my research project. I am grateful to the National Science Foundation of China and Zhejiang Province for their generous support in the past ten years. I also take this opportunity to express a particular gratitude to my teachers, past and present, for introducing me to the joy of mathematics. Last, but not least, I wish to thank deeply my family for their kindness and love which is indispensable in completing this project.

I apologize for all the omissions and errors, and invite the readers to report any remarks, mistakes and misprints.

Zhonggen Su
Hangzhou
December 2014

Contents

Chapter 1

Normal Convergence

1.1 Classical central limit theorems

Throughout the book, unless otherwise specified, we assume that $(\Omega, \mathcal{A}, P)$ is a large enough probability space to support all random variables of study. E will denote mathematical expectation with respect to P.

Let us begin with Bernoulli's law, which is widely recognized as the first mathematical theorem in the history of probability theory. In modern terminology, the Bernoulli law reads as follows. Assume that $\xi_n, n \geq 1$ is a sequence of independent and identically distributed (i.i.d.) random variables, $P(\xi_n = 1) = p$ and $P(\xi_n = 0) = 1 - p$, where $0 < p < 1$. Denote $S_n = \sum_{k=1}^{n} \xi_k$. Then we have

$$\frac{S_n}{n} \xrightarrow{P} p, \quad n \to \infty. \tag{1.1}$$

In other words, for any $\varepsilon > 0$,

$$P\Big(\Big|\frac{S_n}{n} - p\Big| > \varepsilon\Big) \to 0, \quad n \to \infty.$$

It is this law that first provide a mathematically rigorous interpretation about the meaning of probability p that an event A occurs in a random experiment. To get a feeling of the true value p (unknown), what we need to do is to repeat independently a trial n times (n large enough) and to count the number of A occurring. According to the law, the larger n is, the higher the precision is.

Having the Bernoulli law, it is natural to ask how accurate the frequency S_n/n can approximate the probability p, how many times one should repeat the trial to attain the specified precision, that is, how big n should be.

With this problem in mind, De Moivre considered the case $p = 1/2$ and proved the following statement:

$$P\Big(a \leq \frac{S_n - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} \leq b\Big) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx. \tag{1.2}$$

Later on, Laplace further extended the work of De Moivre to the case $p \neq 1/2$ to obtain

$$P\Big(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\Big) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx. \tag{1.3}$$

Formulas (1.2) and (1.3) are now known as De Moivre-Laplace central limit theorem (CLT).

Note $ES_n = np$, $Var(S_n) = np(1-p)$. So $(S_n - np)/\sqrt{np(1-p)}$ is a normalized random variable with mean zero and variance one. Denote $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$, $x \in \mathbb{R}$. This is a very nice function from the viewpoint of function analysis. It is sometimes called bell curve since its graph looks like a bell, as shown in Figure 1.1.

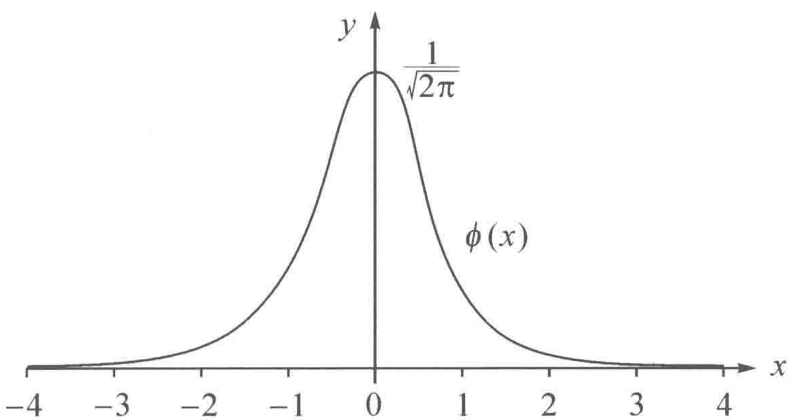

Fig. 1.1 Bell curve

The Bernoulli law and De Moivre-Laplace CLT have become an indispensable part of our modern daily life. See Billingsley (1999a, b), Chow (2003), Chung (2000), Durrett (2010) and Fischer (2011) for a history of the central limit theorem and the link to modern probability theory. But what is the proof? Any trick? Let us turn to De Moivre's original proof of (1.2). To control the left hand side of (1.2), De Moivre used the binomial formula

$$P(S_n = k) = \binom{n}{k} \frac{1}{2^n}$$

and invented together with Stirling the well-known Stirling formula (it actually should be called De Moivre-Stirling formula)

$$n! = n^n e^{-n}\sqrt{2\pi n}(1+o(1)).$$

Setting $k = n/2 + \sqrt{n}x_k/2$, where $a \le x_k \le b$, we have

$$P\Big(\frac{S_n - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} = x_k\Big) = \frac{1}{2^n}\cdot\frac{n^n e^{-n}\sqrt{2\pi n}(1+o(1))}{k^k e^{-k}(n-k)^{n-k}e^{-(n-k)}\sqrt{2\pi k}\sqrt{2\pi(n-k)}}$$
$$= \frac{1}{\sqrt{2\pi n}}e^{-x_k^2/2}(1+o(1)).$$

Taking sum over k yields the integral of the right hand side of (1.2).

Given a random variable X, denote its distribution function $F_X(x)$ under P. Let X, X_n, $n \ge 1$ be a sequence of random variables. If for each continuity point x of F_X,

$$F_{X_n}(x) \to F_X(x), \quad n\to\infty,$$

then we say X_n converges in distribution to X, and simply write $X_n \xrightarrow{d} X$. In this terminology, (1.3) is written as

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1), \quad n\to\infty,$$

where $N(0,1)$ stands for a standard normal random variable.

As the reader may notice, the Bernoulli law only deals with frequency and probability, i.e., Bernoulli random variables. However, in practice people are faced with a lot of general random variables. For instance, measure length of a metal rod. Its length, μ, is intrinsic and unknown. How do we get to know the value of μ? Each measurement is only a realization of μ. Suppose that we measure repeatedly the metal rod n times and record the observed values $\xi_1, \xi_2, \cdots, \xi_n$. It is believed that $\sum_{k=1}^n \xi_k/n$ give us a good feeling of how long the rod is. It turns out that a claim similar to the Bernoulli law is also valid for general cases. Precisely speaking, assume that ξ is a random variable with mean μ. ξ_n, $n \ge 1$ is a sequence of i.i.d. copy of ξ. Let $S_n = \sum_{k=1}^n \xi_k$. Then

$$\frac{S_n}{n} \xrightarrow{P} \mu, \quad n\to\infty. \tag{1.4}$$

This is called the Khinchine law of large numbers. It is as important as the Bernoulli law. As a matter of fact, it provides a solid theoretic support for a great deal of activity in daily life and scientific research.

The proof of (1.4) is completely different from that of (1.1) since we do not know the exact distribution of ξ_k. To prove (1.4), we need to invoke the following Chebyshev inequality. If X is a random variable with finite mean μ and variance σ^2, then for any positive $x > 0$

$$P(|X-\mu| > x) \le \frac{\sigma^2}{x^2}.$$

In general, we have

$$P(X > x) \le \frac{Ef(X)}{f(x)},$$

where $f : \mathbb{R} \mapsto \mathbb{R}$ is a nonnegative nondecreasing function. We remark that the Chebyshev inequalities have played a fundamental role in proving limit theorems like the law of large numbers.

Having (1.4), we next naturally wonder what the second order fluctuation is of S_n/n around μ? In other words, is there a normalizing constant $a_n \to \infty$ such that $a_n(S_n - n\mu)/n$ converges in distribution to a certain random variable? What is the distribution of the limit variable? To attack these problems, we need to develop new tools and techniques since the De Moivre argument using binomial distribution is no longer applicable.

Given a random variable X with distribution function F_X, define for every $t \in \mathbb{R}$,

$$\begin{aligned}\psi_X(t) &= Ee^{\mathbf{i}tX}\\ &= \int_{\mathbb{R}} e^{\mathbf{i}tx} dF_X(x).\end{aligned}$$

Call $\psi_X(t)$ the characteristic function of X. This is a Fourier transform of $F_X(x)$. In particular, if X has a probability density function $p_X(x)$, then

$$\psi_X(t) = \int_{\mathbb{R}} e^{\mathbf{i}tx} p_X(x)dx;$$

while if X takes only finitely or countably many values, $P(X = x_k) = p_k$, $k \ge 1$, then

$$\psi_X(t) = \sum_{k=1}^{\infty} e^{\mathbf{i}tx_k} p_k.$$

Note the characteristic function of any random variable is always well defined no matter whether its expectation exists.

Example 1.1. (i) If X is a normal random variable with mean μ and variance σ^2, then $\psi_X(t) = e^{\mathbf{i}\mu t - \sigma^2 t^2/2}$;
(ii) If X is a Poisson random variable with parameter λ, then $\psi_X(t) = e^{\lambda(e^{\mathbf{i}t}-1)}$;
(iii) If X is a standard Cauchy random variable, then $\psi_X(t) = e^{-|t|}$.