creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

이학석사 학위논문

비가환확률공간에서 중심극한정리

Central Limit Theorem in Non-commutative Probability Space

2018년 2월

서울대학교 대학원

통계학과

연 혜 민

비가환확률공간에서 중심극한정리

지도교수 임 요 한

이 논문을 이학석사 학위논문으로 제출함

2017년 12월

서울대학교 대학원

통계학과

연 혜 민

연혜민의 이학석사 학위논문을 인준함

2017년 12월

위 원 장	김 용 대	(인)
부위원장	임 요 한	(인)
위 원	이 재 용	(인)

Abstract

Hyemin Yeon
The Department of Statistics
The Graduate School
Seoul National University

The assumption of commutativity of random variables is very natural in classical probability theory. However, many objects such as random matrices do not satisfy commutativity. Thus, we need to build a non-commutative probabilistic structure to handle these objects. Then, many properties in classical probability theory differs from those in this new theory, titled "free probability theory", including independence and convergence in distribution. In this paper, we introduce algebraically a non-commutative probability space and provides the notion of "free independence", which is the analogue of independence in classical probability theory. Furthermore, we prove the free version of Central Limit Theorem by using such new concepts and several tools in combinatorics, and compare it to the classical Central Limit Theorem.

Keyword : *Non-commutative probability space, Free independence, Central limit theorem, Random matrix theory, Combinatorics*
Student Number : 2015-20304

Contents

Chapter 1

Introduction

In a classical probability theory, we need the notion of measuarability of sets and functions. Also, we define expectation and independence of random variables in an analytic sense. It means that we need the notion of integration. To illustrate, for a real random variable X in a probability space $(\Omega, \mathcal{F}, P)$, the expectation of X is defined as

$$E[X] = \int_{\Omega} X(\omega) dP(\omega).$$

However, we can define expectation and independence in an algebraic sense ignoring the notion of measure and integration. A linear functional defined on an algebraic structure acts as a role of expectation operator. This way is more powerful when our probability space is non-commutative. In a non-commutative probability space, we define "freeness", which is an analogue of the concept "independence" in classical probability theory, and it is easier to handle non-commutative objects such as random matrices.

In free probability theory, created by Dan Voiculescu (1), combinatoric notions are pivotal including Catalan numbers and non-crossing partitions. Thus, in chapter 2, some concepts in combinatorics are reviewed as well as basic notions in classical probability theory. Chapter 3 introduces the definition of distribution and independence in a non-commutativel probability space. Finally, we prove the free version of Central Limit Theorem (CLT) compared to that in classical probability theory in chapter 4.

Chapter 2

Preliminaries

In this chapter, we review basic concepts in classical probability theory and introduce notions in combinatorics which will be used in the proof of free CLT.

2.1. Classical Probability Theory

In classical probability theory, we first define a proability space $(\Omega, \mathcal{F}, P)$ and a random variable on this probability space $(\Omega, \mathcal{F}, P)$. Moreover, the notions of independence, distributions, and convergence are built in this framework. The details of classical probability theory are explained in the appendix.

Our final goal is to prove the free version of CLT, so let us review the classical version of CLT. We assume that $(\Omega, \mathcal{F}, P)$ is a probability space throughout this paper.

Theorem 2.1 (Central Limit Theorem (CLT))**.** *Let* $X_1, X_2, ...$ *be identically and independently distributed random variables on a probability space* $(\Omega, \mathcal{F}, P)$ *with* $EX_1 = \mu$ *and* $var[X_1] = \sigma^2 > 0$*. Set* $\bar{X}_n := \sum_{i=1}^{n} X_i$ *for each* $n \in \mathbb{N}$*. Then, we have*

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow[n\to\infty]{d} Z, \tag{2.1}$$

or equivalently,

$$\frac{Y_1 + \cdots + Y_n}{\sqrt{n}} \xrightarrow[n\to\infty]{d} \sigma Z, \tag{2.2}$$

where $Y_i = X_i - \mu$ *for all* $i \in \mathbb{N}$ *and* Z *is a random varaible following the standard normal distribution,*

As we know, the assumption of i.i.d. can be relaxed into a weak condition such as the Lindeberg's condition. Note that the limit distribution is the normal distribution. There is an analogue of normal distribution in free CLT, titled *semi-circular law.*

Definition 2.2. For a real or complex random varible X on a probability space $(\Omega, \mathcal{F}, P)$, we say that X has *semicircular distribution (or law)* of radius r if its probability density function is defined as

$$f(t) = \frac{2}{\pi r^2}\sqrt{r^2 - t^2} I(|t| \leq 2)$$

for $t \in \mathbb{R}$ or $t \in \mathbb{C}$. Further, if the semicircular distrubition of radius 2 is called the *standard semicircular distribution.*

2.2. Some Concepts in Combinatorics

It is commonly to use some tools in combinatorics to deal with problems in free probability theory. We now provide several notions used in the proof of free CLT.

Definition 2.3 (Catalan Numbers)**.** For $n \in \mathbb{N}_0$, we define the *n-th Catalan number* as

$$C_n := \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n!(n+1)!}$$

with the convention that $C_0 = 1$.

There is another equivalent condition to define the Catalan numbers recursively.

Proposition 2.4. The sequence $\{C_n\}$ in $\mathbb{N}$ can be the sequence of Catalan numbers if it satisfies the following conditions:

(1) $C_0 = C_1 = 1$, and

(2) $C_p = \sum_{j=1}^{p} C_{j-1}C_{p-j}$ for all $p \geq 2$.

The semi-circular distribution plays an important role in proving our new CLT, and it has special property related to Catalan numbers.

Proposition 2.5. The n-th Catalan number C_n is the $(2n)$-th moment of the standard semicircular distribution.

Definition 2.6. Let S be a finite totally ordered set.

(1) We call $\pi = \{V_1, ..., V_r\}$ a *partition* of S if V_i's are pairwise disjoint and non-empty subsets of S such that $\bigcup_{i=1}^{r} V_i = S$.

(2) $V_1, ..., V_r$ are called *blocks* of π.

(3) For $p, q \in S$, we write $p \sim_\pi q$ if p and q belong to the same block of π.

(4) The set of all partitions of S is denoted by $\mathcal{P}(S)$. In particular, if $S = \{1, ..., n\}$, we denote it by $\mathcal{P}(n)$.

(5) A partition π of S is called *crossing* if there exist $p_1, p_2, q_1, q_2 \in S$ such that $p_1 < q_1 < p_2 < q_2$ and $p_1 \sim_\pi p_2 \nsim_\pi q_1 \sim_\pi q_2$.

(6) If π is not crossing, it is said to be *non-crossing.*

(7) The set of all non-crossing partitions of S is denoted by $NC(S)$. In particular, if $S = \{1, ..., n\}$, we denote it by $NC(n)$.

(8) Let π be a partition of S and $\pi = \{V_1, ..., V_r\}$. If $|V_i| = 2$ for all $i = 1, ..., r$, then π is called a *pair partition* or *pairing.*

(9) The set of all pairings of S is denoted by $\mathcal{P}_2(S)$, and the set of all non-crossing pairings of S by $NC_2(S)$. If $S = \{1, ..., n\}$, then they are denoted by $\mathcal{P}_2(n)$ and $NC_2(n)$, respectively.

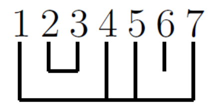
The notion of non-crossing partitions can be understood better with graphical explanations.

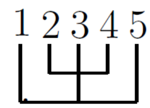
(1) Consider a set $S = \{1, 2, 3, 4, 5, 6, 7\}$ and a partition

$$\pi = \{\{1, 4, 5, 7\}, \{2, 3\}, \{6\}\}$$

of S. Then we can see that π is a non-crossing partition of S. This figurative representation allow us to know why the name is non-crossing.

(2) Consider another example of partition $\pi = \{\{1, 3, 5\}, \{2, 4\}\}$ in a set $S = \{1, 2.3, 4, 5\}$. This partition is evidently crossing.





One can find a strong relation bewteen Catalan numbers and the number of non-crossing pair partitions as follows.

Lemma 2.7. *For $k \in \mathbb{N}$, let D_{2k} be the number of non-crossing pair partitions. Then, we have $D_{2k} = C_k$.*

Chapter 3

Non-commutative Probability Spaces

In classical probability theory, probability spaces are defined in an analytic way with the notion of measure or integration. However, such concepts are not necessary and we can define probability spaces in an algebraic way just with an algebraic structure and a linear operator. Thus, we first review concepts of an algebraic structure, named an "algebra."

Definition 3.1. Let K be a field, and $\mathcal{A}$ a vector space over K equipped with an additional binary operation from $\mathcal{A} \times \mathcal{A}$ to $\mathcal{A}$ denoted by $\cdot$. (i.e., if x and y are any two elements of $\mathcal{A}$, $x \cdot y$ is the *product* of x and y.)

(1) Then, $\mathcal{A}$ is called an *algebra* over K if for any $x, y, z \in \mathcal{A}$ and $a, b \in K$,

- (i) right distributivity: $(x + y) \cdot z = x \cdot z + y \cdot z$;
- (ii) left distributivity: $x \cdot (y + z) = x \cdot y + x \cdot z$;
- (iii) compatibility with scalars: $(ax) \cdot (by) = (ab)(x \cdot y)$.

These three axioms are another way of saying that the binary operation is *bilinear*. An algebra over K is sometimes also called a K-*algebra*, and in this case, K is called the *base field* of $\mathcal{A}$. The binary operation is often referred to as *multiplication* in $\mathcal{A}$.

(2) An algebra $\mathcal{A}$ over $\mathbb{C}$ is said to be *unital* if it has the identity element with respect to the multiplication. Also, a map on a unital algebra $\mathcal{A}$ over a field K is said to be *unital* if it preserves the identity element.

(3) An algebra $\mathcal{A}$ over a field K is called a $*$-*algebra* if $\mathcal{A}$ is endowed with an antilinear $*$-operation $\mathcal{A} \ni a \mapsto a^* \in \mathcal{A}$ such that $(a^*)^* = a$ and $(ab)^* = b^*a^*$ for all $a, b \in \mathcal{A}$.

We now define a non-commuatative probability space without any analytic concept.

Definition 3.2. A pair of an unital algebra $\mathcal{A}$ over $\mathbb{C}$ and an unital linear functional φ on $\mathcal{A}$ is called a *non-commutative probability space (NCPS)*, and is denoted by $(\mathcal{A}, \varphi)$. Further, an element a in $\mathcal{A}$ is called a *(non-commutative) random variable* in $(\mathcal{A}, \varphi)$.

The linear functional φ on the algebra $\mathcal{A}$ plays a role of the expectation in classical probability theory. However, since random varaibles in a NCPS may be non-commutative, we need more definitions on this linear functional.

Definition 3.3. Let $(\mathcal{A}, \varphi)$ be a NCPS.

(1) φ is called a *trace* if $\varphi(ab) = \varphi(ba)$ for all $a, b \in \mathcal{A}$. In this case, $(\mathcal{A}, \varphi)$ is said to be *tracial*.

(2) Assume that $\mathcal{A}$ is a $*$-algebra. If we have that $\varphi(a^*a) \geq 0$ for all $a \in \mathcal{A}$, then φ is said to be *positive*, and $(\mathcal{A}, \varphi)$ is called a *$*$-probability space.*

(3) If $\varphi(a^*a) = 0$ implies that $a = 0$ for all $a \in \mathcal{A}$, then φ is said to be faithful.

In the framework of a $*$-probability space, we can define various types of random variables due to $*$-operation.

Definition 3.4. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space.

(1) $x \in \mathcal{A}$ is said to be *selfadjoint* if $x = x^*$.

(2) $u \in \mathcal{A}$ is said to be *unitary* if $u^*u = uu^* = 1$.

(3) $a \in \mathcal{A}$ is called a *normal* random variable if $a^*a = aa^*$.

Example 3.5. Consider a probability space $(\Omega, \mathcal{F}, P)$ in classical probability theory, and let $\mathcal{RV}$ be a set of all real-valued random variables on $(\Omega, \mathcal{F}, P)$ such that all their moments are finite.

(1) Note that for all $X, Y \in \mathcal{RV}$ and $a, b \in \mathbb{R}$, $aX + bY \in \mathcal{RV}$. Thus, $\mathcal{RV}$ is a vector space over $\mathbb{R}$. Furthermore, the space of random variables satisfies three axioms for the definition of algebra since the set of measurable functions are closed under multiplication of functions. This implies that $\mathcal{RV}$ is an algebra over $\mathbb{R}$.

(2) It should be noted that the unit element in $\mathcal{RV}$ is a measurable function $1_{\mathcal{RV}} : \Omega \to \mathbb{R}$ defined as $1_{\mathcal{RV}}(\omega) = 1$ for all $\omega \in \Omega$. We now consider the expectation operator E. Then, it is obvious that E is an unital linear

functional on $\mathcal{RV}$. Hence, our classical probability space $(\mathcal{RV}, E)$ accords with the definition of a NCPS.

(3) Since random variables actually are commutative, it is evident that $(\mathcal{RV}, E)$ is tracial.

(4) One can show that $(\mathcal{RV}, E)$ is a tracial NCPS even if $\mathcal{RV}$ is the set of all complex-valued random variables. Also, the $*$-operation on $(\mathcal{RV}, E)$ is given as complex conjugation. Then, it is also obivous that $\mathcal{RV}$ is a $*$-algebra, and since

$$E[X^*X] = E\left[\|X\|^2\right] \geq 0$$

for all $X \in \mathcal{RV}$, E is positive. Thus, $(\mathcal{RV}, E)$ is a $*$-probability space.

The next example is also familiar, but this space is non-commutative.

Example 3.6.

(1) Consider the set of all $d \times d$ complex matrices and we denote it by $M_d = M_d(\mathbb{C})$. The linear functional on this space is given as the normalized trace $\mathrm{tr} : M_d \to \mathbb{C}$ defined by

$$tr(a) = \frac{1}{d}\sum_{i=1}^{d} \alpha_{ii} \text{ for } a = (\alpha_{ij})_{i,j=1}^{d} \in M_d.$$

Then, it is easy to check that (M_d, tr) is a $*$-probability space when the $*$-operation is given as the transpose with complex conjutation of all entries.

(2) Let $\mathcal{A} = M_d(\mathcal{RV})$ where $\mathcal{RV}$ is the set of all complex-valued random variables with all finite moments. Define a linear functional φ on $\mathcal{A}$ as

$$\varphi(a) := \int \operatorname{tr}(a(\omega))dP(\omega) \text{ for } a \in \mathcal{A}.$$

This is also a $*$-probability space, and the space of *random matrices* over $(\Omega, \mathcal{F}, P)$.

We are now ready to define a distribution of a random variable.

Definition 3.7. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space and let $a \in \mathcal{A}$.

(1) Assume that a is a normal element of $\mathcal{A}$. If there exists a compactly supported probability measure μ on $\mathbb{C}$ such that

$$\int z^k \bar{z}^l d\mu(z) = \varphi(a^k (a^*)^l)$$

for every $k, l \in \mathbb{N}$, then μ is uniquely determined and called the *$*$-distribution* of a (*in analytic sense*).

(2) For $k \in \mathbb{N}_0$ and $\varepsilon(1), ..., \varepsilon(k) \in \{1, *\}$, $\varphi(a^{(\varepsilon(1))} \cdots a^{e(k)})$ is called a $*$-moment of a. If a is selfadjoint, $\varphi(a^k)$ is called the *k-th moment* of a for $k \in \mathbb{N}$. Furthermore, in the same way, $\varphi(a)$ is called the *mean* of a and

$$var(a) := \varphi\left(\{a - \varphi(a)\}^2\right) = \varphi(a^2) - \varphi(a)^2$$

is called the *variance* of a.

(3) The *$*$-distribution* of a is defined as the linear functional

$$\mu : \mathbb{C}\langle x, x^* \rangle \to \mathbb{C}$$

determined by the equations

$$\mu(x^{(\varepsilon(1))}\cdots x^{e(k)}) = \varphi(a^{(\varepsilon(1))}\cdots a^{e(k)})$$

for every $k \in \mathbb{N}_0$ and all $\varepsilon(1), ...\varepsilon(k) \in \{1, *\}$ where $\mathbb{C}\langle x, x^*\rangle$ is denoted by the unital algebra which is freely generated by two non-commutating indeterminates x and x^*. This is called $*$-distribution *in algebraic sense* to avoid ambiguity.

Refer to ANRS and Fraleighi for the details of definition of a $*$-distribution such as the existence of such compactly supported measure μ or the notion of a freely generated algebra.

We now define the notion of independence in free probability theory. As in classical probability theory, we first define independence among algebraic structures, and then, indenpendence among random variables.

Definition 3.8. Let $(\mathcal{A}, \varphi)$ be a NCPS and I an index set. Subalgebras $(\mathcal{A}_i)_{i\in I}$ are said to be *freely independent* or *free* if $\varphi(a_1 \cdots a_k) = 0$ whenever

(i) $i(1) \neq i(2), i(2) \neq i(3), ..., i(k-1) \neq i(k)$ for $k \in \mathbb{N}$,

(ii) $a_j \in \mathcal{A}_{i(j)}$ for all $j = 1, ..., k$,

(iii) and $\varphi(a_j) = 0$ for all $j = 1, ..., k$.

Definition 3.9. Let $(\mathcal{A}, \varphi)$ be a NCPS and $a_i \in \mathcal{A}$ for some index set I.

(1) Random variables $(a_i)_{i\in I}$ are said to be *freely independent* or *free* if the unital algberabs $\mathcal{A}_i := \mathrm{alg}(1, a_i)$ generated by $a_i \in \mathcal{A}$ $(i \in I)$ are free.

(2) More specifically, $(a_i)_{i\in I}$ are said to be *freely independent* or *free* if

$$\varphi\left(P_1(a_{i(1)})\cdots P_k(a_{i(k)})\right) = 0$$

for all polynomials $P_1, ..., P_k \in \mathbb{C}\langle X\rangle$ in one indeterminate X and all $i(1), ..., i(k) \in I$ such that $i(1) \neq i(2), i(2) \neq i(3), ..., i(k-1) \neq i(k)$ and $\varphi(P_j(a_{i(j)})) = 0$ for all $j = 1, ..., k$.

(3) In the case of a $*$-probability space, random variables $(a_i)_{i\in I}$ are said to be *$*$-freely independent* or *$*$-free* if the unital $*$-algberabs $\mathcal{A}_i := \mathrm{alg}(1, a_i, a_i^*)$ generated by $a_i \in \mathcal{A}$ $(i \in I)$ are free.

Example 3.10. Consider a NCPS $(\mathcal{A}, \varphi)$, and let $a, a' \in \tilde{\mathcal{A}}$ and $b \in \tilde{\mathcal{B}}$ where $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$ are free subalgebras of $\mathcal{A}$.

(1) Let $a^0 := a - \varphi(a)1$ and $b^0 := b - \varphi(b)1$, which are called the *centering* of random variables. Then, since $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$ are free, a^0 and b^0 are free and hence we have $\varphi(a^0b^0) = 0$. Then,

$$\begin{aligned}0 = \varphi(a^0b^0) &= \varphi\left((a-\varphi(a)1)(b-\varphi(b)1)\right)\\ &= \varphi(ab) - \varphi(a1)\varphi(b) - \varphi(a)\varphi(1b) + \varphi(a)\varphi(b)1\\ &= \varphi(ab) - \varphi(a)\varphi(b).\end{aligned}$$

This implies that

$$\varphi(ab) = \varphi(a)\varphi(b) \tag{3.1}$$

and we have the same result in classical probability theory like the proposition A.6 in the appendix.

(2) In the same way, we use $\varphi(a^0 b^0 (a')^0) = 0$ to show that

$$\varphi(aba') = \varphi(aa')\varphi(b). \tag{3.2}$$

(3) We now assume that a and b are commuting, i.e., $ab = ba$. Note that a^0 and b^0 are free as we see in (1). Thus, by definition, we get

$$\varphi(a^0 b^0 a^0 b^0) = 0$$

, and obtain the following equation:

$$\varphi(abab) = \varphi(a^2)\varphi(b)^2 + \varphi(a)^2\varphi(b^2) - \varphi(a)^2\varphi(b)^2.$$

Since $a^2 \in \tilde{\mathcal{A}}$ and $b^2 \in \tilde{\mathcal{B}}$, we can apply the result (1) to a^2 and b^2. Then, since a and b are commuting, we have

$$\varphi(a^2)\varphi(b^2) = \varphi(a^2 b^2) = \varphi(abab).$$

Combining these equations, we have

$$\begin{aligned} 0 &= \left(\varphi(a^2) - \varphi(a)^2\right)\left(\varphi(b^2) - \varphi(b)^2\right) \\ &= \varphi\left((a - \varphi(a)1)^2\right)\varphi\left((b - \varphi(b)1)^2\right). \end{aligned}$$

This implies that $\varphi\left((a - \varphi(a)1)^2\right) = 0$ or $\varphi\left((b - \varphi(b)1)^2\right) = 0$.

(4) Recall that $\varphi\left((a - \varphi(a)1)^2\right) = 0$ or $\varphi\left((b - \varphi(b)1)^2\right) = 0$ means that at least one of these random variables is almost surely constant. This indicates that freeness cannot be considered a non-commutative version of classical independence.

Chapter 4

Central Limit Theorem

We are all ready to state and prove the free version of CLT. Since CLT says the limit behavior of random variables, we need the notion of the limit of distributions.

Definition 4.1. Let $(\mathcal{A}_N, \varphi_N)$ $(N \in \mathbb{N})$ and $(\mathcal{A}, \varphi)$ be NCPSs, and $a_N \in \mathcal{A}_N$ for each $N \in \mathbb{N}$ and $a \in \mathcal{A}$. We say that $(a_N)_{N \in \mathbb{N}}$ *converges in distribution* to a if

$$\lim_{N \to \infty} \varphi_N(a_N^n) = \varphi(a^n)$$

for each $n \in \mathbb{N}$, and denote it by

$$a_N \xrightarrow[N \to \infty]{d} a.$$

The above definition of convergence in distribution says that it is enough to check the convegence of all moments to prove CLT. We now state the free version of CLT.

Theorem 4.2 (Free Central Limit Theorem)**.** *Let $(\mathcal{A}, \varphi)$ be a $*$-probability space and $(a_n)_{n\in\mathbb{N}}$ be a sequence of free and indentically distributed selfadjoint random variables in $\mathcal{A}$ with $\varphi(a_n) = 0$ and $\sigma^2 := \varphi(a_n^2)$ for all $n \in \mathbb{N}$. Then, we have*

$$\frac{a_1 + \cdots + a_N}{\sqrt{N}} \xrightarrow[N\to\infty]{d} s$$

where s is a semicircular element of variance σ^2.

Proof. Let $N, n \in \mathbb{N}$. Then,

$$\varphi\left((a_1 + \cdots + a_N)^n\right) = \sum_{1\leq r(1),\ldots,r(n)\leq N} \varphi\left(a_{r(1)} \cdots a_{r(n)}\right)$$

Note that since all random variables are identically distributed, we observe that

$$\varphi\left(a_{r(1)} \cdots a_{r(n)}\right) = \varphi\left(a_{p(1)} \cdots a_{p(n)}\right)$$

whenever

$$r(i) = r(j) \iff p(i) = p(j) \ \forall i, j = 1, ..., n.$$

Consider a prtition $\pi \in \mathcal{P}(n)$. Then we can consider a relation $\hat{=}$ defined as

$$(r(1), ..., r(n)) \hat{=} \pi \iff [r(p) = r(q) \iff p \sim_\pi q]\,. \tag{4.1}$$

Also, we denote this common value of $\varphi\left(a_{r(1)} \cdots a_{r(n)}\right)$ for all tuples $(r(1), ..., r(n))$ with $(r(1), ..., r(n)) \hat{=} \pi$ by κ_π. Let

$$A_\pi^N := |\{(r(1), ..., r(n)) \hat{=} \pi : 1 \leq r(1), ..., r(n) \leq N\}|$$

for each partition $\pi \in \mathcal{P}(n)$. Then, we can express moment of $a_1 + \cdots + a_N$ as

$$\varphi\left((a_1 + \cdots + a_N)^n\right) = \sum_{\pi \in \mathcal{P}(n)} \kappa_\pi A_\pi^N.$$

It remains to check what value A_π^N and κ_π can have for each partition π of $\{1, ...n\}$. Now, let us $\pi \in \mathcal{P}(n)$ and denote $\pi = \{V_1, ..., V_s\}$ for $s = |\pi|$.

First, we assume that there exists $m \in \{1, ..., s\}$ such that $|V_m| = 1$. Then, there exists $r \in \{1, ..., n\}$ such that $V_m = \{r\}$, and thus, we get

$$\begin{aligned}
\kappa_\pi &= \varphi(a_{r(1)} \cdots a_r \cdots a_{r(n)}) \\
&= \varphi(a_r)\varphi(a_{r(1)} \cdots \check{a}_r \cdots a_{r(n)}) \\
&= 0
\end{aligned}$$

since $\varphi(a_r) = 0$ and a_n ($n \in \mathbb{N}$) are all independent (recall the equation 3.2). This implies that for all $m \in \{1, ..., s\}$, $|V_m| \geq 2$, and thus $s \leq n/2$.

Second, we have that

$$A_\pi^N = N(N-1)\cdots(N-s+1).$$

Since this number grows asymptotically like N^s for large N, we obtain

$$\begin{aligned}
\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\right)^n\right) &= \lim_{N\to\infty} \sum_{\pi \in \mathcal{P}(n)} \frac{A_\pi^N \kappa_\pi}{N^{n/2}} \\
&= \lim_{N\to\infty} \sum_{\pi \in \mathcal{P}(n)} N^{s-(n/2)} \kappa_\pi.
\end{aligned}$$

We just consider the following two cases because $s \leq n/2$: $s < n/2$ and $s = n/2$. If $s < n/2$, then $s - n/2 < 0$ which implies that

$$\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\right)^n\right) = 0.$$

Hence, we get $s = n/2$ and this means that for each $m \in \{1, ..., s\}$, $|V_m| = 2$, i.e., π is a pair partition of $\{1, ..., n\}$.

Therefore, we can express the result as

$$\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\right)^n\right) = \sum_{\pi\in\mathcal{P}_2(n)} \kappa_\pi. \tag{4.2}$$

It is also obvious that if n is odd, then the value of the limit is zero since $\{1, ..., n\}$ cannot have any pair partition for odd n. [4.1]

We now may assume that n is even. For $\pi \in \mathcal{P}_2(n)$, consider an index-tuple $((r(1), ..., r(n))$ with $((r(1), ..., r(n))\hat{=}\pi$. If all consecutive indicies are different, i.e., $r(1) \neq r(2), ..., r(n-1) \neq r(n)$, then since a_n $(n \in \mathbb{N})$ are all centered, we have $\kappa_\pi = 0$ by definition of freeness. Thus, we do not need to consider such case.

Next, we assume that there exist two consecutive indices $r(m)$ and $r(m+1)$ such that $r(m) = r(m+1) = r$ for some $m \in \{1, ..., n-1\}$. Then, by the equation 3.2, we have

$$\begin{aligned}
\kappa_\pi &= \varphi(a_{r(1)} \cdots a_r a_r \cdots a_{r(n)}) \\
&= \varphi(a_{r(1)} \cdots a_{r(m-1)} a_{r(m+2)} \cdots a_{r(n)})\varphi(a_r^2) \\
&= \sigma^2 \varphi(a_{r(1)} \cdots a_{r(m-1)} a_{r(m+2)} \cdots a_{r(n)}).
\end{aligned}$$

Since we can repeat it iteratively, we obtain that $\kappa_\pi = 0$ or $\kappa_\pi = \sigma^n$.

Then, when does κ_π have zero? The fact $\kappa_\pi = 0$ means that it happens that all consecutive indices are different during the above iteration. Thus, for some $l \in \mathbb{N}$, there exists $b_1, ..., b_l \in \{a_{r(1)}, ..., a_{r(n)}\}$ such that there is no pair of consecutive equal elements in $\{b_1, ..., b_l\}$. Pick any pair $(b_{k(1)}, b_{k(2)})$

[4.1] It should be noted that we did not use any property of free independence.

such that $b_{k(1)} < b_{k(2)}$. Then, we can find another pair $(b_{k(3)}, b_{k(4)})$ such that $b_{k(1)} < b_{k(3)} < b_{k(4)} < b_{k(2)}$ until there is no such pair in $\{b_1, ..., b_l\}$. Finally, we choose an element $a_0 \in \{a_{r(1),...,a_{r(n)}}\}$ such that $b_{k(3)} < a_0 < b_{k(4)}$. Otherwise, $b_{k(3)}$ and $b_{k(4)}$ must be consecutive and equal. Furthermore, there exits $a'_0 \in \{a_{r(1)}, ..., a_{r(n)}\}$ such that $a'_0 < b_{k(3)}$ or $b_{k(4)} < a'_0$, and a_0 is paired with a'_0. In brief, π should be crossing.

Thus, we represent the limit of moments as

$$\begin{aligned}\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\right)^n\right) &= \sum_{\pi \in NC_2(n)} \sigma^n \\ &= D_n \sigma^n\end{aligned}$$

where D_n is the number of non-corssing pair partitions of $\{1, ..., n\}$, i.e., $D_n = |NC_2(n)|$. By the lemma 2.7,

$$\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\right)^n\right) = \begin{cases} 0 & \text{if } n = 2k-1, k \in \mathbb{N} \\ C_k & \text{if } n = 2k, k \in \mathbb{N}. \end{cases}$$

Note that the density function of a semicircular random variable is even, and thus their all $(2k-1)$-th moments are zero. Therefore, by the proposition 2.5, the proof is complete. □

One can observe that any property of freeness does not used until we show that the given partition π is a pair partition of $\{1, ..., n\}$. Thus, in the framework of classical probability theory, we obtain the equation 4.2. Also, since all random varaibles are commutative in classical probability theory, we get $\kappa_\pi = \sigma^n$ in the similar way, and thus,

$$\begin{aligned}\lim_{N\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\right)^n\right) &= \sigma^n \, |\mathcal{P}_2(n)| \\ &= (n-1)(n-3)\cdots 5\cdot 3\cdot 1 \\ &= (n-1)!!.\end{aligned}$$

By the Wick's theorem, the mixed moments of standard normal random variable Z are given as

$$E[Z^n] = \begin{cases} 0 & \text{if } n \text{ is odd} \\ (n-1)!! & \text{if } n \text{ is even.} \end{cases}$$

Thus, we obtain the theorem 2.1.

Chapter 5

Conclusion

We have proved the free version of CLT by using non-crossing partitions and Catalan numbers in an algebraic way. Also, in the same way, we proved the classical CLT in the same framework. One can be aware that the difference between two CLTs is just the type of partitions of $\{1, ..., n\}$ that the summands κ_π contribute. In other words, in free CLT, moments are represented as sums over all non-crossing partitions whearas moments in classical CLT are written by sums over all partitions.

One can consider the multivariate version of free CLT. To prove this, we need to introduce the notion of joint distributions and joint moments in free probability theory. We can prove the multivraite version in a similar way, and the multivariate version of κ_π appears, which called the *free cumulants* of random variables. In the proof, we obtain the free version of *moment-cumulant formula*. In this case, we find that the difference between two versions of moment-cumulant formula is also whether moments are written by sums over all partitions or over all non-crossing partitions.

Appendix

Review of Classical Probability Theory

Definition A.1. A *probability space* in classical probability theory consists of a set Ω, which is called a *sample space*, a collection $\mathcal{F}$ of measurable sets in the sample space Ω, which is called a *σ-field* on the sample space Ω, and a *probability measure* P on a measurable space $(\Omega, \mathcal{F})$. We deonte it by $(\Omega, \mathcal{F}, P)$.

We assume that $(\Omega, \mathcal{F}, P)$ is a probability space throughout the appendix. Then, we can define a random variable on this probability space, and its distribution and expectation.

Definition A.2. A function $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called a *random variable* on the sample space Ω if X is mearuable where $\mathcal{B}(\mathbb{R})$ is the *Borel σ-field* on the real line $\mathbb{R}$ (i.e., the smallest σ-field containing the topology of $\mathbb{R}$).

Definition A.3. Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$.

(1) Define a function $F : \mathbb{R} \to [0, 1]$ as $F(x) = P(X \leq x)$ for all $x \in \mathbb{R}$. Then, we have the unique probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$\mu((a,b]) = F(b) - F(a)$ for $a, b \in \mathbb{R}$ with $a \leq b$. In this situation, Either F or μ is called the *distribution* of X.

(2) Also, the *expectation* EX of a random variable X is defined as the Lebesgue integral of X by P:

$$EX := \int X dP = \int_{\Omega} X(\omega) dP(\omega).$$

We can also express the expectation by using the distribution of random variable, i.e,

$$\begin{aligned} EX &= \int_{\Omega} X(\omega) dP(\omega) \\ &= \int_{\mathbb{R}} x d\mu(x) \\ &= \int_{-\infty}^{\infty} x dF(x). \end{aligned}$$

We now assume that all random variables in the appendix is defined on a probability space $(\Omega, \mathcal{F}, P)$. It must be noted that the operator E is linear as follows since the integration is a linear operator.

Proposition A.4. For all $a, b \in \mathbb{R}$ and random variables X and Y, we have

$$E[aX + bY] = aEX + bEY.$$

The notion of independence in classical probability theory can be expressed by using a probability measure in an analytic sense. We can define the concept of independence step by step as follows.

Definition A.5.

(1) Two measurable sets A and B in the σ-field $\mathcal{F}$ are said to be *independent* if

$$P(A \cap B) = P(A)P(B)$$

(2) Finite measurable sets $A_1, ..., A_n \in \mathcal{F}$ are said to be *independent* if for all $i_1, ..., i_k \in \{1, ..., n\}$ such that $i_1 < ... < i_k$,

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

(3) Measurable sets $A_i \in \mathcal{F}$ $(i \in \mathbb{N})$ are said to be *independent* if $A_1, ..., A_n \in \mathcal{F}$ are independent for all $n \geq 2$.

(4) Sub σ-fields $\mathcal{G}_i \subset \mathcal{F}$ $(i \in \mathbb{N})$ are said to be *independent* if for any $A_i \in \mathcal{G}_i$ $(i \in \mathbb{N})$, A_i $(i \in \mathbb{N})$ are independent.

(5) Random variables X_i $(i \in \mathbb{N})$ are said to be *independet* if $\sigma(X_i)$ $(i \in \mathbb{N})$ are independet.

We also know the following properties of independent random variables, but the converse is not true.

Proposition A.6. Assume that random variables X and Y on a probability space $(\Omega, \mathcal{F}, P)$ are independent. Then, $E[XY] = (EX) \cdot (EY)$.

In brief, the expectation operator is a linear operator on a "space" of random variables with the following property: for independent random variables X and Y, $EX = 0 = EY$ implies that $E[XY] = 0$.

Moreover, we need the notion of convergence in distribution for CLT.

Definition A.7. Let X and X_n ($n \in \mathbb{N}$) be random varaibles on $(\Omega, \mathcal{F}, P)$. Assume that X and X_n ($n \in \mathbb{N}$) have distribution functions F and F_n ($n \in \mathbb{N}$), i.e., $F(x) = P(X \leq x)$ and $F_n(x) = P(X_n \leq x)$ for all $x \in \mathbb{R}$. Then, let μ and μ_n ($n \in \mathbb{N}$) be the corresponding probability Borel measures to F and F_n ($n \in \mathbb{N}$).

(1) We say that $(\mu_n)_{n\in\mathbb{N}}$ *converges weakly* to μ if there exists a dense subset D of $\mathbb{R}$ such that for all $a, b \in D$ with $a < b$,

$$\lim_{n\to\infty} \mu_n\left((a, b]\right) = \mu\left((a, b]\right),$$

and we denote it by

$$\mu_n \xrightarrow[n\to\infty]{w} \mu.$$

(2) We say that $(X_n)_{n\in\mathbb{N}}$ *converges in distribution* to X if we have

$$\mu_n \xrightarrow[n\to\infty]{w} \mu,$$

or equivalently,

$$\lim_{n\to\infty} F_n(x) = F(x) \text{ for all } x \in \mathbb{R},$$

and we denote it by

$$X_n \xrightarrow[n\to\infty]{d} X.$$

We have the famous equivalent condition, which is called the *Portmanteau lemma*, and sometimes use this as the defnition of convergence in distribution.

Proposition A.8 (Portmanteau lemma)**.** Let X and X_n ($n \in \mathbb{N}$) be random varaibles on $(\Omega, \mathcal{F}, P)$, and μ and μ_n ($n \in \mathbb{N}$) be the corresponding probability Borel measures to X and X_n ($n \in \mathbb{N}$). Then, the followings are equivalent.

(1) $(X_n)_{n\in\mathbb{N}}$ converges in distribution to X;

(2) for any bounded continuous function $f : \mathbb{R} \to \mathbb{R}$, we have

$$\lim_{n\to\infty} \int f(t) d\mu_n(t) = \int f(t) d\mu(t),$$

or equivalently,

$$\lim_{n\to\infty} E\left[f(X_n)\right] = E\left[f(X)\right].$$

Bibliography

[1] D. Voiculescu, "Free probability theory: random matrices and von neumann algebras," in *Proceedings of the International Congress of Mathematicians*, vol. 1, p. 2, 1994.

[2] R. Durrett, *Probability: theory and examples*. Cambridge university press, 2010.

[3] A. Nica and R. Speicher, *Lectures on the combinatorics of free probability*, vol. 13. Cambridge University Press, 2006.

[4] J. B. Fraleigh, *A first course in abstract algebra*. Pearson Education India, 2003.

[5] M. Bhattacharjee, A. Bose, *et al.*, "Large sample behaviour of high dimensional autocovariance matrices," *The Annals of Statistics*, vol. 44, no. 2, pp. 598–628, 2016.

[6] R. Speicher, "Free probability and random matrices," *arXiv preprint arXiv:1404.3393*, 2014.

국문초록

고전확률론에서 확률변수 사이의 가환성을 가정하는 것은 매우 자연스럽다. 그러나 랜덤행렬의 집합과 같은 확률적 구조에서는 가환성을 만족하지 못 하고, 이를 위해 가환성을 제외한 확률적 구조를 만들 필요가 있다. 이때, 고전확률론에서의 많은 성질들은 이런 비가환 구조에서 새롭게 만들어진 "자유확률론"이라는 이론에서는 성립하지 않는다. 본 논문에서는 비가환 확률공간과 기존 확률론에서 독립성에 대응되는 "자유 독립성"의 개념을 대수적으로 정의하는 방법을 소개한다. 또한, 이런 비가환 확률공간에서의 새로운 개념과 조합론에서 사용되는 몇 가지 성질을 이용하여 자유확률론에서 중심극한정리를 증명하고 이를 기존의 중심극한정리와 비교해본다.

주요어 : 비가환 확률 공간, 자유 독립성, 중심 극한 정리, 무작위 행렬 이론, 조합론

학 번 : 2015-20304