The Free Central Limit Theorem: A Combinatorial Approach

by
Dennis Stauffer

A project submitted to the Department of Mathematical Sciences in conformity with the requirements for Math 4301 (Honour's Seminar)

Lakehead University
Thunder Bay, Ontario, Canada

Abstract

The free central limit theorem is a key result in free probability theory. In this work, we present a proof of the free central limit theorem. The theorem applies to freely independent random variables, which are non-commutative. The proof uses a combinatorial approach. We also show how the free central limit theorem is similar to the classic central limit theorem for classically independent random variables. In particular, we show that the role of the number of total pair partitions of a given set in the proof of the classic central limit theorem is analogous to the role of the number of non-crossing pair partitions of a given set in the proof of the free central limit theorem.

Acknowledgements

I would like to thank Dr. Grazia Viola for her advice, instruction, and detailed editorial comments in regards to this work. I would also like to thank Dr. Adam Van Tuyl for his comments, suggestions, and advice during the composition of this paper.

Contents

CHAPTER 1

Introduction

Free probability is a mathematical theory that is used to study non-commutative random variables. In some ways, free probability theory is similar to classic probability theory. In particular, the free independence of non-commutative random variables is comparable to the classic independence of commutative random variables. Free probability theory was first developed by Dan Voiculescu in 1986 as a tool to solve a problem dealing with free groups and operator algebras. Specifically, it is known that the free group with two generators, denoted by $\mathcal{F}_2$, is not isomorphic to the free group with three generators, denoted $\mathcal{F}_3$. We let $L(\mathcal{F}_2)$ and $L(\mathcal{F}_3)$ denote the von Neumann algebras associated with $\mathcal{F}_2$ and $\mathcal{F}_3$, respectively. Voiculescu attempted to answer the question of whether or not $L(\mathcal{F}_2)$ is isomorphic to $L(\mathcal{F}_3)$.

Voiculescu discovered that the generators of the free groups exhibited a special kind of independence, which he termed free independence. He also found that free independence was analogous to classic independence in many ways. For example, just as classic independence gives a rule for calculating mixed moments of classically independent random variables from the moments of the individual random variables, free independence gives a rule for calculating mixed moments of freely independent random variables from the moments of the individual random variables. Various applications of free probability to other areas of mathematics have been found. For example, random matrices are matrices whose entries are random variables. In 1955, Eugene Wigner showed that, subject to certain hypotheses, the limiting distribution of the eigenvalues of symmetric random $N \times N$ matrices is a semicircular distribution. The hypotheses on the random matrices are that the entries are classically independent, real, have the same first and second moments, and have the moments uniformly bounded by some constants B_n, i.e., for any $n \geq 1$, $\phi(a_{ij}^n) < B_n$ for all $1 < i, j < N$. Later on, Voiculescu demonstrated that, subject to certain conditions, random $N \times N$ matrices are asymptotically free independent. The conditions are specified as follows: The entries are complex functions $f(x)$ on some space $\mathcal{X}$ with the property that the integral of $|f(x)|^p$ on the space $\mathcal{X}$ with respect to some measure is finite for any $p \in \mathbb{N}$; the matrix is self-adjoint, i.e. for its entries, $a_{ij} = \overline{a_{ji}}$ for $i, j = 1, ..., N$; each entry has mean 0; the joint density of the family given by $\{\text{Re}(a_{ij}), \text{Im}(a_{ij})\}$ for each $1 \leq i, j \leq N$ is Gaussian; and $\phi(a_{ij}\overline{a_{ij}}) = 1/N$ for $1 \leq i, j \leq N$. Large random matrices are used in many contexts such as wireless communications, traffic flow models, and plane boarding models. For instance, a large random matrix modeling traffic flow could have columns corresponding to different times of day, rows corresponding to different locations, and entries representing the number of cars passing a location at a given time. Free probability theory could then be used to optimize traffic flow on a given route.

Naturally, Voiculescu also wondered whether or not some of the key results in classic probability theory had equivalent results in the context of free random variables. He found that the central limit theorem in classic probability theory had an analogue in free probability theory. He called the related result the free central limit theorem. In classic probability theory, the central limit theorem asserts that the limiting distribution of the sum of N independent, identically distributed random variables, each with mean 0, divided by $\sqrt{N}$ is the normal distribution. Similarly, the free central limit theorem asserts that the limiting distribution of the sum of N freely independent, identically distributed non-commutative random variables, each with mean 0, divided by $\sqrt{N}$ is the semicircular distribution. The normal distribution has the same role in classic probability that the semicircular distribution has in free probability.

To highlight these similarities, we prove both the free and classic central limit theorems using a combinatorial approach. In Chapter 2, we present some basic definitions and concepts in free probability. Specifically, we define a non-commutative probability space, classic independence, and free independence. In Chapter 3, we compute the moments of a normally distributed random variable and the moments of a semicircular random variable. In Chapter 4, we introduce partitions and pair partitions of the set $\{1,..,n\}$. We explain the difference between crossing and non-crossing pair partitions. In Chapter 5, we give proofs of the classic and free central limit theorems using a combinatorial approach. We show that the main difference in the proofs of the two theorems is that in the classic central limit theorem, we have to count the pair partitions of the set $\{1,...,n\}$, while in the free central limit theorem, we have to count the non-crossing pair partitions of this set. While proofs in classic probability use all pair partitions of a given set, proofs in free probability use non-crossing pair partitions.

CHAPTER 2

Preliminaries

1. Non-commutative Probability Space

A non-commutative probability space consists of an algebra and a linear functional that maps the elements of the algebra to complex numbers. This map is analogous to the map that we have in classic probability, which sends a random variable to the expected value of the random variable. In this section, we define the properties of an algebra and the moments of random variables. We also define and compare classic and free independence.

Definition 2.1. Let $\mathcal{A}$ be a set with three operations. The addition and multiplication of two elements a, b in $\mathcal{A}$ is denoted by $a + b$ and $a \cdot b$ respectively. For any complex number λ and any element a in $\mathcal{A}$, we denote the scalar multiplication of a by λ as $a \cdot \lambda$. Assume also that for $a, b, c \in \mathcal{A}$ and scalars $\lambda_1, \lambda_2 \in \mathbb{C}$ the following holds:

(1) $(a + b) + c = a + (b + c)$ (associative property)
(2) $a + b = b + a$ (commutative property)
(3) $0 + a = a + 0 = a$
(4) $a + (-a) = -a + a = 0$

(5) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associative property)
(6) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ (distributive property)

(7) $(a \cdot \lambda_1) \cdot \lambda_2 = a \cdot (\lambda_1 \cdot \lambda_2)$ (associative property)
(8) $(a + b) \cdot \lambda = (a \cdot \lambda) + (b \cdot \lambda)$ (distributive property)

Then we call $\mathcal{A}$ an **algebra**. If $\mathcal{A}$ also has a multiplicative identity element 1 such that $a \cdot 1 = 1 \cdot a = a$, then $\mathcal{A}$ is a **unital algebra**.

Definition 2.2. Let $\mathcal{A}$ be an algebra. Let $\mathcal{A}_i$ be a subset of $\mathcal{A}$ that is closed under the operations of $\mathcal{A}$. Then $\mathcal{A}_i$ is a **subalgebra** of $\mathcal{A}$.

Definition 2.3. An algebra $\mathcal{A}$ is a $*$-algebra if it is closed under a $*$-operation satisfying the following properties:

(1) $(a + b)^* = a^* + b^*$,
(2) $(\lambda a)^* = \bar{\lambda} a^*$,

(3) $(ab)^* = b^*a^*$,
(4) $a^{**} = (a^*)^* = a$.

Remark 2.4. If an element $a \in \mathcal{A}$ has the property that $a^* = a$, then we say that a is **self-adjoint**.

Example 2.5. An example of a unital $*$-algebra is the algebra of complex numbers $\mathbb{C}$. All the axioms of a unital algebra can be readily verified if the elements are complex numbers and the $*$-operation is the conjugate of a complex number. Both multiplication between elements and multiplication by a scalar coincide with the multiplication of complex numbers.

Example 2.6. A more interesting example of a $*$-algebra is the algebra of $n \times n$ matrices with complex entries, denoted $M_n(\mathbb{C})$. Addition, multiplication, and scalar multiplication are defined as matrix addition, matrix multiplication, and multiplication of a matrix by a scalar. The $*$-operation is equivalent to taking the transpose of a matrix and the conjugate of its entries. We note that matrix multiplication is usually not commutative. In other words, for $n \times n$ matrices A and B, we usually have $AB \neq BA$.

Example 2.7. The set of all continuous functions on the interval $[0,1]$, denoted $\mathrm{C}[0,1]$, is an algebra. We define the addition and the multiplication of two elements in $\mathrm{C}[0,1]$ as function addition and function multiplication, respectively. Scalar multiplication of an element in $\mathrm{C}[0,1]$ is defined as multiplication of a function by a scalar, and is denoted $(\lambda \cdot f)$, where $f \in \mathrm{C}[0,1]$ and $\lambda \in \mathcal{C}$.

Definition 2.8. Let $\mathcal{A}$ be an algebra. A map $\phi : \mathcal{A} \to \mathbb{C}$ is called a **functional**. We say that ϕ a **linear functional** if, for any two elements $a, b \in \mathcal{A}$ and scalars λ_1, λ_2, we have

$$\phi(\lambda_1 a + \lambda_2 b) = \lambda_1 \phi(a) + \lambda_2 \phi(b).$$

Example 2.9. A common example of a linear functional is the integration of continuous functions on the interval $[0,1]$. We define $\phi : \mathrm{C}[0,1] \to \mathbb{C}$ by $\phi(f(x)) = \int_0^1 f(x)\mathrm{dx}$ for any $f(x) \in \mathrm{C}[0,1]$.

Definition 2.10. A **non-commutative probability space** $(\mathcal{A}, \phi)$ consists of a unital algebra $\mathcal{A}$ and a linear functional $\phi : \mathcal{A} \to \mathbb{C}$, such that $\phi(1_{\mathcal{A}}) = 1$. Any element $a \in \mathcal{A}$ is called a **free random variable**.

Definition 2.11. Let $(\mathcal{A}, \phi)$ be a non-commutative probability space, and let $a \in \mathcal{A}$ be a random variable. Then the n-th **moment** of a is defined as $\phi(a^n)$.

Definition 2.12. A **mixed moment** is any expression of the form $\phi(x)$ where x is a product of two or more random variables. For example, let $a, b \in \mathcal{A}$ be distinct random variables. Then for non-negative integers n and m, $\phi(a^m b^n)$ is a mixed moment.

Definition 2.13. Let $(\mathcal{A}, \phi)$ be a non-commutative probability space. If $\mathcal{A}$ is a $*$-algebra, and $\phi(a^*a) \geq 0$ for all a in $\mathcal{A}$, then ϕ is said to be **positive** and $(\mathcal{A}, \phi)$ is called a $*$-**probability space**.

Example 2.14. The algebra $\mathrm{M}_n(\mathbb{C})$ of $n \times n$ matrices with complex entries together with the linear functional $\phi : \mathrm{M}_n(\mathbb{C}) \to \mathbb{C}$, given by the **normalized trace** of a matrix, is a $*$-probability space. The normalized trace of a matrix in $\mathrm{M}_n(\mathbb{C})$ is the sum of the complex numbers in the main diagonal divided by n. To see that ϕ is positive, we take an arbitrary matrix A in $\mathrm{M}_n(\mathbb{C})$ where $A = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ \vdots & & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{pmatrix}$ and $z_{i,j} \in \mathbb{C}$ for $1 \leq i, j \leq n$. Then, $A^* = \begin{pmatrix} \overline{z_{11}} & \overline{z_{21}} & \cdots & \overline{z_{n1}} \\ \vdots & & \ddots & \vdots \\ \overline{z_{1n}} & \overline{z_{2n}} & \cdots & \overline{z_{nn}} \end{pmatrix}$. If we write $z \cdot \bar{z} = |z|^2$, then we have that

$$A^*A = \begin{pmatrix} |z_{11}|^2+|z_{12}|^2+\cdots+|z_{1n}|^2 & \vdots & \\ & |z_{21}|^2+|z_{22}|^2+\cdots+|z_{2n}|^2 & \vdots \\ \vdots & \vdots & \ddots \\ & & |z_{n1}|^2+|z_{n2}|^2+\cdots+|z_{nn}|^2 \end{pmatrix}. \tag{2.1}$$

Applying the functional ϕ, which takes the normalized trace of A^*A, we have

$$\phi(A^*A) = \frac{1}{n}\left(|z_{11}|^2 + |z_{12}|^2 + \cdots + |z_{nn}|^2 \right) = \frac{1}{n}\left(\sum_{1 \leq i,j \leq n} |z_{i,j}|^2 \right). \tag{2.2}$$

Because all the terms in the above sum are non-negative, the sum is non-negative. Therefore, $\phi(A^*A) \geq 0$, so $\mathrm{M}_n(\mathbb{C})$ is a $*$-probability space.

2. Classic Independence

Classically independent random variables are commutative with respect to multiplication. In contrast, free random variables are not commutative. In this section, we define classic independence and compute some mixed moments of classically independent random variables. We also show that our definition of classic independence is equivalent to the common definition used in classic probability theory.

Definition 2.15. Let $(\mathcal{B}, \phi)$ be a commutative probability space, and let $b_1, b_2, b_3, ..., b_n$ in $\mathcal{B}$ be distinct random variables. Then $b_1, b_2, b_3, ..., b_n$ are (classically) **independent** if $b_i b_j = b_j b_i$ for $1 \leq i, j \leq n$, and

$$\phi(b_1 b_2 b_3 \cdots b_n) = 0 \tag{2.3}$$

whenever $\phi(b_i) = 0$ for $1 \leq i \leq n$.

Definition 2.16. The element $b^0 = b - \phi(b) \cdot 1$ has the property that $\phi(b^0) = 0$. To see this, we note that $\phi(b^0) = \phi(b - \phi(b) \cdot 1) = \phi(b) - \phi(b) = 0$. We refer to b^0 as the **centering** of b in $\mathcal{B}$.

Remark 2.17. The above definition of classical independence is equivalent to the more common definition, which for simplicity we write only in the case that $n = 2$:

$$\phi(b_1 b_2) = \phi(b_1)\phi(b_2) \qquad \text{for } b_1, b_2 \in \mathcal{B}.$$

To show that the two definitions are equivalent, assume that $\phi(b_1) = \phi(b_2) = 0$. If $\phi(b_1 b_2) = \phi(b_1)\phi(b_2)$, then we have that $\phi(b_1 b_2) = 0$.
Conversely, assume that $\phi(b_1 b_2) = 0$ whenever $\phi(b_1) = \phi(b_2) = 0$. Since $b = b^0 + \phi(b) \cdot 1$, then we have that

$$\begin{aligned} \phi(b_1 b_2) &= \phi((b_1^0 + \phi(b_1) \cdot 1)(b_2^0 + \phi(b_2) \cdot 1)) \\ &= \phi(b_1^0 b_2^0 + \phi(b_2) b_1^0 + \phi(b_1) b_2^0 + \phi(b_1)\phi(b_2) \cdot 1) \\ &= \phi(b_1^0 b_2^0) + \phi(b_2)\phi(b_1^0) + \phi(b_1)\phi(b_2^0) + \phi(b_1)\phi(b_2) \\ &= \phi(b_1)\phi(b_2). \end{aligned} \tag{2.4}$$

where we have used that $\phi(b_1^0) = \phi(b_2^0) = 0$ and $\phi(b_1^0 b_2^0) = 0$ by equation (2.3).
We use the former definition to illustrate the similarity between classic independence and free independence.

3. Free Independence

In this section, we define free independence and compute mixed moments of free random variables. The non-commutativity of free independent random variables makes computations of moments more complicated than for classically independent random variables. The mixed moments of free and classically independent random variables are identical for the product of three or less random variables. However, products involving four or more random variables produce different moments for the two types of random variables.

Definition 2.18. Let $(\mathcal{A}, \phi)$ be a non-commutative probability space and let I be a fixed index set. For each $i \in I$, let $\mathcal{A}_i \subset \mathcal{A}$ be a subalgebra. The subalgebras $(\mathcal{A}_i)_{i \in I}$ are called **freely independent** if

$$\phi(a_1 a_2 a_3 \cdots a_k) = 0$$

whenever we have the following:
(1) k is a positive integer
(2) $a_j \in \mathcal{A}_{i(j)}$ with $i(j) \in I$ for all $j = 1, 2, ..., k$
(3) $\phi(a_j) = 0$ for all $j = 1, 2, ..., k$

(4) neighbouring elements are from different subalgebras,
i.e. $i(1) \neq i(2), i(2) \neq i(3), ..., i(k-1) \neq i(k)$.

We can think of free independence as a rule for computing mixed moments. Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ be freely independent algebras, and let $a \in \mathcal{A}$, $b \in \mathcal{B}$, and $c \in \mathcal{C}$ be free random variables. Using the definition of free independence, we have

$$\begin{aligned} 0 &= \phi(a^0 b^0) \\ &= \phi((a - \phi(a)\cdot 1)(b - \phi(b)\cdot 1)) \\ &= \phi(ab - a\phi(b) - b\phi(a) + \phi(a)\phi(b)\cdot 1) \\ &= \phi(ab) - 2\phi(a)\phi(b) + \phi(a)\phi(b) \\ &= \phi(ab) - \phi(a)\phi(b). \end{aligned} \tag{2.5}$$

Therefore, we have $\phi(ab) = \phi(a)\phi(b)$.
To compute a mixed moment with three random variables, we have

$$\begin{aligned} 0 &= \phi(a^0 b^0 a^0) \\ &= \phi((a - \phi(a)\cdot 1)(b - \phi(b)\cdot 1)(a - \phi(a)\cdot 1)) \\ &= \phi(aba - ab\phi(a) - a^2\phi(b) + a\phi(a)\phi(b) - ab\phi(a) + b\phi(a)^2 \\ &\quad + a\phi(a)\phi(b) - \phi(a)^2\phi(b)\cdot 1) \\ &= \phi(aba) - \phi(ab)\phi(a) - \phi(a^2)\phi(b) + \phi(a)^2\phi(b) - \phi(ab)\phi(a) \\ &\quad + \phi(b)\phi(a)^2 + \phi(a)^2\phi(b) - \phi(a)^2\phi(b) \\ &= \phi(aba) - \phi(b)\phi(a^2) - 3\phi(a)^2\phi(b) + 3\phi(a)^2\phi(b) \\ &= \phi(aba) - \phi(b)\phi(a^2) \end{aligned} \tag{2.6}$$

where we have used (2.5) in step (5). Therefore, we have $\phi(aba) = \phi(b)\phi(a^2)$. Replacing the second a with the random variable $c \in \mathcal{C}$ and performing the same computation gives the result $\phi(abc) = \phi(b)\phi(ac)$.

Both of these mixed moment calculations involving free independent random variables give the same result as with classically independent random variables. However, if we calculate a mixed moment of a product of four random variables, we obtain

$$\phi(abab) = \phi(b)^2\phi(a^2) + \phi(a)^2\phi(b^2) - \phi(a)^2\phi(b)^2. \tag{2.7}$$

We compare this result to the classic moment. When a and b are classically independent, they commute. Consequently,

$$\phi(abab) = \phi(a^2b^2) = \phi(a^2)\phi(b^2). \tag{2.8}$$

CHAPTER 3

Distributions and Moments of Random Variables

Our proofs of the central limit theorems involve calculating the moments of normal or semicircular random variables. In this chapter, we define the distributions of the standard normal and semicircular random variable, and derive the formulas for the moments of these random variables.

1. The Normally Distributed Random Variable and its Moments

In this section, we give a formula for calculating the moments of a normally distributed random variable with mean 0 and variance σ^2. We prove this formula by induction. In the proof of the classic central limit theorem, we will use the fact that the n-th moment of a standard normal random variable (with a variance of 1) corresponds to the number of pair partitions of a set of n elements.

Definition 3.1. A **probability density function** is a function $f(t)$ such that, for all t in $\mathbb{R}$,

1. $f(t) \geq 0$
2. $\displaystyle\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$
3. $P[a \leq t \leq b] = \displaystyle\int_{a}^{b} f(t)\mathrm{d}t.$

where $P[a \leq t \leq b]$ is the probability that t assumes a value between a and b.

Definition 3.2. The normal probability density function is

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2\sigma^2} \qquad \text{for} \quad -\infty < t < \infty \tag{3.1}$$

with a mean of 0 and a variance of σ^2.

Definition 3.3. Let $(\mathcal{A}, \phi)$ be a non-commutative probability space and $T \in \mathcal{A}$. We say that T is a normally distributed random variable with mean 0 and variance σ^2 if the

n-th moment of T is given by

$$\phi(T^n) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t^n e^{-t^2/2\sigma^2}\,\mathrm{d}t. \tag{3.2}$$

Lemma 3.4. *The n-moments of the normally distributed random variable T are given by:*

$$\phi(T^n) = \begin{cases} 0 & \text{for } n \text{ odd} \\ (n-1)!!\sigma^n & \text{for } n \text{ even} \end{cases} \tag{3.3}$$

where $(n-1)!! = (n-1)(n-3)(n-5)\cdots 5\cdot 3\cdot 1$.

Proof. We use a proof by induction. For $n = 0$, we have that $\phi(T^0) = 1$ by definition of probability distribution. Since $(-1)!!$ is defined to be 1, then $(-1)!!\sigma^0 = 1$. For $n = 1$, we set $u = \frac{t^2}{2\sigma^2}$. Then, $t\,\mathrm{d}t = \sigma^2\,\mathrm{d}u$, and

$$\begin{aligned}
\phi(T) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} te^{-t^2/2\sigma^2}\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{0} te^{-t^2/2\sigma^2}\,\mathrm{d}t + \frac{1}{\sqrt{2\pi}\sigma}\int_{0}^{\infty} te^{-t^2/2\sigma^2}\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}\sigma}\lim_{a\to-\infty}\int_{\frac{a^2}{2\sigma^2}}^{0} e^{-u}\sigma^2\mathrm{d}u + \frac{1}{\sqrt{2\pi}\sigma}\lim_{b\to\infty}\int_{0}^{\frac{b^2}{2\sigma^2}} e^{-u}\sigma^2\mathrm{d}u \\
&= \frac{\sigma}{\sqrt{2\pi}}\lim_{a\to-\infty}\left(-1+e^{-a^2/2\sigma^2}\right) + \frac{\sigma}{\sqrt{2\pi}}\lim_{b\to\infty}\left(-e^{-b^2/2\sigma^2}+1\right) \\
&= \frac{\sigma}{\sqrt{2\pi}}(-1+1) = 0.
\end{aligned}$$

So the formula holds for $n = 0$ and $n = 1$.

For $n \geq 1$, we assume that $\phi(T^{n-1}) = \begin{cases} 0 & \text{if } n \text{ is even} \\ (n-2)!!\sigma^{n-1} & \text{if } n \text{ is odd} \end{cases}$

and show that $\phi(T^{n+1}) = \begin{cases} 0 & \text{if } n \text{ is even} \\ n!!\sigma^{n+1} & \text{if } n \text{ is odd.} \end{cases}$

We set $u = t^n$ and $\mathrm{d}v = te^{-t^2/2\sigma^2}\,\mathrm{d}t$. Then, $\mathrm{d}u = nt^{n-1}\mathrm{d}t$ and $v = -\sigma^2 e^{-t^2/2\sigma^2}$, and

$$
\begin{aligned}
\phi(T^{n+1}) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t^{n+1}e^{-t^2/2\sigma^2}\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}\sigma}\left(\lim_{a\to-\infty}\int_a^0 t^{n+1}e^{-t^2/2\sigma^2}\,\mathrm{d}t + \lim_{b\to\infty}\int_0^b t^{n+1}e^{-t^2/2\sigma^2}\,\mathrm{d}t\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma}\lim_{a\to-\infty}\left(-\sigma^2 t^n e^{-t^2/2\sigma^2}\Big|_a^0 + \int_a^0 \sigma^2 e^{-t^2/2\sigma^2}(nt^{n-1})\,\mathrm{d}t\right) \\
&\quad + \frac{1}{\sqrt{2\pi}\sigma}\lim_{b\to\infty}\left(-\sigma^2 t^n e^{-t^2/2\sigma^2}\Big|_0^b + \int_0^b \sigma^2 e^{-t^2/2\sigma^2}(nt^{n-1})\,\mathrm{d}t\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma}\sigma^2 n\left(\lim_{a\to-\infty}\int_a^0 t^{n-1}e^{-t^2/2\sigma^2}\,\mathrm{d}t + \lim_{b\to\infty}\int_0^b t^{n-1}e^{-t^2/2\sigma^2}\,\mathrm{d}t\right) \\
&= \sigma^2 n\left(\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t^{n-1}e^{-t^2/2\sigma^2}\,\mathrm{d}t\right) = \sigma^2 n\Big(\phi(T^{n-1})\Big).
\end{aligned}
$$

By our assumption, we have that

$$
\phi(T^{n+1}) = \sigma^2 n\Big(\phi(T^{n-1})\Big) = \begin{cases} 0 & \text{if } n \text{ is even} \\ \sigma^2 n(n-2)!!\sigma^{n-1} = n!!\sigma^{n+1} & \text{if } n \text{ is odd.} \end{cases}
$$

Thus, the result follows by induction. □

2. The Semicircular Random Variable and its Moments

We derive a formula for calculating the moments of a standard semicircular random variable. We prove the formula by induction. We show that the $2k$-th moment of the semicircular random variable is the Catalan number C_k, which is also the number of non-crossing pair partitions of a set of $2k$ elements.

Definition 3.5. The standard semicircular probability density function is

$$
f(t) = \begin{cases} \frac{1}{2\pi}\sqrt{4-t^2} & \text{for } -2 \le t \le 2 \\ 0 & \text{otherwise} \end{cases} \tag{3.4}
$$

with a mean of 0 and a variance of 1.

Definition 3.6. Let $(\mathcal{A}, \phi)$ be a non-commutative probability and $T \in \mathcal{A}$. We say that T is a standard semicircular random variable if the n-th moment of T is given by

$$\phi(T^n) = \frac{1}{2\pi}\int_{-2}^{2} t^n \sqrt{4-t^2}\,\mathrm{d}t. \tag{3.5}$$

Lemma 3.7. *The n-moments of the standard semicircular variable T are given by:*

$$\phi(T^n) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \dfrac{1}{k+1}\dbinom{2k}{k} & \text{if } n = 2k \text{ for some } k. \end{cases} \tag{3.6}$$

Proof. We use a proof by induction. For $n = 0$, we have that $\phi(T^0) = 1$ by definition of probability distribution. Since $\binom{0}{0} = \frac{0!}{0!0!} = 1$, then $\frac{1}{0+1}\binom{0}{0} = 1$.
For $n = 1$, we set $u = 4 - t^2$. Then, $t\,\mathrm{d}t = \frac{\mathrm{d}u}{-2}$ and

$$\phi(t) = \frac{1}{2\pi}\int_{-2}^{2} t\sqrt{4-t^2}\,\mathrm{d}t = \frac{-1}{4\pi}\int_{0}^{0}\sqrt{u}\,\mathrm{d}u = 0.$$

The formula holds for $n = 0$ and $n = 1$.

For $n \geq 1$, we assume that $\phi(T^{n-1}) = \begin{cases} 0 & \text{if } n \text{ is even} \\ \dfrac{1}{k+1}\dbinom{2k}{k} & \text{if } n = 2k+1 \end{cases}$

and show that $\phi(t^{n+1}) = \begin{cases} 0 & \text{if } n \text{ is even} \\ \dfrac{1}{k+2}\dbinom{2k+2}{k+1} & \text{if } n = 2k+1. \end{cases}$

In steps (6) and (8) in the following computation, we use the identity

$$\int_{-\pi/2}^{\pi/2} \sin^n(u)\,\mathrm{d}u = -\frac{1}{n}\sin^{n-1}(u)\cos(u)\Big|_{-\pi/2}^{\pi/2} + \frac{n-1}{n}\int_{-\pi/2}^{\pi/2}\sin^{n-2}(u)\,\mathrm{d}u. \tag{3.7}$$

We set $t = 2\sin(u)$. Then $\mathrm{d}t = 2\cos(u)\,\mathrm{d}u$ and

$$\begin{aligned}
\phi(T^{n+1}) &= \frac{1}{2\pi}\int_{-2}^{2} t^{n+1}\sqrt{4-t^2}\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} 2^{n+1}\sin^{n+1}(u)\sqrt{4(1-\sin^2(u))}\,(2\cos(u))\,\mathrm{d}u \\
&= \frac{2^{n+3}}{2\pi}\int_{-\pi/2}^{\pi/2} \sin^{n+1}(u)\cos^2(u)\,\mathrm{d}u = \frac{2^{n+3}}{2\pi}\int_{-\pi/2}^{\pi/2} \sin^{n+1}(u)(1-\sin^2(u))\,\mathrm{d}u \\
&= \frac{2^{n+3}}{2\pi}\left(\int_{-\pi/2}^{\pi/2} \sin^{n+1}(u)\,\mathrm{d}u - \int_{-\pi/2}^{\pi/2} \sin^{n+3}(u)\,\mathrm{d}u\right) = \frac{2^{n+3}}{2\pi}\left(\int_{-\pi/2}^{\pi/2} \sin^{n+1}(u)\,\mathrm{d}u + \right. \\
&\qquad \left. -\left(-\frac{1}{n+3}\sin^{n+2}(u)\cos(u)\,\Big|_{-\pi/2}^{\pi/2} + \frac{n+2}{n+3}\int_{-\pi/2}^{\pi/2} \sin^{n+1}(u)\,\mathrm{d}u\right)\right) \\
&= \frac{2^{n+3}}{2\pi}\left(\frac{1}{n+3}\right)\int_{-\pi/2}^{\pi/2} \sin^{n+1}(u)\,\mathrm{d}u \\
&= \frac{2^{n+3}}{2\pi}\left(\frac{1}{n+3}\right)\left(\frac{n}{n+1}\right)\int_{-\pi/2}^{\pi/2} \sin^{n-1}(u)\,\mathrm{d}u \\
&= \frac{4n}{n+3}\left(\frac{2^{n+1}}{2\pi}\right)\left(\frac{1}{n+1}\right)\int_{-\pi/2}^{\pi/2} \sin^{n-1}(u)\,\mathrm{d}u.
\end{aligned}$$

Similarly, using equation (3.7) in step (6), we have that

$$\begin{aligned}
\phi(T^{n-1}) &= \frac{1}{2\pi}\int_{-2}^{2} t^{n-1}\sqrt{4-t^2}\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} 2^{n-1}\sin^{n-1}(u)\sqrt{4(1-\sin^2(u))}\,(2\cos(u))\,\mathrm{d}u \\
&= \frac{2^{n+1}}{2\pi}\int_{-\pi/2}^{\pi/2} \sin^{n-1}(u)\cos^2(u)\,\mathrm{d}u \\
&= \frac{2^{n+1}}{2\pi}\int_{-\pi/2}^{\pi/2} \sin^{n-1}(u)(1-\sin^2(u))\,\mathrm{d}u \\
&= \frac{2^{n+1}}{2\pi}\left(\int_{-\pi/2}^{\pi/2} \sin^{n-1}(u)\,\mathrm{d}u - \int_{-\pi/2}^{\pi/2} \sin^{n+1}(u)\,\mathrm{d}u\right) \\
&= \frac{2^{n+1}}{2\pi}\left(\int_{-\pi/2}^{\pi/2} \sin^{n-1}(u)\,\mathrm{d}u - \frac{n}{n+1}\int_{-\pi/2}^{\pi/2} \sin^{n-1}(u)\,\mathrm{d}u\right)
\end{aligned}$$

$$= \frac{2^{n+1}}{2\pi}\left(\frac{1}{n+1}\right)\int_{-\pi/2}^{\pi/2}\sin^{n-1}(u)\,du.$$

Comparing these two results, we have that $\phi(T^{n+1}) = \dfrac{4n}{n+3}\left(\phi(T^{n-1})\right)$. If n is odd, then by our assumption, and taking $n-1 = 2k$, we have

$$\begin{aligned}
\frac{4n}{n+3}\left(\phi(T^{n-1})\right) &= \frac{4(2k+1)}{2k+4}\left(\frac{1}{k+1}\right)\binom{2k}{k} \\
&= \frac{4(2k+1)}{2k+4}\cdot\frac{2}{2k+2}\binom{2k}{k} \\
&= \frac{2}{2k+4}\cdot\frac{4(2k+1)}{2k+2}\cdot\frac{(2k)!}{k!k!} \\
&= \frac{1}{k+2}\cdot\frac{(2k+1)(2k+2)}{(k+1)(k+1)}\cdot\frac{(2k)!}{k!k!} \\
&= \frac{1}{k+2}\cdot\frac{(2k+2)!}{(k+1)!(k+1)!} \\
&= \frac{1}{k+2}\binom{2k+2}{k+1}.
\end{aligned}$$

Therefore, we have that

$$\phi(t^{n+1}) = \frac{4n}{n+3}\left(\phi(t^{n-1})\right) = \begin{cases} 0 & \text{if } n \text{ is even} \\ \dfrac{1}{k+2}\dbinom{2k+2}{k+1} & \text{if } n = 2k+1. \end{cases}$$

□

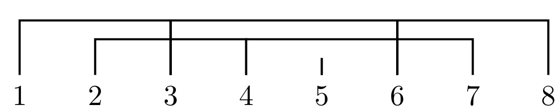
CHAPTER 4

Pair Partitions and Non-Crossing Pair Partitions

The number of pair partitions and non-crossing pair partitions of the set $\{1, 2, ..., k\}$ is a key element in our proofs of the central limit theorems. In this chapter, we define and illustrate non-crossing and crossing pair partitions. We also obtain formulas for the number of pair paritions and non-crossing pair partitions of a given set. We show that the number of non-crossing pair partitions of a set of $2k$ elements is equal to the Catalan number C_k.

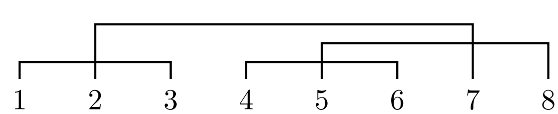
Definition 4.1. The set $\{V_1, V_2, ..., V_n\}$ is a **partition** of the set $\{1, ..., k\}$ if each block V_i for $1 \leq i \leq n$ is a subset of $\{1, ..., k\}$, $V_i \cap V_j = \emptyset$ for $1 \leq i, j \leq n$, and $\displaystyle\bigcup_{i=1}^{n} V_i = \{1, ..., k\}$. If a block V_i for some $1 \leq i \leq n$ of the partition contains only one element, we say that the partition has a **singleton**.

Example 4.2. Given the set $\{1, 2, 3, 4, 5, 6, 7, 8\}$, we define a partition $\pi = \{\{1, 3, 6, 8\}, \{2, 4, 7\}, \{5\}\}$. Then, $\pi = \{V_1, V_2, V_3\}$, where $V_1 = \{1, 3, 6, 8\}, V_2 = \{2, 4, 7\}$, and $V_3 = \{5\}$. We note that π has a singleton. The partition π is visually represented by the figure below. The elements that belong to the same block of the partition are connected.



Definition 4.3. We say that a set $\{V_1, V_2, ..., V_k\}$ is a **pair partition** of the set $\{1, ..., 2k\}$ if each V_i for $1 \leq i \leq k$ consists of two distinct elements of the set $\{1, ..., 2k\}$.

Example 4.4. Given the set $\{1, 2, 3, 4, 5, 6, 7, 8\}$, $\pi = \{\{1, 3\}, \{2, 7\}, \{6, 4\}, \{5, 8\}\}$ is a pair partition. In the representation of this pair partition below, vertical lines from some elements cross horizontal lines joining other elements.

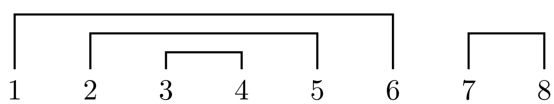


Lemma 4.5. *The number of pair partitions of the set* $\{1, 2, ..., k\}$ *is* $(k-1)!!$.

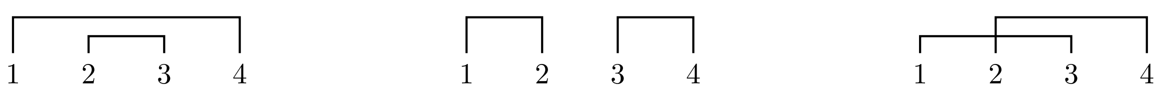
PROOF. Let $A = \{1, 2, .., k\}$. Given an element $j \in A$, there are $k-1$ possible distinct pairings of j with an element of A. For each of these pairings, there are $k-3$ possible pairings of another element $l \in A$ with an element of A excluding j and the element with which j has been paired. Repeating this argument until we exhaust the elements of A, we see that there are $(k-1)(k-3)\cdots 5 \cdot 3 \cdot 1 = (k-1)!!$ possible distinct pair partitions of A. □

Definition 4.6. Let $\pi = \{V_1, V_2, ..., V_k\}$ be a pair partition of the set $\{1, ..., 2k\}$. Let $V_m = \{p_1, p_2\}$ and let $V_n = \{q_1, q_2\}$ be blocks of π for any $1 \leq m, n \leq k$ and $p_1, p_2, q_1, q_2 \in \{1, ..., 2k\}$. Then $\{V_1, V_2, ..., V_k\}$ is a **non-crossing pair partition** if we never have $p_1 < q_1 < p_2 < q_2$ or $q_1 < p_1 < q_2 < p_2$. If neither of these conditions are verified, we say that the pair partition is **crossing**.

Example 4.7. An example of a non-crossing pair partition of the set $\{1, 2, 3, 4, 5, 6, 7, 8\}$, illustrated below, is $\pi = \{\{1, 6\}, \{2, 5\}, \{3, 4\}, \{7, 8\}\}$. We note that none of the vertical lines from elements of the pair partition cross any horizontal line joining other elements of the pair partition. The pair partition in the previous example is a crossing pair partition.

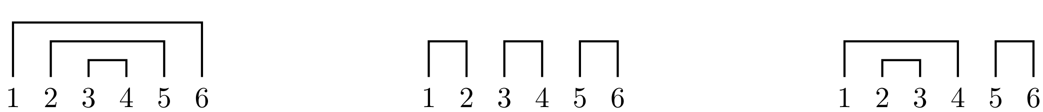


Visually, we can count the number of non-crossing pair partitions for small values of $2k$. For $k = 2$, there are only three possible pair partitions: (a) $\{\{1, 4\}, \{2, 3\}\}$, (b) $\{\{1, 2\}, \{3, 4\}\}$, and (c) $\{\{1, 3\}, \{2, 4\}\}$.

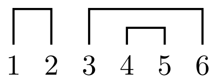


(a) (b) (c)

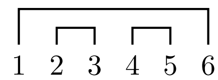
We see that pair partitions (a) and (b) are non-crossing, while (c) is crossing. Thus, there are 2 non-crossing partitions for a set of 4 elements. Similarly, for $k = 3$, there are $6!! = 15$ total pair partitions: 10 crossing and 5 non-crossing. The non-crossing pair partitions are: (a) $\{\{1, 6\}, \{2, 5\}, \{3, 4\}\}$, (b) $\{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$, (c) $\{\{1, 4\}, \{2, 3\}, \{5, 6\}\}$, (d) $\{\{1, 2\}, \{3, 6\}, \{4, 5\}\}$, and (e) $\{\{1, 6\}, \{2, 3\}, \{4, 5\}\}$.



(a) (b) (c)



(d)



(e)

The number of non-crossing pair partitions of the set $\{1, ..., 2k\}$ has a central role in our proof of the free central limit theorem. We obtain a formula for the number of non-crossing pair partitions of the set $\{1, ..., 2k\}$ based on the Catalan numbers.

Definition 4.8. The k-th **Catalan number** is given by the formula

$$C_k = \frac{1}{k+1}\binom{2k}{k}, \quad \text{for } k \geq 0. \tag{4.1}$$

The Catalan numbers can also be defined recursively as

$$C_0 = 1, \; C_1 = 1, \qquad C_{k+1} = \sum_{i=0}^{k} C_i C_{k-i}, \quad \text{for } n \geq 1. \tag{4.2}$$

The first seven Catalan numbers are 1, 1, 2, 5, 14, 42, and 132. Referring to Lemma 3.7, we see that the Catalan numbers are exactly the moments of the standard semicircular distribution.

Lemma 4.9. *The number D_{2k} of non-crossing pair partitions of the set $\{1, ..., 2k\}$ is given by the Catalan numbers C_k.*

Proof. Let $\pi = \{V_1, V_2, ..., V_k\}$ be a pair partition of the set $\{1, ..., 2k\}$, and let D_{2k} denote the number of non-crossing pair partitions of this set. Let V_1 be the block containing 1, so that $V_1 = \{1, m\}$ for some $1 < m \leq 2k$. Since the partition is non-crossing, for any other block $V_j = \{p, q\}$ of π with $2 \leq p < q \leq 2k$ and $1 < j \leq k$, we cannot have $1 < p < m < q$. Therefore, any other block V_j for $1 < j \leq k$ must be contained in either $\{2, ..., m-1\}$ or $\{m+1, ..., 2k\}$. Since each element in the set $\{2, ..., m-1\}$ belongs to a block of π that is entirely contained in the set $\{2, ..., m-1\}$, then the number of elements in this set must be even. Since there are $m-2$ elements in $\{2, ..., m-1\}$, we see that $m-2$, and therefore m, must be even. Thus, $m = 2l$ for some integer $1 \leq l \leq k$. We let D_{2l-2} denote the number of non-crossing pair partions of $\{2, ..., 2l-1\}$ and let D_{2k-2l} denote the number of non-crossing pair partitions of $\{2l+1, ..., 2k\}$. Since each non-crossing pair partition of these two sets occurs independently, and since l can be any integer in $\{1, ..., k\}$, we have that

$$D_{2k} = \sum_{l=1}^{k} D_{2l-2} D_{2k-2l} = \sum_{l=1}^{k} D_{2(l-1)} D_{2(k-l)} = \sum_{l=0}^{k-1} D_{2l} D_{2(k-1-l)}. \tag{4.3}$$

Replacing k with $k+1$, we obtain

$$D_{2(k+1)} = \sum_{l=0}^{k} D_{2l} D_{2(k-l)}. \tag{4.4}$$

Comparing this result with the recursive formula for the Catalan number C_k, it follows that $D_{2k} = C_k$. □

CHAPTER 5

The Central Limit Theorems

We provide proofs of the classic and free central limit theorems. We show that the moments of $\frac{a_1+\cdots+a_N}{\sqrt{N}}$, where $a_1, ..., a_N$ are indentically distributed classically independent random variables, converge to the moments of a normally distributed random variable as $N \to \infty$. We also show that the moments of $\frac{a_1+\cdots+a_N}{\sqrt{N}}$, where $a_1, ..., a_N$ are identically distributed freely independent random variables, converge to the moments of a semicircular random variable as $N \to \infty$. Since the first part of the proofs of the classic and free central limit theorems is identical, we present first the common argument followed by the statement of the theorems.

Definition 5.1. A collection of random variables $\{a_i\}_{i\in\mathbb{N}}$ is called **identically distributed** if each random variable a_i for $i \in \mathbb{N}$ has the same probability distribution $\left\{\phi(a_i^n)\right\}_{n=0}^{\infty}$.

Definition 5.2. Let $(\mathcal{A}_N, \phi_N)_{N\in\mathbb{N}}$ and $(\mathcal{A}, \phi)$ be non-commutative probability spaces and consider random variables $a_N \in \mathcal{A}_N$ for each $N \in \mathbb{N}$, and $a \in \mathcal{A}$. We say that a_N **converges in distribution** towards a for $N \to \infty$, and denote this by

$$a_N \overset{distr}{\to} a$$

if we have

$$\lim_{N\to\infty} \phi_N(a_N^n) = \phi(a^n) \quad \text{for all } n \in \mathbb{N}.$$

Since the convergence in distribution of $\frac{a_1+\cdots+a_N}{\sqrt{N}}$ means the convergence of all moments of $\frac{a_1+\cdots+a_N}{\sqrt{N}}$, we need to calculate the limit $N \to \infty$ of all moments of $\frac{a_1+\cdots+a_N}{\sqrt{N}}$. We begin by calculating such moments for finite N.

Example 5.3. We let $N = 3$ and $n = 2$. Then, we have that

$$(a_1 + a_2 + a_3)^2 = a_1a_1 + a_1a_2 + a_1a_3 + a_2a_1 + a_2a_2 + a_2a_3 + a_3a_1 + a_3a_2 + a_3a_3. \tag{5.1}$$

The right side of the above equation is the sum of all distinct products of two random variables with indices 1, 2, or 3. Allowing the indices $r(1)$ and $r(2)$ to assume any value in the set $\{1, 2, 3\}$, this equation can also be written as

$$(a_1 + a_2 + a_3)^2 = \sum_{1\le r(1),r(2)\le 3} a_{r(1)}a_{r(2)}. \tag{5.2}$$

Extending the above example, we fix a positive integer n. We have that

$$(a_1 + \cdots + a_N)^n = \sum_{1 \le r(1),\ldots,r(n) \le N} a_{r(1)} \cdots a_{r(n)}. \tag{5.3}$$

Since ϕ is linear, it respects addition and we get that

$$\phi\big((a_1 + \cdots + a_N)^n\big) = \sum_{1 \le r(1),\ldots,r(n) \le N} \phi\big(a_{r(1)} \cdots a_{r(n)}\big). \tag{5.4}$$

We note that all the a_r have the same distribution, and therefore the same moments. Both classic and free independence give a rule for calculating mixed moments from the values of the moments of the individual variables.

Example 5.4. As an illustration, we consider the following example. Let a_1 and a_2 be freely independent random variables. By equation (2.7) we have that

$$\phi(a_1 a_2 a_1 a_2) = \phi(a_2)^2 \phi(a_1^2) + \phi(a_1)^2 \phi(a_2^2) - \phi(a_1)^2 \phi(a_2)^2. \tag{5.5}$$

By equation (2.7), we also have that

$$\phi(a_2 a_1 a_2 a_1) = \phi(a_1)^2 \phi(a_2^2) + \phi(a_2)^2 \phi(a_1^2) - \phi(a_2)^2 \phi(a_1)^2. \tag{5.6}$$

Since a_1 and a_2 have the same probability distribution, $\phi(a_1^n) = \phi(a_2^n)$ for every n. Thus, we see that $\phi(a_1 a_2 a_1 a_2) = \phi(a_2 a_1 a_2 a_1)$.

If all the random variables a_r are independent and identically distributed, then the value of $\phi(a_{r(1)} \cdots a_{r(n)})$ depends only on which of the indices are the same and which are different. In other words, we have

$$\phi\big(a_{r(1)} \cdots a_{r(n)}\big) = \phi\big(a_{p(1)} \cdots a_{p(n)}\big) \tag{5.7}$$

whenever, $r(i) = r(j)$ if and only if $p(i) = p(j)$ for all $1 \le i, j \le n$.

We can denote the common value of the moments in Example 5.4 as

$$\kappa_{\{(1,3),(2,4)\}} = \phi(a_1 a_2 a_1 a_2) = \phi(a_2 a_1 a_2 a_1). \tag{5.8}$$

We see that $\kappa_{\{(1,3),(2,4)\}}$ denotes the common value of all products of 4 random variables such that the indices of the first and third random variables are the same and the indices of the second and fourth random variables are the same. We note that $\pi = \{\{1,3\},\{2,4\}\}$ is a partition of the set $\{1,2,3,4\}$. Using this notation, we can write

$$\phi\big((a_1 + \cdots + a_N)^n\big) = \sum_{\pi \text{ partition of } \{1,\ldots,n\}} \kappa_\pi \cdot A_\pi^N \tag{5.9}$$

where A_π^N is the number of products of random variables corresponding to the same partition π of $\{1,...,n\}$. Referring to the example above in the case that $n = 4$, $A^N_{\{(1,3),(2,4)\}}$ is the number of all possible products of 4 random variables such that the indices of the first and third random variables are the same, and the indices of the second and fourth random variables are the same. We note that the only term in equation (5.9) that depends on the value of N is A_π^N.

We examine the contributions of different partitions to the sum in equation (5.9). First, we show that partitions with singletons do not contribute to the sum. Consider a partition $\pi = \{V_1,...,V_m\}$ where $V_k = \{j\}$ for some $1 \leq k \leq m$ and some $1 \leq j \leq n$. Then we have that

$$\kappa_\pi = \phi(a_{r(1)} \cdots a_{r(j)} \cdots a_{r(n)}) = \phi(a_{r(j)}) \cdot \phi(a_{r(1)} \cdots a_{r(j-1)} a_{r(j+1)} \cdots a_{r(n)}). \tag{5.10}$$

In the case of classically independent random variables, the above result follows from the factorization $\phi(ab) = \phi(a) \cdot \phi(b)$ for independent random variables a and b. Since $a_{r(j)}$ in equation (5.10) is classically independent from $a_{r(1)} \cdots a_{r(j-1)} a_{r(j+1)} \cdots a_{r(n)}$, and since classically independent random variables commute, we have that

$$\begin{aligned} \phi(a_{r(1)} \cdots a_{r(j)} \cdots a_{r(n)}) &= \phi(a_{r(j)} a_{r(1)} \cdots a_{r(j-1)} a_{r(j+1)} \cdots a_{r(n)}) \\ &= \phi(a_{r(j)}) \cdot \phi(a_{r(1)} \cdots a_{r(j-1)} a_{r(j+1)} \cdots a_{r(n)}). \end{aligned}$$

In the case of free independent random variables, the subalgebra generated by $a_{r(j)}$ is freely independent from the subalgebra generated by $\{a_{r(1)},...,a_{r(j-1)},a_{r(j+1)},....,a_{r(n)}\}$. Consequently, the factorization follows from the comment after the computation of the product of three freely independent random variables in equation (2.6). Since the random variables are centered, i.e. $\phi(a_{r(j)}) = 0$ for $1 \leq j \leq n$, we get $\kappa_\pi = 0$. Therefore, only partitions which have no singletons contribute to the sum in equation (5.9). In particular, we can restrict the partitions in our sum to partitions $\pi = \{V_1,...,V_m\}$ for which $m \leq n/2$.

Next, we consider a partition $\pi = \{V_1,...,V_m\}$. We have N possible choices for the common index $r(i)$ of the random variables $a_{r(i)}$ corresponding to the block V_1. We have $N - 1$ possible choices for the common index of the random variables corresponding to the second block V_2, and so on. If we denote by $|\pi|$ the number of blocks of the partition

π, then we have that

$$A_\pi^N = N(N-1)\cdots(N-|\pi|+1). \tag{5.11}$$

We note that, for large N, A_π^N grows asymptotically like $N^{|\pi|}$. By equation (5.9), we have that

$$\lim_{N\to\infty} \phi\Big(\Big(\frac{a_1+\cdots+a_N}{\sqrt{N}}\Big)^n\Big) = \lim_{N\to\infty}\sum_\pi \frac{A_\pi^N}{N^{n/2}}\kappa_\pi = \lim_{N\to\infty}\sum_\pi N^{|\pi|-(n/2)}\kappa_\pi. \tag{5.12}$$

We saw before that the only partitions π that contribute non-trivially to our sum are the ones which have the property that $|\pi| \le n/2$. For these partitions, either $|\pi|-(n/2)<0$ or $|\pi|-(n/2)=0$. In the first case, the factor $N^{|\pi|-(n/2)}$ has limit 0, and in the second case, $N^{|\pi|-(n/2)}$ has limit 1 as $N\to\infty$. Thus, all partitions with $|\pi|<n/2$ contribute a zero term to the sum in equation (5.9) in the limit. Only the partitions without singletons and with a number of blocks equal to $n/2$ give a non-zero contribution κ_π. If $|\pi|=n/2$ for a particular partition π that has no singletons, then each block of π must contain exactly two elements. Thus, π must be a pair partition of $\{1,...,n\}$. Therefore, we have the following result:

$$\lim_{N\to\infty} \phi\Big(\Big(\frac{a_1+\cdots+a_N}{\sqrt{N}}\Big)^n\Big) = \sum_{\substack{\pi \text{ pair partition of} \\ \{1,...,n\}}} \kappa_\pi. \tag{5.13}$$

Since there are no possible pair partitions of a set with an odd number of elements, then we have that

$$\lim_{N\to\infty} \phi\Big(\Big(\frac{a_1+\cdots+a_N}{\sqrt{N}}\Big)^n\Big) = 0 \qquad \text{for } n \text{ odd}. \tag{5.14}$$

After this point, the proofs of the classic and free central limit theorems are no longer identical. We present the rest of the proof in two cases.

1. Classic Central Limit Theorem

Theorem 5.5. *Let $(\mathcal{A},\phi)$ be a $*$-probability space and let $a_1, a_2, ... \in \mathcal{A}$ be a sequence of independent and identically distributed self-adjoint random variables. Furthermore, assume that all variables are centered, i.e. $\phi(a_r) := 0$ for all $r\in\mathbb{N}$, and denote by $\sigma^2 := \phi(a_r^2)$ for all $r\in\mathbb{N}$ the common variance of the variables. Then we have*

$$\tfrac{a_1+\cdots+a_N}{\sqrt{N}} \overset{distr}{\to} x,$$

where x is a normally distributed random variable with mean 0 and variance σ^2.

PROOF. In view of equation (5.13), we proceed by determining the contribution κ_π of each pair partition π of the set $\{1,...,n\}$. Since π is a pair partition, then n must be even. Consider a mixed moment $\kappa_\pi = \phi(a_{r(1)}\cdots a_{r(n)})$ corresponding to a pair partition π. Each index $r(i)$ for $1 \le i \le n$ occurs exactly twice in the mixed moment. Since the random variables all have the same moments, the rule for computing moments of classically independent random variables gives that $\phi(a_{r(1)}\cdots a_{r(n)})$ factorizes into a product of $n/2$ moments of the form $\phi(a^2_{r(i)})$. Since σ^2 denotes the common variance of the variables (i.e. $\phi(a^2_{r(i)}) = \sigma^2$ for all $1 \le r(i) \le N$), then $\phi(a_{r(1)}\cdots a_{r(n)}) = \sigma^n$. Since each mixed moment κ_π corresponding to a pair partition of the set $\{1,...,n\}$ contributes a value of σ^n, then equation (5.13) can be written as

$$\lim_{N\to\infty} \phi\Big(\Big(\frac{a_1+\cdots+a_N}{\sqrt{N}}\Big)^n\Big) = \sigma^n \cdot (\#\text{ of pair partitions of } \{1,...,n\}). \tag{5.15}$$

By Lemma 4.5 and equation (5.14), we have that

$$\lim_{N\to\infty} \phi\Big(\Big(\frac{a_1+\cdots+a_N}{\sqrt{N}}\Big)^n\Big) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ (n-1)!!\sigma^n & \text{if } n \text{ is even.} \end{cases} \tag{5.16}$$

According to Lemma 3.4, we see that $\lim_{N\to\infty} \phi\Big(\Big(\frac{a_1+\cdots+a_N}{\sqrt{N}}\Big)^n\Big)$ is equal to the n-th moment of a normally distributed random variable of mean 0 and variance σ^2 for each n. The result follows by the definition of convergence in distribution. □

2. Free Central Limit Theorem

Theorem 5.6. *Let $(\mathcal{A},\phi)$ be a $*$-probability space and let $a_1, a_2, ... \in \mathcal{A}$ be a sequence of freely independent and identically distributed self-adjoint random variables. Assume that $\phi(a_r) := 0$ for all $r \in \mathbb{N}$, and denote by $\sigma^2 := \phi(a_r^2)$ for all $r \in \mathbb{N}$ the common variance of the variables. Then we have*

$$\tfrac{a_1+\cdots+a_N}{\sqrt{N}} \overset{distr}{\to} s,$$

where s is a semicircular element with mean 0 and variance σ^2.

PROOF. As in the proof of the classic central limit theorem, we proceed by determining the contributions of each pair partition π of the set $\{1,...,n\}$. Consider a mixed moment $\kappa_\pi = \phi(a_{r(1)}\cdots a_{r(n)})$ corresponding to a pair partition π. We have two cases:

(1) All of the consecutive indices are different, i.e. $r(j) \neq r(j+1)$ for all $j = 1, ..., n-1$.
(2) At least one pair of adjacent indices are the same, i.e. $r(j) = r(j+1)$ for some $1 \leq j \leq n$.

Using free independence (Definition 2.18) in the first case, we have $\kappa_\pi = \phi(a_{r(1)} \cdots a_{r(n)}) = 0$ because no adjacent indices are the same and $\phi(a_{r(j)}) = 0$ for all $1 \leq j \leq n$. In the second case, we have that

$$\kappa_\pi = \phi(a_{r(1)} \cdots a_{r(j)} a_{r(j+1)} \cdots a_{r(n)}) = \phi(a_{r(1)} \cdots a_{r(j)} a_{r(j)} \cdots a_{r(n)}) \text{ for some } 1 \leq j \leq n-1.$$

Because π is a pair partition, the index $r(j)$ does not appear again among the indices of the mixed moment κ_π. The subalgebra generated by $a^2_{r(j)}$ is freely independent from the subalgebra generated by $\{a_{r(1)}, ..., a_{r(j-1)}, a_{r(j+2)}, ..., a_{r(n)}\}$. The factorization of the product of three freely independent random variables, as in the comment following equation (2.6), gives that

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

If we obtain a non-zero term and repeat the previous argument, then either $\kappa_\pi = 0$ as in the first case, or we can reduce the length of the mixed moment further to obtain another non-zero term as in the second case. We repeat the argument until we either get $\kappa_\pi = 0$ after a certain number of iterations, or until we get $\kappa_\pi = \phi(1) \cdot (\sigma^2)^{n/2} = \sigma^n$. We see that the only mixed moments which contribute non-trivially to our sum in equation (5.13) are those in which we can successively find a pair of adjacent random variables with the same index.

We consider the pair partitions corresponding to mixed moments κ_π that contribute a zero term to our sum. The fact that $\kappa_\pi = 0$ means that we are in the first case for some iteration of the above argument. Let π denote the pairing of elements in the mixed moment to which case (1) applies. Take the pair of elements $a_{r(i)} = a_{r(j)}$ corresponding to the block $V_i = \{i, j\}$ of π, and assume that $i < j$. Because $a_{r(i)}$ is not adjacent to $a_{r(j)}$, by condition (1), there must be some elements between $a_{r(i)}$ and $a_{r(j)}$. If there is another pair $a_{r(k)}$ and $a_{r(l)}$, with $k < l$, corresponding to the block $V_j = \{k, l\}$ of π between $a_{r(i)}$ and $a_{r(j)}$, i.e. $i < k < l < j$, then we rename the pair $(a_{r(k)}, a_{r(l)})$ as $(a_{r(i)}, a_{r(j)})$. We continue in this way until there is no pair of elements between $a_{r(i)}$ and $a_{r(j)}$. Because these elements cannot be adjacent, there must exist elements $a_{r(m)}$ and $a_{r(p)}$ in the mixed product $a_{r(1)} a_{r(2)} \cdots a_{r(n)}$ that correspond to the block $V_k = \{m, p\}$ of π such that $i < m < j$ and either $p < i$ or $p > j$. Referring to Definition 4.6, we see that π must be crossing because there exist elements $a_{r(i)}, a_{r(j)}, a_{r(m)}$, and $a_{r(p)}$ and blocks $V_i = \{i, j\}$ and $V_k = \{m, p\}$ of π such that either $i < m < j < p$ or $p < i < m < j$.

We conclude that the only pair partitions π that contribute non-trivially to our sum are the non-crossing pair partitions. Each non-crossing pair partition gives a contribution of σ^n. Therefore, equation (5.13) can be written as

$$\lim_{N\to\infty} \phi\Bigg(\Big(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\Big)^n\Bigg) = \sigma^n \cdot (\# \text{ of non-crossing pair partitions of } \{1, ..., n\}).$$

Because n must be even (see the comment before the beginning of section 1), we set $n = 2k$ for some $k \in \mathbb{N}$. According to Lemma 4.9, the number of non-crossing pair partitions of $\{1, ..., 2k\}$ is equal to the Catalan number C_k. Using this fact and equation (5.14), we have that

$$\lim_{N\to\infty} \phi\Bigg(\Big(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\Big)^n\Bigg) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \sigma^{2k} \cdot C_k & \text{if } n = 2k. \end{cases} \tag{5.17}$$

Taking $\sigma^2 = 1$ and referring to Lemma 3.7, we have that the moments of $\dfrac{a_1 + \cdots + a_N}{\sqrt{N}}$ converge to the moments of a standard semicircular random variable as $N \to \infty$. The result follows by definition of convergence in distribution. $\square$

CHAPTER 6

Conclusion

In this work, we have given a basic introduction to free probability theory. We have examined the similarities and differences between classic and free independence. We have computed the moments of normally distributed random variables and semicircular random variables. Finally, we have given proofs of the classic and free central limit theorems using a combinatorial approach.

This paper provides an understanding of concepts that are essential to the further study of free probability theory. In particular, one can explore the formula for finding mixed moments of freely independent random variables from the free cumulants rather than from direct computations. The expression $\kappa_{\{(1,3),(2,4)\}}$ in equation (5.8) is an example of a cumulant. Cumulants are determined by the block structure of the partition of the set $\{1, ..., n\}$ corresponding to the mixed moment $\phi(a_{r(1)} \cdots a_{r(n)})$. One can also advance to the study of random matrices and prove that a large random self-adjoint $N \times N$ matrix, with additonal conditions as specified in the second paragraph of the introduction, converges to a semicircular element when $N \to \infty$.

Moreover, there are various conjectures and open questions related to free probability. A particularly interesting open problem is a formal proof of the replica method. The replica method is used in computations involving functions of a variable that can be expressed as a power series of the variable. In other words, this technique reduces a function $f(z)$ to powers of z, and the same computation which is to be done on $f(z)$ is done on the powers of z. This method has accurately reproduced results from random matrix theory and free probability theory. However, the method fails in some cases. Free probability theory could potentially be used to establish necessary and sufficient conditions for the validity of the replica method. Another open problem is determining under which extra conditions symmetric random $N \times N$ matrices with real entries are asymptotically free independent. Random $N \times N$ matrices with real entries are often used in statistics. As these examples suggest, free probability theory provides much potential for new applications in the areas of engineering, physics, statistics, and mathematics.

Bibliography

[1] R. Couillet, M. Debbah, *Random Matrix Methods for Wireless Communications*, Cambridge University Press, 2011.

[2] K. Dykema, A. Nica, D. Voiculescu, *Free Random Variables*, American Mathematical Society, 1992.

[3] F. Hiai, D. Petz, *The Semicircle Law, Free Random Variables, and Entropy*, American Mathematical Society, 2000.

[4] A. Nica, R. Speicher, *Lectures on the Combinatorics of Free Probability*, Cambridge University Press, 2006.

[5] R. Speicher, *Free Probability Theory and Random Matrices*, Lect. Notes Math., **1815**(2003), pp. 53-73.

[6] D. Voiculescu, *Free Probability and the von Neumann Algebras of Free groups*, Rep. Math. Phys., **55**(2005), pp. 127-133.

[7] D. Voiculescu, *Limit Laws for Random Matrices and Free Products*, Invent. Math., **104**(1991), pp. 201-220.