University of Applied Sciences Munich
Department of Computer Science and Mathematics

Bachelor Thesis
Scientific Computing

N.V. Krylov's Proof of the de Moivre-Laplace Theorem

N.V. Krylovs Beweis des Satzes von de Moivre-Laplace

Mario Teixeira Parente
(2169 5010)
21st August 2013

Supervised by:
Prof. Dr. Manfred Gruber
Department of Computer Science and Mathematics
University of Applied Sciences Munich

Abstract

The Central Limit Theorem is an outstanding discovery of mathematics. It is not only a theoretical construct from probability theory, but simplifies also many calculations in everyday work.

This thesis treats a proof of the de Moivre-Laplace theorem which is a special case of the (classical) Central Limit Theorem for sums of Bernoulli distributed random variables. The proof was created by Nicolai V. Krylov (University of Minnesota) and shall, according to the title of the corresponding paper, serve as a lecture for undergraduate students.

In this proof Krylov makes no use of constructs from measure theory, but uses only contents from a typical lecture about calculus. Krylov knows how to use these simple concepts, such that the proof's complexity does not grow too much. It is this aspect that makes the proof so impressive and exciting.

Zusammenfassung

Der Zentrale Grenzwertsatz ist eine herausragende Entdeckung der Mathematik. Er stellt nicht nur ein theoretisches Konstrukt der Wahrscheinlichkeitstheorie dar, sondern vereinfacht auch viele Rechnungen in der Praxis.

Diese Arbeit behandelt einen Beweis des Satzes von de Moivre-Laplace, welcher ein Spezialfall des (klassischen) Zentralen Grenzwertsatzes für die Summe Bernoulli-verteilter Zufallsgrößen darstellt. Der Beweis wurde von Nicolai V. Krylov (Universität Minnesota) geführt und soll laut dem Titel der zugehörigen Schrift einer Vorlesung für Bachelor-Studenten dienen.

Krylov benutzt in diesem Beweis keinerlei Konstrukte aus der Maßtheorie, sondern verwendet lediglich Inhalte aus einer typischen Analysis-Vorlesung. Krylov versteht es, diese einfachen Konzepte so zu verwenden, dass die Komplexität des Beweises im Rahmen bleibt. Genau dieser Aspekt macht den Beweis so eindrucksvoll und spannend.

"The scientist finds his reward in what Henri Poincaré calls the joy of comprehension, and not in the possibilities of application to which any discovery of his may lead."

"Der Wissenschaftler findet seine Belohnung in dem, was Poincaré die Freude am Verstehen nennt, nicht in den Anwendungsmöglichkeiten seiner Erfindung."

- Albert Einstein

Preface

The topic of this thesis was offered by my supervisor *Prof. Dr. Manfred Gruber (University of Applied Sciences Munich, Department for Computer Science and Mathematics)*. He discovered a paper of *Nicolai V. Krylov (University of Minnesota)* about the Central Limit Theorem for undergraduate students ([Kry]). It was my wish to treat a topic in my thesis which is purely mathematical and this paper gave me the opportunity to do that. My task was to go through the proof, to understand it and to make it more accessible even for not purely mathematical interested students. But I want to note that the thesis is an interpretation in no small part. It was quite difficult to capture all thoughts of Prof. Krylov and probably I did not capture them all. Even if it was not that easy to follow his calculations and ideas, it was a great experience to do something like that and above all to remain stubbornly. Of course, there is a possibility for some mistakes and inexactness for which I assume all responsibility.

This thesis would not have been possible without my supervisor Prof. Dr. Manfred Gruber who spent a lot of time with me explaining some mathematical details for my thesis.

I also would like to thank the *German National Academic Foundation* which supported me with a scholarship during the last three terms. It put the financial aspects in the background and let me completely focus on my work.

Last but not least I want to thank all the persons from the University of Applied Sciences Munich who contributed to the success of my studies.

This thesis is dedicated to all people who believe in me. First of all, I would like to thank my family for their support during my whole life. But I would not be the same character as today without my friends. All the times we spent talking, joking, laughing and crying made me the person I am today.

Thank you!

Mario Teixeira Parente — Munich, 21st August 2013

Contents

List of Figures

Chapter 1

Introduction

In this introductory chapter, we will get to know the basics for understanding the proof of the de Moivre-Laplace Theorem as a special case of the Central Limit Theorem for sums of Bernoulli distributed random variables. Without discussing some important facts, it would be quite difficult to understand the statement of the theorem and the proof even more.

1.1 Binomial Distribution

Before we start with the binomial distribution, let us first consider a simpler distribution called Bernoulli distribution $Ber(p)$ with parameter $p \in (0,1)$. This distribution gives the value 1 for success with probability p and value 0 for failure with probability $q = 1 - p$. In the following sections, we will use X as a Bernoulli distributed random variable $\to X \sim Ber(p)$. The expected value and the variance are easily computed to

$$E(X) = 1 \cdot p + 0 \cdot q = p,$$

$$Var(X) = E(X^2) - E(X)^2 = \left(1^2 \cdot p + 0^2 \cdot q\right) - p^2 = p(1-p).$$

In this thesis, we will only consider cases with $p = q = \frac{1}{2}$ without loss of generality. Thus, it is

$$E(X) = \frac{1}{2}$$

and

$$Var(X) = \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}.$$

Now, let us look at the binomial distribution $B(n, p)$ with parameters $n \in \mathbb{N}$ and $p = \frac{1}{2}$. (We will use S_n as a special name for binomially distributed random variables $\rightarrow S_n \sim B\left(n, \frac{1}{2}\right)$).

As in the case of the Bernoulli distribution, it is a discrete probability distribution. The binomial distribution looks at the number of successes (or failures) of n independent Bernoulli trials with the same success probability p which we set to $\frac{1}{2}$ again.

It has thus the following probability mass function (for k successes with $k = 0, 1, \ldots, n$) which is illustrated in Figure 1.1:

$$\begin{aligned} P(S_n = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \binom{n}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{n-k} \\ &= \binom{n}{k} \frac{1}{2^n} \\ &= \frac{1}{2^n} \cdot \frac{n!}{k!(n-k)!} \end{aligned} \tag{1.1}$$

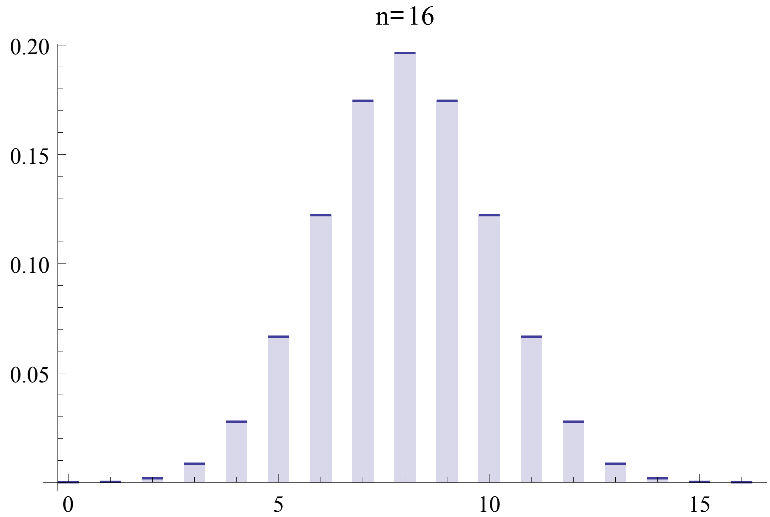

Figure 1.1: Probability mass function of binomial distribution with $p = \frac{1}{2}$

For the calculation of the expected value and the variance, it is useful to regard the binomial distribution S_n as a sum of independent Bernoulli trials X_i. The following calculation, adjusted from [Irl05, p.59], shows exactly this fact:

$$\begin{aligned} P(S_n = k) &= P\left(\sum_{i=1}^{n} X_i = k\right) \\ &= \sum_{(X_1,\ldots,X_n):\sum_{i=1}^{n} X_i = k} p^k(1-p)^{n-k} \\ &= \binom{n}{k} p^k(1-p)^{n-k} \\ &= B(n,p)(k) \end{aligned} \tag{1.2}$$

It is easy to see that the Bernoulli distribution is a special case of the binomial distribution with $n = 1$.

Using the fact from Equation (1.2), we can now easily compute the expected value and the variance of S_n with the ones from the Bernoulli distribution:

$$E\left(S_n\right) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E\left(X_i\right) = np = \frac{n}{2}$$
$$Var\left(S_n\right) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var\left(X_i\right) = np(1-p) = \frac{n}{4}$$

Changing the order of summation and the functions for the expected value and the variance is allowed due to the independency of the random variables X_i.

1.2 Normal Distribution

The other distribution we will need, is the normal distribution $N(\mu, \sigma^2)$. It is a continuous probability distribution and one of the most important distributions in applied statistics. The parameter μ denotes the expected value and σ^2 denotes the variance. Its density function has the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

The distribution $N(0,1)$ is called standard normal distribution and has thus the density function

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

In Figure 1.2, you can see the density function of different normal distributions including the standard normal distribution.

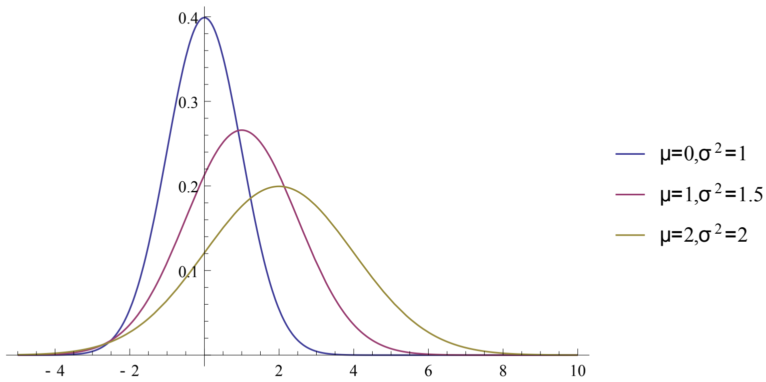

Figure 1.2: Probability density function of some normal distributions

It is nice to know that every normal distribution is a special version of the standard normal distribution. So,

$$f(x) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)$$

is a density function of some normal distribution $N(\mu, \sigma^2)$. The domain of the standard normal distribution was only stretched by the factor σ and translated by μ.

The cumulative distribution function (CDF) of the general normal distribution is

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt.$$

If we substitute with $z = \frac{t-\mu}{\sigma}$, the general CDF $F(x)$ can be represented in terms of the CDF of the standard normal distribution. That is,

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{1}{2}z^2} dz = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

with

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt.$$

Figure 1.3 shows the cumulative distribution functions of the corresponding density functions from Figure 1.2.

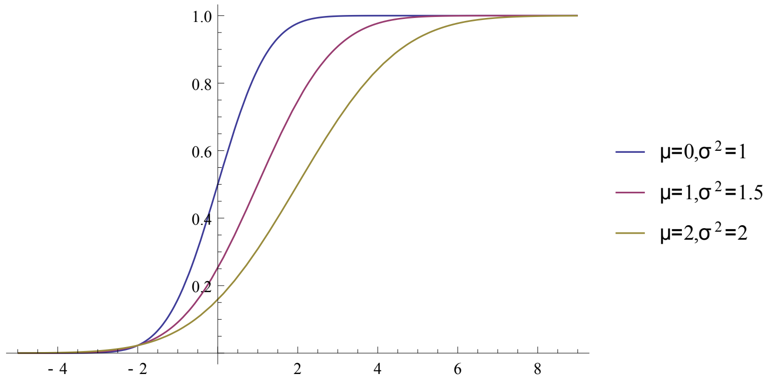

Figure 1.3: Cumulative distribution function of normal distributions from Figure 1.2

1.3 De Moivre-Laplace Theorem

The de Moivre-Laplace Theorem states that sums of Bernoulli distributed random variables, which can be treated as binomially distributed as we saw before, converge to the normal distribution for $n \to \infty$ and probabilities $0 < p < 1$. Figure 1.4 illustrates this fact. It shows two binomial distributions for $n = 4$ and $n = 16$ with $p = \frac{1}{2}$ and the „according" normal distributions with the same expected value and the same variance.

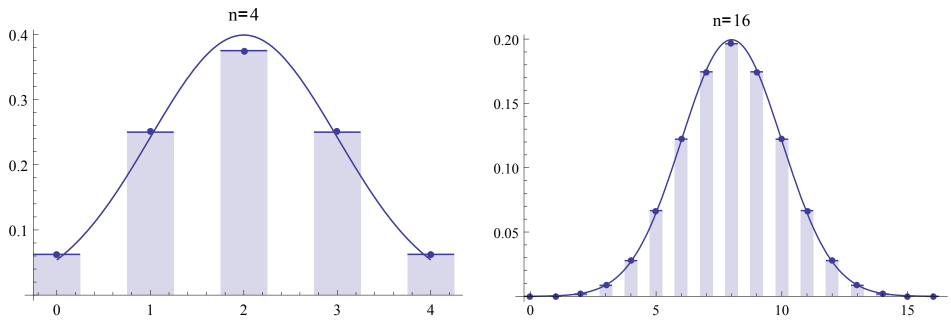

Figure 1.4: Two binomial distributions ($p = \frac{1}{2}$) with their „according" normal distribution

This theorem is a special case of the classical Central Limit Theorem which says that under certain conditions the sum of random variables of *every* distribution converges to the normal distribution. This is the reason why the normal distribution is so important.

The mathematical description of the de Moivre-Laplace Theorem is the following.

Theorem (de Moivre-Laplace Theorem)**.** *Let S_n be a binomially distributed random variable with parameters n and $0 < p < 1$ and let Z be a normal distributed random variable with parameters $\mu = np$ and $\sigma^2 = np(1-p)$. Then, it holds*

$$\lim_{n\to\infty} P(S_n < t) = F(t)$$

where $F(t)$ is the cumulative distribution function of Z.

An alternative way of defining this theorem is to consider normalized random variables which leads to

Theorem (de Moivre-Laplace Theorem with normalized random variables)**.** *Let S_n be a binomially distributed random variable with parameters n and $0 < p < 1$. Then, it holds*

$$\lim_{n\to\infty} P\left(\frac{S_n - np}{\sqrt{np(1-p)}} < z\right) = \Phi(z)$$

where $\Phi(z)$ is the cumulative distribution function of the standard normal distribution.

Both versions of the theorem have the same statement, namely that the difference (or error) between the cumulative values of the binomial distribution and the ones of the normal distribution gets smaller for increasing n and finally converges to zero. The question *how* it converges to zero is the topic of the next section.

1.4 Error Analysis

In this section, we try to find out how the error between the normal distribution and the binomial distribution behaves with increasing n.

We calculate the error in the following way. For fixed n and p, let us compute the expected value μ and the variance σ^2 of a binomially distributed random variable $S_n \sim B(n,p)$ (see Section 1.1). Then we can set up the „according“ normal distributed random variable $Z \sim N(\mu, \sigma^2)$. After that we look at the values of S_n at the points $k = 0, 1, \ldots, n$ and subtract them from the values of Z at the same points to get the error. The euclidean norm of the resulting error vector is the value we look at for increasing n. The mathematical way describing the euclidean norm of the resulting error vector is

$$\sqrt{\sum_{k=0}^{n} [Z(k) - S_n(k)]^2}.$$

In Figure 1.5, we can see the error sums for $p = \frac{1}{2}$ and $n = 4, 6, 8, \ldots, 50$.

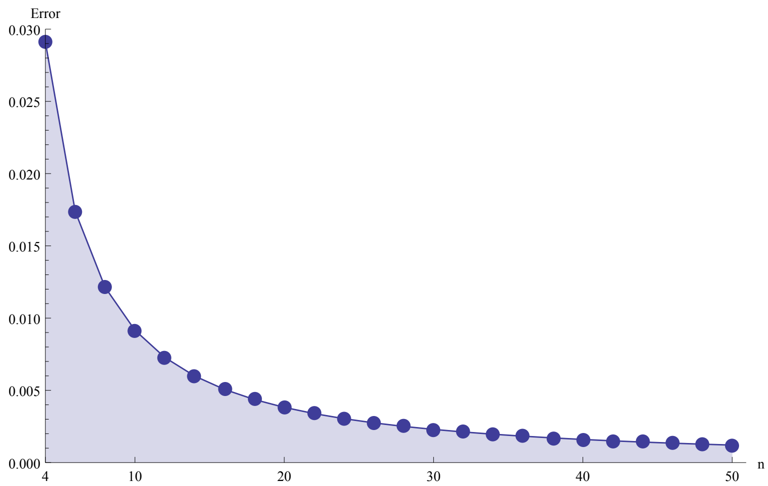

Figure 1.5: Error sums for $p = \frac{1}{2}$

The error decreases exponentially and converges to zero with increasing n. Hopefully we can verify this result in Section 2 which deals with the proof. But before we start, let us take a look at some outstanding historical facts.

1.5 Some History

The first approach to approximate the binomial with the normal distribution was made by Abraham de Moivre (1667-1754). In 1738, he published the second edition of his book *The Doctrine of Chances* which contained the theorem for the first time.

De Moivre looked at „the number of heads coming from many tosses of a fair coin“ [Tij04, p.169]. The probability of tossing a head is Bernoulli distributed which says that the number of heads, as the sum of n independent Bernoulli trials, is distributed binomially as we saw from Equation (1.2). De Moivre also gave a first suggestion for the „normal curve“. He tried to find an approximation for the binomial coefficient and proved that

$$\sum_{\left|x-\frac{n}{2}\right|\leq d} \binom{n}{x} \left(\frac{1}{2}\right)^n \approx \frac{4}{\sqrt{2\pi}} \int_0^{\frac{d}{\sqrt{n}}} e^{-2y^2} dy.$$

But to get a proper probability distribution, it was necessary to norm the integral $\int_{-\infty}^{\infty} e^{\frac{1}{2}^2} dx$. It was Pierre-Simon Laplace (1749-1827) who did the computation. His result was

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt = \sqrt{2\pi}.$$

Two different ways of computing this integral can be seen in Appendix A.1.

In 1809, Carl Friedrich Gauß (1777-1855) defined the well-known normal distribution in his book *Theoria motus corporum coelestium in sectionibus conicis solem ambientium* (engl.: Theory of motion of the celestial bodies moving in conic sections around the sun) as we know it today to

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

We can see Gauß in Figure 1.6, which shows him and the normal curve on the front of an old German bank note.

Figure 1.6: Bank note of 10 Deutsche Mark with Carl Friedrich Gauß and the normal curve on its front

Finally, it was again Laplace, who proved the Central Limit Theorem in 1810 and hence completed de Moivre's work on the theorem for binomial distributions.

Chapter 2

Proof by Krylov

For the proof, Krylov uses only some aspects from undergraduate calculus and thus avoids concepts from measure theory which would make the proof even more complex. There is also another method to prove the de Moivre-Laplace Theorem using Stirling's formula which gives an approximation for large factorials:

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n}$$

The proof using this formula can be seen in [Irl05, p.64-66].

2.1 Motivation

In this section, we give a short setup for the following section and try to motivate the exact proof. Thus, here we will show in an intuitive way that the de Moivre-Laplace Theorem holds for distributions with probability $p = \frac{1}{2}$.

First, let us introduce some variables for latter calculations. We will center the random variable S_n and make its variance constant. The new random variable is called Z_n and defined as

$$Z_n := \frac{S_n - \frac{n}{2}}{\sqrt{n}}.$$

With some simple steps, we can see that Z_n is indeed centered and has constant variance (which means that it does not depend on n):

$$\begin{aligned}
E(Z_n) &= E\left(\frac{S_n - \frac{n}{2}}{\sqrt{n}}\right)\\
&= \frac{1}{\sqrt{n}} E\left(S_n - \frac{n}{2}\right)\\
&= \frac{1}{\sqrt{n}} \left(E\left(S_n\right) - E\left(\frac{n}{2}\right)\right)\\
&= \frac{1}{\sqrt{n}} \left(\frac{n}{2} - \frac{n}{2}\right) = 0
\end{aligned}$$

$$\begin{aligned}
Var(Z_n) &= Var\left(\frac{S_n - \frac{n}{2}}{\sqrt{n}}\right)\\
&= \frac{1}{n} Var(S_n)\\
&= \frac{1}{n} \cdot \frac{n}{4} = \frac{1}{4}
\end{aligned}$$

We may hope that the distributions of Z_n converge to something as $n \to \infty$. This means that for $a < b$

$$P\left(a < \frac{S_n - \frac{n}{2}}{\sqrt{n}} \leq b\right) \tag{2.1}$$

has a limit as $n \to \infty$. In order to express the probability above, it is necessary to transform the variable k like it was done with the random variable S_n. It is

$$y_{kn} := \frac{k - \frac{n}{2}}{\sqrt{n}}, \qquad k = 0, 1, \ldots, n.$$

Informally saying, y_{kn} is the "centered variable k with constant variance". It has proven useful to know what $y_{k+1,n}$ means. So,

$$y_{k+1,n} = \frac{k + 1 - \frac{n}{2}}{\sqrt{n}} = \frac{k - \frac{n}{2}}{\sqrt{n}} + \frac{1}{\sqrt{n}} = y_{kn} + \frac{1}{\sqrt{n}}. \tag{2.2}$$

This says that the distance between y_{kn} and $y_{k+1,n}$ is $\frac{1}{\sqrt{n}}$ which converges to 0 as $n \to \infty$.

By the way, this is the reason why the variance was made constant. Without constant variance, the y_{kn}'s would spread way to far. This fact is illustrated in Figure 2.1. Red points display y_{kn}'s with non-constant variance (depending on n), blue points show y_{kn}'s having constant variance.

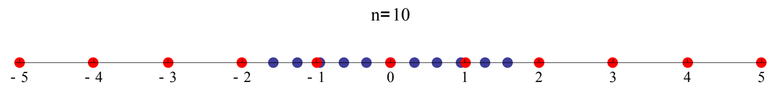

Figure 2.1: y_{kn}'s with constant and non-constant variance

At this point, we can notice an important property of the k-y_{kn}-transformation:

$$k = \frac{n}{2} \Leftrightarrow y_{kn} = 0 \tag{2.3}$$

This means that $\frac{n}{2}$ is the center for the k's and 0 is the center for the y_{kn}'s.

We also will need the reverse transformation from y_{kn} to k:

$$\begin{aligned} y_{kn} &= \frac{k - \frac{n}{2}}{\sqrt{n}} \\ \Leftrightarrow \sqrt{n}\, y_{kn} &= k - \frac{n}{2} \\ \Leftrightarrow \qquad k &= \sqrt{n}\, y_{kn} + \frac{n}{2} \end{aligned} \tag{2.4}$$

The probability $P(S_n = k)$ now arises to

$$P(S_n = k) = P\left(\frac{S_n - \frac{n}{2}}{\sqrt{n}} = y_{kn}\right).$$

This probability is the definition for f_n which is a function of y_{kn}. Hence, it is

$$f_n(y_{kn}) := P\left(\frac{S_n - \frac{n}{2}}{\sqrt{n}} = y_{kn}\right) = P(S_n = k). \tag{2.5}$$

Now, the key idea is to put $f_n(y_{kn})$ and $f_n(y_{k+1,n})$ in relation (or similar $P(S_n = k)$ and $P(S_n = k+1)$). As we can guess, our goal is indeed to derive a recurrence relation. For $k \leq n-1$, a simple algebraic manipulation of Equation (1.1) leads to

$$\begin{aligned} P(S_n = k) &= \frac{1}{2^n} \cdot \frac{n!}{k!(n-k)!} \\ &= \frac{1}{2^n} \cdot \frac{n!}{(k+1)!(n-(k+1))!} \cdot \frac{k+1}{n-k} \\ &= P(S_n = k+1)\, \frac{k+1}{n-k}. \end{aligned} \tag{2.6}$$

So, here we have the relation between the probability for k and $k+1$ successes. If we substitute Equations (1.1), (2.4) and (2.5) in the equation above, we get

$$\begin{aligned} && f_n(y_{kn}) &= f_n(y_{k+1,n})\, \frac{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1}{n - \sqrt{n}\, y_{kn} - \frac{n}{2}} \\ \Leftrightarrow && f_n(y_{kn}) &= f_n\left(y_{kn} + \frac{1}{\sqrt{n}}\right) \frac{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1}{\frac{n}{2} - \sqrt{n}\, y_{kn}} \\ \Leftrightarrow && f_n\left(y_{kn} + \frac{1}{\sqrt{n}}\right) &= f_n(y_{kn})\, \frac{\frac{n}{2} - \sqrt{n}\, y_{kn}}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1}. \end{aligned} \tag{2.7}$$

Subtracting $f_n(y_{kn})$ from both sides leads to

$$\begin{aligned} f_n\left(y_{kn} + \frac{1}{\sqrt{n}}\right) - f_n(y_{kn}) &= f_n(y_{kn}) \left(\frac{\frac{n}{2} - \sqrt{n}\, y_{kn}}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1} - 1 \right) \\ &= f_n(y_{kn})\, \frac{\frac{n}{2} - \sqrt{n}\, y_{kn} - (\sqrt{n}\, y_{kn} + \frac{n}{2} + 1)}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1} \\ &= f_n(y_{kn})\, \frac{-2\sqrt{n}\, y_{kn} - 1}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1} \\ &= -f_n(y_{kn})\, \frac{2\sqrt{n}\, y_{kn} + 1}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1}. \end{aligned} \tag{2.8}$$

Here, we can see the aspect of the recurrence relation. The natural conclusion is to make this relation to an ordinary differential equation (ODE). For this, let us believe that for a function ϕ holds

$$\phi_n(y_{kn}) := a_n^{-1} f_n(y_{kn}) \to \phi(y)$$

as n $\to \infty$ and $y_{kn} \to y$. The sequence a_n stands for the order of f_n's terms and has to be determined. To find this order, let us regard the probability in Equation (2.1) which is

$$\begin{aligned} P\left(a < \frac{S_n - \frac{n}{2}}{\sqrt{n}} \leq b\right) &= \sum_{k: a < \frac{k - \frac{n}{2}}{\sqrt{n}} \leq b} P(S_n = k) \\ &= \sum_{k: a < y_{kn} \leq b} f_n(y_{kn}). \end{aligned}$$

The sum above has about $(b-a)\sqrt{n}$ terms. To see this, let us concentrate on the summation variable k and its condition $a < y_{kn} \leq b$ and do some simple algebraic steps:

$$\begin{aligned} a <& \quad y_{kn} \quad \leq b \\ a <& \frac{k - \frac{n}{2}}{\sqrt{n}} \leq b \\ a\sqrt{n} <& \; k - \frac{n}{2} \leq b\sqrt{n} \\ a\sqrt{n} + \frac{n}{2} <& \quad k \quad \leq b\sqrt{n} + \frac{n}{2} \end{aligned}$$

Thus, the difference of k's boundaries is

$$\begin{aligned} b\sqrt{n} + \frac{n}{2} - \left(a\sqrt{n} + \frac{n}{2}\right) &= b\sqrt{n} - a\sqrt{n} \\ &= (b-a)\sqrt{n}. \end{aligned}$$

This means that every term of f_n should be of order $\frac{1}{\sqrt{n}}$, which furthermore says that $a_n^{-1} = \sqrt{n}$.

For ϕ_n, Equation (2.8) becomes

$$\phi_n\left(y_{kn} + \frac{1}{\sqrt{n}}\right) - \phi_n(y_{kn}) = -\phi_n(y_{kn})\, \frac{2\sqrt{n}\, y_{kn} + 1}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1}. \tag{2.9}$$

Now, let us pretend that Equation (2.9) holds for ϕ rather than for ϕ_n and for all

y in place of y_{kn}. That leads to

$$\phi\left(y+\frac{1}{\sqrt{n}}\right)-\phi(y)=-\phi(y)\,\frac{2y\sqrt{n}+1}{y\sqrt{n}+\frac{n}{2}+1}. \tag{2.10}$$

Here, the ODE aspect comes in. If we divide the equation by $\frac{1}{\sqrt{n}}$ and let $n\to\infty$, we get an ODE whose solution is $\phi(y)$ to which $\phi_n(y)$ should be close. In order to specify the ODE, we first have to determine the limit of the fraction:

$$\lim_{n\to\infty}\frac{2y\sqrt{n}+1}{y\sqrt{n}+\frac{n}{2}+1}\cdot\frac{1}{\frac{1}{\sqrt{n}}}=\lim_{n\to\infty}\frac{2y\,n+\sqrt{n}}{y\sqrt{n}+\frac{n}{2}+1}=4y. \tag{2.11}$$

Thus, the ODE looks like

$$\phi'(y)=-4y\,\phi(y). \tag{2.12}$$

To get the solution of this ODE, we can use the *method of integrating factors.* We use the integrating factor $\psi(y)$ and multiply it by the equation:

$$\psi(y)\cdot\phi'(y)+\psi(y)\cdot 4y\,\phi(y)=0 \tag{2.13}$$

If we assume that $\psi(y)\,4y=\psi'(y)$, the equation can be simplified with some manipulations:

$$\begin{aligned}
&& \psi(y)\cdot\phi'(y)+\psi'(y)\cdot\phi(y)&=0\\
\Leftrightarrow && (\psi(y)\cdot\phi(y))'&=0\\
\Leftrightarrow && \int(\psi(y)\cdot\phi(y))'\,dy&=\int 0\,dy\\
\Leftrightarrow && \psi(y)\cdot\phi(y)&=c\\
\Leftrightarrow && \phi(y)&=c\cdot\psi(y)^{-1}
\end{aligned}$$

As mentioned above, the integrating factor $\psi(y)$ satisfies the equation

$$\psi(y)\,4y=\psi'(y)\Leftrightarrow\frac{\psi'(y)}{\psi(y)}=4y. \tag{2.14}$$

If we notice that $\frac{\psi'(y)}{\psi(y)} = \ln(\psi(y))'$, the integrating factor can be easily computed:

$$\begin{aligned}
&& \frac{\psi'(y)}{\psi(y)} &= 4y \\
&\Leftrightarrow & \ln(\psi(y))' &= 4y \\
&\Leftrightarrow & \int \ln(\psi(y))'\,dy &= \int 4y\,dy \\
&\Leftrightarrow & \ln\psi(y) &= 2y^2 + c \\
&\Leftrightarrow & \psi(y) &= c\,e^{2y^2}
\end{aligned}$$

It follows that

$$\phi(y) = c\,e^{-2y^2}.$$

And here, we have derived the aspect of the normal distribution for this special instance of the Central Limit Theorem.

Now, you may wonder why the coefficient of the exponent is -2 and thus different from the one in the Gaussian function which is $-\frac{1}{2}$. That is, because we did not use the correct standard deviation for normalizing the random variable S_n. We used $\sqrt{n}$ instead of $\frac{\sqrt{n}}{2}$. You can find the previous calculation with the correctly normalized random variable in Appendix A.2 to see that then the correct coefficient arises.

2.2 Krylov's „First Rigorous Result“

First, it is useful to notice a fact coming from the result of the motivation chapter. The result stated that

$$\phi(y) = c\,e^{-2y^2}.$$

But what is the value of c? To get this value, it is necessary that e^{-2y^2} becomes one. This is true for $y = 0$. Thus, $c = \phi(0)$ which leads to the equation

$$\phi(y) = \phi(0)\,e^{-2y^2}. \tag{2.15}$$

2.2.1 Part 1

In the first part of the proof, our (and Krylov's) goal is to show that Equation (2.15) is also true for our discrete case as $n \to \infty$. Expressed in a mathematical way, this means we have to prove that

$$\phi_n(0)^{-1}\, e^{2y_{kn}^2}\, \phi_n(y_{kn}) = e^{\mathcal{O}(z(n))} = 1 + \mathcal{O}\left(z(n)\right) \tag{2.16}$$

where $z(n) \to 0$ as $n \to \infty$.

Let us look at Krylov's version of the statement. He wants to show that

$$c_n^{-1}\, e^{2y_{kn}^2}\, f_n(y_{kn}) = e^{\mathcal{O}(z(n))} = 1 + \mathcal{O}\left(z(n)\right)$$

holds for *even* n where $c_n = P(S_n = \frac{n}{2})$ and $z(n) = n^{4\beta-3}$ for $\beta \in \left[\frac{1}{2}, \frac{3}{4}\right]$.

Apart from his choice of $z(n)$, he makes the same statement as we do, what does not seem so at first glance. c_n can be written as $f_n(0)$ as we know from Equation (2.3) and $\phi_n(y_{kn}) = \sqrt{n}\, f_n(y_{kn})$ as we found out in the motivation chapter. Our choice of $z(n)$ will be developed during the proof.

Now, let us begin with the proof of the first part. Krylov decides to consider all k's such that $\left|k - \frac{n}{2}\right| \leq n^\beta$ for some β and so do we. In terms of y_{kn}, this is

$$\left|\frac{k - \frac{n}{2}}{\sqrt{n}}\right| \leq n^{\beta - \frac{1}{2}} \Leftrightarrow |y_{kn}| \leq n^{\beta - \frac{1}{2}}.$$

The values of n^β and $n^{\beta-\frac{1}{2}}$ have to go to ∞ as $n \to \infty$ because we want to capture the whole real axis. Thus, a first useful condition for β is $\beta > \frac{1}{2}$.

Let us do some manipulations of Equation (2.16) which will simplify the proof:

$$
\begin{aligned}
& & \phi_n(0)^{-1}\, e^{2y_{kn}^2}\, \phi_n(y_{kn}) &= e^{\mathcal{O}(z(n))} \\
&\Leftrightarrow & \ln \phi_n(0)^{-1} + 2y_{kn}^2 + \ln \phi_n(y_{kn}) &= \mathcal{O}\left(z(n)\right) \\
&\Leftrightarrow & \ln \frac{1}{\sqrt{n}\, f_n(0)} + 2y_{kn}^2 + \ln(\sqrt{n}\, f_n(y_{kn})) &= \mathcal{O}\left(z(n)\right) \\
&\Leftrightarrow & -\ln \sqrt{n} - \ln f_n(0) + 2y_{kn}^2 + \ln \sqrt{n} + \ln f_n(y_{kn}) &= \mathcal{O}\left(z(n)\right) \\
&\Leftrightarrow & \underbrace{\ln f_n(y_{kn})}_{g_n(y_{kn})} - \underbrace{\ln f_n(0)}_{g_n(0)} &= -2y_{kn}^2 + \mathcal{O}\left(z(n)\right) \qquad (2.17)
\end{aligned}
$$

This is the equation which we finally want to prove. The important part of this equation is the left hand side $g_n(y_{kn}) - g_n(0)$. Because we know that f_n has to converge to $e^{-2y_{kn}^2}$ as $n \to \infty$, we also know that g_n has to converge to $-2y_{kn}^2$ which is a parabola (see Figure 2.2).

Figure 2.2: f_n and g_n

Figure 2.3 illustrates how the left hand side is computed.

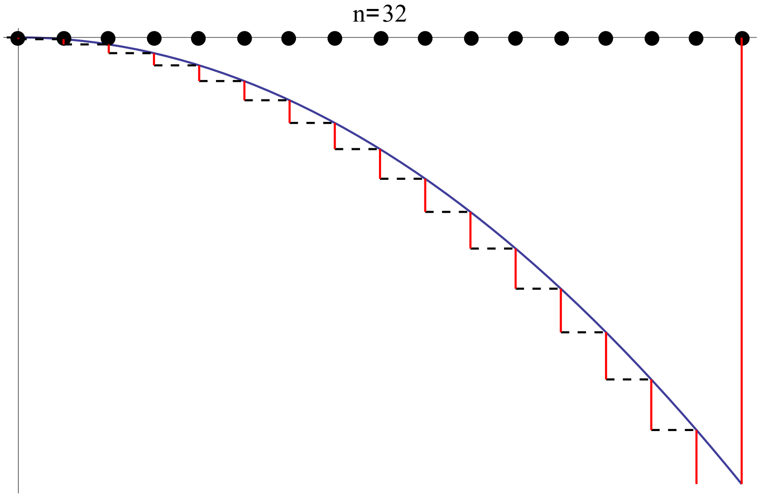

Figure 2.3: Splitted differences of g_n

The difference is splitted into several easier ones by the following steps:

$$\begin{aligned} g_n(y_{k+1,n}) - g_n(0) &= \left[g_n(y_{\frac{n}{2}+1,n}) - g_n(\overbrace{y_{\frac{n}{2},n}}^{0})\right] + \left[g_n(y_{\frac{n}{2}+2,n}) - g_n(y_{\frac{n}{2}+1,n})\right] + \\ &\quad \ldots + [g_n(y_{kn}) - g_n(y_{k-1,n})] + [g_n(y_{k+1,n}) - g_n(y_{kn})] \\ &= \sum_{i=\frac{n}{2}}^{k} [g_n(y_{i+1,n}) - g_n(y_{in})] \end{aligned} \tag{2.18}$$

Thus, it is important to know the value of $g_n(y_{k+1,n}) - g_n(y_{kn})$. If we recognize that Equation (2.7) is exactly what we need, we can transform it to

$$\frac{f_n(y_{k+1,n})}{f_n(y_{kn})} = \frac{\frac{n}{2} - \sqrt{n}\, y_{kn}}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1}.$$

After some algebraic steps, we arrive at a first result:

$$\begin{aligned}
\ln \frac{f_n(y_{k+1,n})}{f_n(y_{kn})} &= \ln \frac{\frac{n}{2} - \sqrt{n}\, y_{kn}}{\sqrt{n}\, y_{kn} + \frac{n}{2} + 1} \\
\Leftrightarrow \ln f_n(y_{k+1,n}) - \ln f_n(y_{kn}) &= \ln \left(\frac{n}{2} - \sqrt{n}\, y_{kn}\right) - \ln \left(\sqrt{n}\, y_{kn} + \frac{n}{2} + 1\right) \\
\Leftrightarrow \qquad g_n(y_{k+1,n}) - g_n(y_{kn}) &= \ln \left(\frac{n}{2} - \sqrt{n}\, \frac{k - \frac{n}{2}}{\sqrt{n}}\right) - \ln \left(\sqrt{n}\, \frac{k - \frac{n}{2}}{\sqrt{n}} + \frac{n}{2} + 1\right) \\
&= \ln(n-k) - \ln(k+1) \qquad (2.19)
\end{aligned}$$

At this step, Krylov has a perfect idea and introduces a new variable transformation for further calculations. The transformation is

$$x_{kn} = \frac{2y_{kn}}{\sqrt{n}} = \frac{2\left(k - \frac{n}{2}\right)}{n} = \frac{2k}{n} - 1.$$

The reverse transformation from x_{kn} to k is thus

$$\begin{aligned}
x_{kn} &= \frac{2k}{n} - 1 \\
\Leftrightarrow n\, x_{kn} &= 2k \\
\Leftrightarrow \qquad k &= \frac{n}{2}\, x_{kn} + \frac{n}{2}. \qquad (2.20)
\end{aligned}$$

Also here, it is useful to know what $x_{k+1,n}$ means:

$$x_{k+1,n} = \frac{2(k+1)}{n} - 1 = \frac{2k+2}{n} - 1 = \frac{2k}{n} - 1 + \frac{2}{n} = x_{kn} + \frac{2}{n}$$

This means that the distance between x_{kn} and $x_{k+1,n}$ is $\frac{2}{n}$ which converges to 0 as $n \to \infty$.

Finally with Equation (2.20), Equation (2.19) gives

$$\begin{aligned} g_n(y_{k+1,n}) - g_n(y_{kn}) &= \ln\left(n - \frac{n}{2} x_{kn} - \frac{n}{2}\right) - \ln\left(\frac{n}{2} x_{kn} + \frac{n}{2} + 1\right) \\ &= \ln\left(\frac{n}{2}(1 - x_{kn})\right) - \ln\left(\frac{n}{2}\left(1 + x_{kn} + \frac{2}{n}\right)\right) \\ &= \ln(1 - x_{kn}) - \ln\left(1 + x_{kn} + \frac{2}{n}\right) \\ &= \ln(1 - x_{kn}) - \ln\left(1 + x_{k+1,n}\right). \end{aligned}$$

We continue like Krylov and develop the equation above into a Taylor series until order 3:

$$\begin{aligned} g_n(y_{k+1,n}) - g_n(y_{kn}) &= \ln(1 - x_{kn}) - \ln\left(1 + x_{k+1,n}\right) \\ &= -\left(x_{k+1,n} + x_{kn}\right) + \frac{1}{2}\left(x_{k+1,n}^2 - x_{kn}^2\right) - \frac{1}{3}\left(\bar{x}_{k+1,n}^3 + \bar{x}_{kn}^3\right) \end{aligned}$$

where $\bar{x}_{kn} \leq x_{kn}$ and $\bar{x}_{k+1,n} \leq x_{k+1,n}$.

We will put this result in Equation (2.18) and get

$$\begin{aligned} \sum_{i=\frac{n}{2}}^{k} \left[g_n(y_{i+1,n}) - g_n(y_{in})\right] &= \sum_{i=\frac{n}{2}}^{k} \Bigg[- x_{in} - x_{i+1,n} + \frac{1}{2}\left(x_{i+1,n}^2 - x_{in}^2\right) - \\ &\quad \ldots - \frac{1}{3}\left(\bar{x}_{in}^3 + \bar{x}_{i+1,n}^3\right) \Bigg] \\ &= - \sum_{i=\frac{n}{2}}^{k} \left(x_{in} + x_{i+1,n}\right) + \frac{1}{2} \sum_{i=\frac{n}{2}}^{k} \left(x_{i+1,n}^2 - x_{in}^2\right) - \\ &\quad \ldots - \frac{1}{3} \sum_{i=\frac{n}{2}}^{k} \left(\bar{x}_{in}^3 + \bar{x}_{i+1,n}^3\right). \end{aligned} \tag{2.21}$$

We will compute every single sum step by step. Beforehand we will clarify some basics for these computations. First, it is sufficient to concentrate on $k \geq \frac{n}{2}$ (or similarly $y_{kn} \geq 0$) due to symmetry reasons. It follows that for $\frac{n}{2} \leq k \leq \frac{n}{2} + n^{\beta}$:

$$\begin{aligned}
\frac{n}{2} \leq \quad k \quad &\leq \frac{n}{2} + n^{\beta} \\
0 \leq k - \frac{n}{2} &\leq n^{\beta} \\
0 \leq \frac{k - \frac{n}{2}}{\sqrt{n}} &\leq n^{\beta - \frac{1}{2}} \\
0 \leq \quad y_{kn} \quad &\leq n^{\beta - \frac{1}{2}}
\end{aligned}$$

Additionally, if we use the transformation Equation (2.20), we get the order of x_{kn}:

$$\begin{aligned}
0 \leq \quad \frac{\sqrt{n}\, x_{kn}}{2} \quad &\leq n^{\beta - \frac{1}{2}} \\
0 \leq \quad \sqrt{n}\, x_{kn} \quad &\leq 2n^{\beta - \frac{1}{2}} \\
0 \leq \quad x_{kn} \quad &\leq 2n^{\beta - 1} \\
\Rightarrow x_{kn} = \mathcal{O}\left(n^{\beta - 1}\right)
\end{aligned}$$

Using this result, we can also find the order of $\bar{x}_{kn}^3 + \bar{x}_{k+1,n}^3$:

$$\begin{aligned}
\bar{x}_{kn}^3 + \bar{x}_{k+1,n}^3 \leq x_{kn}^3 + x_{k+1,n}^3 &= \mathcal{O}\left(n^{\beta-1}\right)^3 + \mathcal{O}\left(n^{\beta-1}\right)^3 \\
&= \mathcal{O}\left(n^{3\beta-3}\right)
\end{aligned}$$

We will also need two elementary sums for the computations. These are

$$\sum_{i=\frac{n}{2}}^{k} 1 = \underbrace{1 + 1 + \ldots + 1}_{k+1-\frac{n}{2}} = k + 1 - \frac{n}{2}$$

and

$$\begin{aligned}
\sum_{i=\frac{n}{2}}^{k} i &= \frac{n}{2} + \left(\frac{n}{2}+1\right) + \left(\frac{n}{2}+2\right) + \ldots + k \\
&= \frac{n}{2}(\underbrace{1+1+\ldots+1}_{k+1-\frac{n}{2}}) + \left(1+2+\ldots+\left(k-\frac{n}{2}\right)\right) \\
&= \frac{n}{2}\left(k+1-\frac{n}{2}\right) + \sum_{i=1}^{k-\frac{n}{2}} i \\
&= \frac{n\left(k+1-\frac{n}{2}\right)}{2} + \frac{\left(k-\frac{n}{2}\right)\left(k+1-\frac{n}{2}\right)}{2} \\
&= \frac{1}{2}\left(k+1-\frac{n}{2}\right)\left(n+k-\frac{n}{2}\right) \\
&= \frac{1}{2}\left(k+1-\frac{n}{2}\right)\left(k+\frac{n}{2}\right).
\end{aligned}$$

With these basics, we now can compute the three sums from Equation (2.21). The first one is calculated by

$$\begin{aligned}
\sum_{i=\frac{n}{2}}^{k} (x_{i+1,n} + x_{in}) &= \sum_{i=\frac{n}{2}}^{k} \left(\frac{2i+2}{n} - 1 + \frac{2i}{n} - 1\right) = \sum_{i=\frac{n}{2}}^{k} \left(\frac{4i+2}{n} - 2\right) \\
&= \sum_{i=\frac{n}{2}}^{k} \left(\frac{4}{n}\left(i-\frac{n}{2}\right) + \frac{2}{n}\right) = \frac{4}{n}\sum_{i=\frac{n}{2}}^{k}\left(i-\frac{n}{2}\right) + \frac{2}{n}\sum_{i=\frac{n}{2}}^{k} 1 \\
&= \frac{4}{n}\sum_{i=1}^{k-\frac{n}{2}} i + \frac{2}{n}\sum_{i=\frac{n}{2}}^{k} 1 \\
&= \frac{4}{n}\cdot\frac{\left(k+1-\frac{n}{2}\right)\left(k-\frac{n}{2}\right)}{2} + \frac{2}{n}\left(k+1-\frac{n}{2}\right) \\
&= \frac{2}{n}\left(k+1-\frac{n}{2}\right)\left(k-\frac{n}{2}\right) + \frac{2}{n}\left(k+1-\frac{n}{2}\right) \\
&= \frac{2}{n}\left(k+1-\frac{n}{2}\right)\left(1+k-\frac{n}{2}\right) = \frac{2}{n}\left(k+1-\frac{n}{2}\right)^2 \\
&= 2\left(\frac{\left(k+1-\frac{n}{2}\right)^2}{\sqrt{n}^2}\right) = 2\left(\frac{k+1-\frac{n}{2}}{\sqrt{n}}\right)^2 \\
&= 2y_{k+1,n}^2.
\end{aligned}$$

The computation for the second sum is

$$
\begin{aligned}
\sum_{i=\frac{n}{2}}^{k}\left(x_{i+1,n}^{2}-x_{in}^{2}\right) &= \sum_{i=\frac{n}{2}}^{k}\left[\left(\frac{2i+2}{n}-1\right)^{2}-\left(\frac{2i}{n}-1\right)^{2}\right] \\
&= \sum_{i=\frac{n}{2}}^{k}\left[\left(\frac{4i^{2}+8i+4}{n^{2}}-\frac{4i+4}{n}+1\right)-\left(\frac{4i^{2}}{n^{2}}-\frac{4i}{n}+1\right)\right] \\
&= \sum_{i=\frac{n}{2}}^{k}\left(\frac{8i+4}{n^{2}}-\frac{4}{n}\right) \\
&= \frac{1}{n^{2}}\left[8\sum_{i=\frac{n}{2}}^{k} i+4\sum_{i=\frac{n}{2}}^{k} 1\right]-\frac{4}{n}\left(k+1-\frac{n}{2}\right) \\
&= \frac{1}{n^{2}}\left[8\left(\frac{1}{2}\left(k+1-\frac{n}{2}\right)\left(k+\frac{n}{2}\right)\right)+4\left(k+1-\frac{n}{2}\right)\right]-\frac{4}{n}\left(k+1-\frac{n}{2}\right) \\
&= \frac{1}{n^{2}}\left[4\left(k+1-\frac{n}{2}\right)\left(k+\frac{n}{2}\right)+4\left(k+1-\frac{n}{2}\right)\right]-\frac{4}{n}\left(k+1-\frac{n}{2}\right) \\
&= \frac{1}{n^{2}}\left[4\left(k^{2}+\frac{kn}{2}+k+\frac{n}{2}-\frac{kn}{2}-\frac{n^{2}}{4}\right)+4\left(k+1-\frac{n}{2}\right)\right] \\
&\ldots-\frac{1}{n}\left(4k+4-2n\right) \\
&= \frac{1}{n^{2}}\left[4k^{2}+8k+4\right]-1-\frac{1}{n}\left(4k+4\right)+2 \\
&= \frac{\left(2k+2\right)^{2}}{n^{2}}-2\cdot\frac{2k+2}{n}+1 \\
&= \left(\frac{2\left(k+1\right)}{n}-1\right)^{2} \\
&= x_{k+1,n}^{2}=\mathcal{O}\left(n^{2\beta-2}\right).
\end{aligned}
$$

Finally, we get

$$\begin{aligned}
g_n(y_{k+1,n}) - g_n(0) &= \sum_{i=\frac{n}{2}}^{k} \left[g_n(y_{i+1,n}) - g_n(y_{in})\right] \\
&= -\sum_{i=\frac{n}{2}}^{k} (x_{i+1,n} + x_{in}) + \frac{1}{2}\sum_{i=\frac{n}{2}}^{k} \left(x_{i+1,n}^2 - x_{in}^2\right) \\
&\dots - \frac{1}{3}\sum_{i=\frac{n}{2}}^{k} \left(\bar{x}_{i+1,n}^3 + \bar{x}_{in}^3\right) \\
&= -2y_{k+1,n}^2 + \mathcal{O}\left(n^{2\beta-2}\right) + \mathcal{O}\left(n^{3\beta-3}\right) \\
&= -2y_{k+1,n}^2 + \mathcal{O}\left(n^{2\beta-2}\right).
\end{aligned}$$

Regarding the penultimate step, the question was, which of the two orders is greater. Let us try to find out for which values of β, $\mathcal{O}\left(n^{2\beta-2}\right)$ is greater or equal than $\mathcal{O}\left(n^{3\beta-3}\right)$:

$$\begin{aligned}
3\beta - 3 &\leq 2\beta - 2 \\
3\beta &\leq 2\beta + 1 \\
\beta &\leq 1
\end{aligned}$$

Hence, we have a new interval for $\beta \Rightarrow \beta \in \left(\frac{1}{2}, 1\right]$. This says that $\mathcal{O}\left(n^{2\beta-2}\right)$ is, for the given interval of β, greater than $\mathcal{O}\left(n^{3\beta-3}\right)$.

Thus, we have proved Equation (2.17) with $z(n) = n^{2\beta-2}$. Equation (2.16) can be updated to

$$\phi_n(0)^{-1}\, e^{2y_{kn}^2}\, \phi_n(y_{kn}) = e^{\mathcal{O}\left(n^{2\beta-2}\right)} = 1 + \mathcal{O}\left(n^{2\beta-2}\right) \tag{2.22}$$

or similarly

$$f_n(0)^{-1}\, e^{2y_{kn}^2}\, f_n(y_{kn}) = e^{\mathcal{O}\left(n^{2\beta-2}\right)} = 1 + \mathcal{O}\left(n^{2\beta-2}\right).$$

The equality of $e^{\mathcal{O}\left(n^{2\beta-2}\right)}$ and $1+\mathcal{O}\left(n^{2\beta-2}\right)$ can be proved by developing $e^{\mathcal{O}\left(n^{2\beta-2}\right)}$

into a Taylor series:

$$\begin{aligned} e^{\mathcal{O}\left(n^{2\beta-2}\right)} &= 1+\mathcal{O}\left(n^{2\beta-2}\right)+\frac{1}{2}\,\mathcal{O}\left(n^{2\beta-2}\right)^2+\frac{1}{6}\,\mathcal{O}\left(n^{2\beta-2}\right)^3+\mathcal{O}\left(\mathcal{O}\left(n^{2\beta-2}\right)^4\right) \\ &= 1+\mathcal{O}\left(n^{2\beta-2}\right)+\mathcal{O}\left(n^{4\beta-4}\right)+\mathcal{O}\left(n^{6\beta-6}\right)+\mathcal{O}\left(n^{8\beta-8}\right) \\ &= 1+\mathcal{O}\left(n^{2\beta-2}\right). \end{aligned}$$

Regarding the penultimate step again, we had to decide which of the orders is the greatest. Using the method from above, we can see that $n^{2\beta-2}$ has the greatest order.

So finally we proved that for $\beta \in \left(\frac{1}{2}, 1\right]$

$$f_n(0)^{-1}\, e^{2y_{kn}^2}\, f_n(y_{kn}) = 1+\mathcal{O}\left(n^{2\beta-2}\right). \tag{2.23}$$

2.2.2 Part 2

As a second part, Krylov focuses on determining the behavior of $f_n(0)$ as $n \to \infty$. First, let us say again that

$$c_n = f_n(0) = P\left(S_n = \frac{n}{2}\right).$$

Now, let us rewrite Equation (2.23). Multiplying by $\frac{1}{\sqrt{n}}$ and $e^{-2y_{kn}^2}$ gives

$$\frac{1}{c_n\sqrt{n}}\underbrace{f_n(y_{kn})}_{P(S_n=k)} = \frac{1}{\sqrt{n}}\, e^{-2y_{kn}^2}+\mathcal{O}\left(n^{2\beta-2.5}\right). \tag{2.24}$$

In the next step, Krylov not only regards one probability but a sum of those. So, he sums all probabilities within a certain interval which is $|S_n - \frac{n}{2}| \leq n^{\frac{3}{5}}$ in his case. We will consider the more general case $|S_n - \frac{n}{2}| \leq n^{\beta}$ which results in the equation

$$\frac{1}{c_n\sqrt{n}} \sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} P(S_n = k) = \frac{1}{\sqrt{n}} \sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2} + \sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} \mathcal{O}\left(n^{2\beta-2.5}\right). \tag{2.25}$$

The next step is to find out what the sum $\sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} P(S_n = k)$ means. Using Equation (2.5), we get

$$\begin{aligned}
\sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} P(S_n = k) &= \sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} P\left(\frac{S_n - \frac{n}{2}}{\sqrt{n}} = y_{kn}\right) \\
&= \sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} P\left(\left|\frac{S_n - \frac{n}{2}}{\sqrt{n}}\right| = |y_{kn}|\right) \\
&= P\left(\left|\frac{S_n - \frac{n}{2}}{\sqrt{n}}\right| \leq n^{\beta-\frac{1}{2}}\right) \\
&= P\left(\left|S_n - \frac{n}{2}\right| \leq n^{\beta}\right). \qquad (2.26)
\end{aligned}$$

Now, we have to determine the value of $\sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} \mathcal{O}\left(n^{2\beta-2.5}\right)$. Since $\mathcal{O}\left(n^{2\beta-2.5}\right)$ does not depend on k, it is enough to find the cardinality of the set $I = \{k : |y_{kn}| \leq n^{\beta-\frac{1}{2}}\}$ and to multiply it by $\mathcal{O}\left(n^{2\beta-2.5}\right)$. The following steps show how to find the cardinality of I:

$$\begin{aligned}
-n^{\beta-\frac{1}{2}} \leq &\quad y_{kn} &\leq n^{\beta-\frac{1}{2}} \\
-n^{\beta} \leq &\; k - \frac{n}{2} &\leq n^{\beta} \\
-n^{\beta} + \frac{n}{2} \leq &\quad k &\leq n^{\beta} + \frac{n}{2}
\end{aligned}$$

Thus, I contains about $n^{\beta} + \frac{n}{2} - \left(-n^{\beta} + \frac{n}{2}\right) = 2n^{\beta}$ elements.

Equation (2.25) therefore reduces to

$$\frac{1}{c_n\sqrt{n}} P\left(\left|S_n - \frac{n}{2}\right| \leq n^{\beta}\right) = \frac{1}{\sqrt{n}} \sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2} + \underbrace{2n^{\beta} \cdot \mathcal{O}\left(n^{2\beta-2.5}\right)}_{\mathcal{O}\left(n^{3\beta-2.5}\right)}.$$

Now, Krylov regards the limit on both sides. Of course, it would be great if $\mathcal{O}\left(n^{3\beta-2.5}\right)$ disappears. The condition that this holds is

$$3\beta - 2.5 < 0 \Rightarrow \beta < \frac{5}{6}. \qquad (2.27)$$

Thus, we have a new condition for the values of β: $\beta \in \left(\frac{1}{2}, \frac{5}{6}\right)$.

This leads to the equation

$$\lim_{n\to\infty} \frac{1}{c_n\sqrt{n}} P\left(\left|S_n - \frac{n}{2}\right| \leq n^\beta\right) = \lim_{n\to\infty} J_n \tag{2.28}$$

where

$$J_n := \frac{1}{\sqrt{n}} \sum_{k:|y_{kn}|\leq n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2}. \tag{2.29}$$

The probability $P\left(\left|S_n - \frac{n}{2}\right| \leq n^\beta\right)$ can be estimated by Chebyshev's inequality. You can find a short proof of this theorem in Appendix A.3.

Thus it holds,

$$\begin{aligned} 1 \geq P\left(\left|S_n - \frac{n}{2}\right| \leq n^\beta\right) &= 1 - P\left(\left|S_n - \frac{n}{2}\right| \geq n^\beta\right) \\ \geq 1 - \frac{\frac{n}{4}}{n^{2\beta}} &= 1 - \frac{1}{4} n\, n^{-2\beta} = 1 - \frac{1}{4} n^{1-2\beta} \geq 1 - n^{1-2\beta}. \end{aligned}$$

Regarding the limit of the probability, it would be fine if the term $n^{1-2\beta}$ vanishes. For this, it must hold $1-2\beta < 0 \Rightarrow \beta > \frac{1}{2}$. This condition is absolutely compatible with the condition $\beta \in \left(\frac{1}{2}, \frac{5}{6}\right)$.

So, we get

$$\lim_{n\to\infty} P\left(\left|S_n - \frac{n}{2}\right| \leq n^\beta\right) = 1.$$

for $\beta \in \left(\frac{1}{2}, \frac{5}{6}\right)$. Equation (2.28) thus reduces to

$$\lim_{n\to\infty} \frac{1}{c_n\sqrt{n}} = \lim_{n\to\infty} J_n. \tag{2.30}$$

Now, it is time to find $\lim_{n\to\infty} J_n$. Due to the symmetry of the Gaussian function, the definition of J_n (see Equation (2.29)) can be rewritten as

$$\begin{aligned} J_n &= \frac{1}{\sqrt{n}} \sum_{k:|y_{kn}|\le n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2} \\ &= \frac{1}{\sqrt{n}}\left[\sum_{k<\frac{n}{2}:y_{kn}\le n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2} + \underbrace{\sum_{k=\frac{n}{2}} e^{-2y_{kn}^2}}_{1} + \sum_{k>\frac{n}{2}:y_{kn}\le n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2} \right] \\ &= \frac{1}{\sqrt{n}} + 2\cdot\frac{1}{\sqrt{n}} \sum_{k>\frac{n}{2}:y_{kn}\le n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2}. \end{aligned} \tag{2.31}$$

To calculate the sum

$$\sum_{k>\frac{n}{2}:y_{kn}\le n^{\beta-\frac{1}{2}}} e^{-2y_{kn}^2}, \tag{2.32}$$

it is important to notice some properties. By looking at Figure 2.4 we can see that for $k \ge \frac{n}{2}$

$$e^{-2y_{kn}^2} \le e^{-2y^2} \quad \text{for } 0 \le y \le y_{kn}, \qquad e^{-2y_{kn}^2} \ge e^{-2y_{kn}^2} \quad \text{for } 0 \le y_{kn} \le y. \tag{2.33}$$

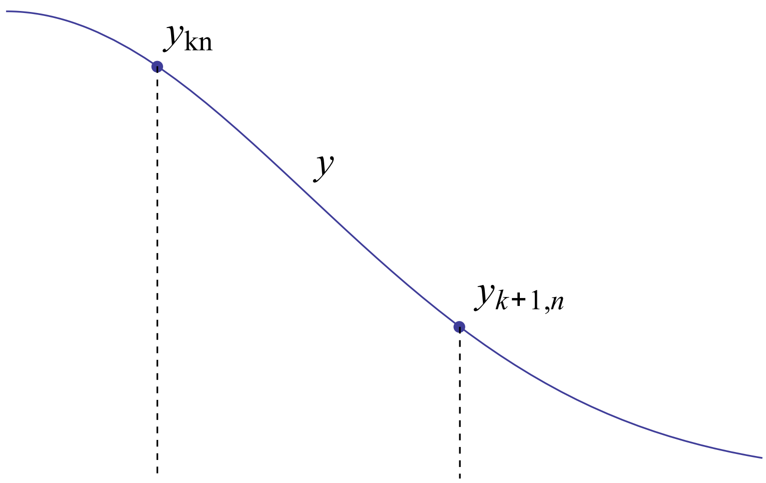

Figure 2.4: y_{kn}, y and $y_{k+1,n}$ on Gaussian function

Thus, for $k > \frac{n}{2}$ we also can observe in Figure 2.5 that

$$\underbrace{\int_{y_{k-1,n}}^{y_{kn}} e^{-2y^2}dy}_{A} \geq \underbrace{\frac{1}{\sqrt{n}}e^{-2y_{kn}^2}}_{B} \geq \underbrace{\int_{y_{kn}}^{y_{k+1,n}} e^{-2y^2}dy}_{C}. \tag{2.34}$$

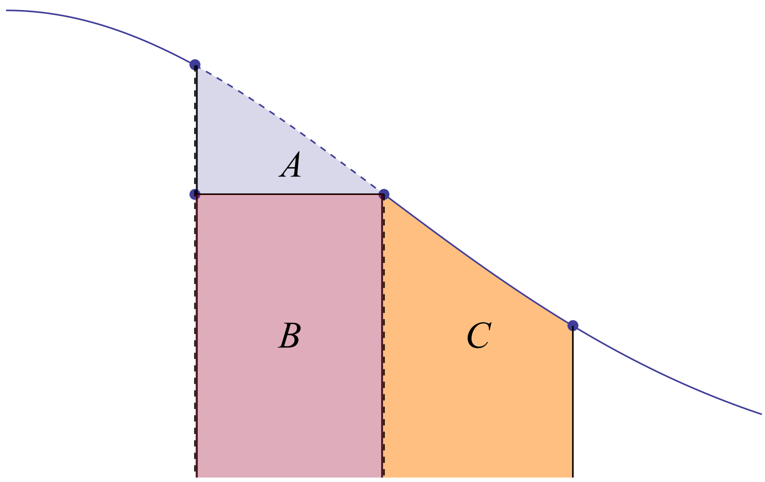

Figure 2.5: Areas under Gaussian function (A: Dashed blue area, B: Violet blue area, C: Orange area)

By Equation (2.31) and Inequality (2.34), we can infer that

$$\sum_{k>\frac{n}{2}:y_{kn}\leq n^{\beta-\frac{1}{2}}} \int_{y_{k-1,n}}^{y_{kn}} e^{-2y^2}dy \geq \frac{J_n - \frac{1}{\sqrt{n}}}{2} \geq \sum_{k>\frac{n}{2}:y_{kn}\leq n^{\beta-\frac{1}{2}}} \int_{y_{kn}}^{y_{k+1,n}} e^{-2y^2}dy \tag{2.35}$$

$$\int_0^{\bar{y}_n} e^{-2y^2}dy \geq \frac{J_n - \frac{1}{\sqrt{n}}}{2} \geq \int_{y_{\frac{n}{2}+1,n}}^{\bar{y}_n} e^{-2y^2}dy \tag{2.36}$$

$$\int_0^{\bar{y}_n} e^{-2y^2}dy \geq \frac{J_n - \frac{1}{\sqrt{n}}}{2} \geq \int_0^{\bar{y}_n} e^{-2y^2}dy - \int_0^{y_{\frac{n}{2}+1,n}} e^{-2y^2}dy \tag{2.37}$$

where $\bar{y}_n$ is the largest y_{kn} such that $y_{kn} \leq n^{\beta-\frac{1}{2}}$. Because of

$$y_{\frac{n}{2}+1,n} = \frac{\frac{n}{2}+1-\frac{n}{2}}{\sqrt{n}} = \frac{1}{\sqrt{n}},$$

by building the limit for $n \to \infty$ of all terms, we get

$$\lim_{n\to\infty} \int_0^{\bar{y}_n} e^{-2y^2} dy \geq \lim_{n\to\infty} \frac{J_n - \frac{1}{\sqrt{n}}}{2} \geq \lim_{n\to\infty} \int_0^{\bar{y}_n} e^{-2y^2} dy - \underbrace{\lim_{n\to\infty} \int_0^{\frac{1}{\sqrt{n}}} e^{-2y^2} dy}_{0}$$

$$\Rightarrow \lim_{n\to\infty} \frac{J_n - \frac{1}{\sqrt{n}}}{2} = \lim_{n\to\infty} \int_0^{\bar{y}_n} e^{-2y^2} dy.$$

It is obviously that $\bar{y}_n \to \infty$ as $n \to \infty$. So, finally by adjusting the calculation in Appendix A.1, we get

$$\begin{aligned} \lim_{n\to\infty} J_n &= 2 \int_0^{\infty} e^{-2y^2} dy \\ &= \sqrt{\frac{\pi}{2}}. \end{aligned} \tag{2.38}$$

Combining this result with Equation (2.30) leads to

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{c_n \sqrt{n}} = \sqrt{\frac{\pi}{2}} \\ \Rightarrow &\lim_{n\to\infty} c_n \sqrt{n} = \frac{1}{\sqrt{\frac{\pi}{2}}} = \sqrt{\frac{2}{\pi}}. \end{aligned} \tag{2.39}$$

2.2.3 Part 3

With few more steps, we will come to another beautiful result. Let us regard Equation (2.28) again and put in the result from Equation (2.39):

$$\sqrt{\frac{\pi}{2}} \lim_{n\to\infty} P\left(\left|S_n - \frac{n}{2}\right| \leq n^{\beta}\right) = \lim_{n\to\infty} J_n$$

Now, we change the limits of the interval for $S_n - \frac{n}{2}$. Let us set $\beta = \frac{1}{2}$ and regard the interval $a\sqrt{n} \leq S_n - \frac{n}{2} \leq b\sqrt{n}$ such that $a < b$. This gives

$$\sqrt{\frac{\pi}{2}} \lim_{n\to\infty} P\left(a \leq \frac{S_n - \frac{n}{2}}{\sqrt{n}} \leq b\right) = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k:a\leq y_{kn}\leq b} e^{-2y_{kn}^2}$$

If we take a closer look at the right hand side of the equation above, we recognize that it is a proper integral expression. $\frac{1}{\sqrt{n}}$ is the width of the „Riemann rectangles“ and $e^{-2y_{kn}^2}$ is the function value of e^{-2y^2} at y_{kn}. As we saw at Equation (2.2), the distance between two y_{kn}'s is indeed $\frac{1}{\sqrt{n}}$. Thus, finally we get the same beautiful result as Krylov, namely

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - \frac{n}{2}}{\sqrt{n}} \leq b\right) = \sqrt{\frac{2}{\pi}} \int_a^b e^{-2y^2} dy. \tag{2.40}$$

The general one-dimensional cumulative distribution function of the normal distribution is

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

where μ is the expectation value and σ^2 the variance of the random variable considered.

In our case, we have $\mu = 0$ and $\sigma^2 = \frac{1}{4}$. If we put these values in the cumulative distribution function from above, we get

$$\begin{aligned} F(x) &= \frac{1}{\sqrt{\frac{\pi}{2}}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(4t^2\right)} dt \\ &= \sqrt{\frac{2}{\pi}} \int_{-\infty}^{x} e^{-2t^2} dt. \end{aligned}$$

Hence, for a normal distributed random variable X it holds that

$$P(a \leq X \leq b) = F(b) - F(a) = \sqrt{\frac{2}{\pi}} \int_a^b e^{-2t^2} dt.$$

But this is the same result as we found at Equation (2.40). Therefore we have shown that the cumulative distribution function of the binomial distribution converges to the one of the normal distribution as $n \to \infty$.

2.2.4 Part 4

As a corollary of Part 2 (2.2.2), we can also show that the maximal error for $\beta \in \left(\frac{1}{2}, \frac{5}{6}\right)$ converges to zero as $n \to \infty$. We have already shown it graphically in Section 1.4.

Let us again regard Equation (2.22) and some manipulations of it to see the result:

$$
\begin{aligned}
& & \phi_n(0)^{-1} e^{-2y_{kn}^2} \phi_n(y_{kn}) &= 1 + \mathcal{O}\left(n^{2\beta-2}\right) \\
\Leftrightarrow & & \frac{1}{c_n\sqrt{n}} e^{-2y_{kn}^2} \sqrt{n} f_n(y_{kn}) &= 1 + \mathcal{O}\left(n^{2\beta-2}\right) \\
\Leftrightarrow & & \max_{|k-\frac{n}{2}|\leq n^\beta} \left| \frac{1}{c_n\sqrt{n}} e^{-2y_{kn}^2} \sqrt{n} f_n(y_{kn}) - 1 \right| &= \max_{|k-\frac{n}{2}|\leq n^\beta} \mathcal{O}\left(n^{2\beta-2}\right) \\
\Leftrightarrow & & \lim_{n\to\infty} \max_{|k-\frac{n}{2}|\leq n^\beta} \left| \sqrt{\frac{n\pi}{2}} e^{-2y_{kn}^2} f_n(y_{kn}) - 1 \right| &= 0
\end{aligned}
$$

This is exactly the result which Krylov also comes to.

Chapter 3

Comments

3.1 Limits of the β-Interval

In Krylov's paper, he uses $\left[\frac{1}{2}, \frac{3}{4}\right]$ as the interval for β. This interval was used for all parts. As we saw, our interval was chosen to $\left(\frac{1}{2}, 1\right]$ for the first part and to $\left(\frac{1}{2}, \frac{5}{6}\right)$ for the remainder parts.

During a step in Part 1 (2.2.1), in which Krylov computed $\sum_{i=\frac{n}{2}}^{k}\left(x_{in}^2 + x_{i+1,n}^2\right)$, he says that $\mathcal{O}\left(n^{2\beta-2}\right) = \mathcal{O}\left(n^{4\beta-3}\right)$. He hence restricts the interval for β due to the order of $n^{4\beta-3}$ to $\left[\frac{1}{2}, \frac{3}{4}\right]$.

I could not find his reason for this restriction, and so I chose my intervals for this thesis.

3.2 The Proof for *odd* n

The proof discussed in Chapter 2 assumes that n is an even number. If we want to do the proof with odd n, we have to be careful at some points. But first, let us regard the most evident difference between a binomial distributed random variable with an even and an odd number of trials.

In Figure 3.1 we can see that the binomial distribution for even n has one value with the maximum probability. In contrast, for an odd number of trials the distribution has two values with the maximum probability.

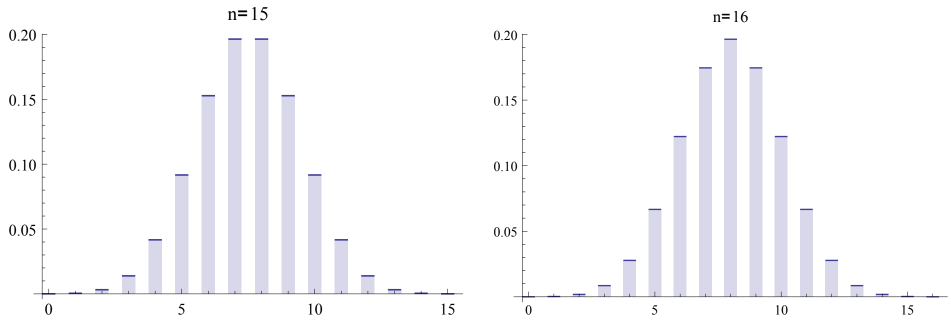

Figure 3.1: Binomial distribution with $p = \frac{1}{2}$ for $n = 15$ and $n = 16$

To go on with the proof, we have to treat the parts 1-4 differently. The number of trails n is now an odd number unless stated otherwise.

In Part 1, Krylov uses $i = \frac{n}{2}$ as the starting index for the sums. This can not be done in the odd case. I tried to do the same calculations as we did for even n but for $i = \frac{n+1}{2}$. Unfortunately, I did not come to any result with this method. I guess that the x_{kn}-transformation has to be adjusted such that we get similar results as in the even case. Let us say that z_{kn} is this transformation. Then it has to satisfy the following equation:

$$\sum_{i=\frac{n+1}{2}}^{k} (z_{in} + z_{i+1,n}) = 2y_{k+1,n}^2.$$

It is getting easier when we regard Part 2 (2.2.2). The sequence J_n was defined as

$$J_n := \frac{1}{\sqrt{n}} \sum_{k:|y_{kn}| \le n^{\beta - \frac{1}{2}}} e^{-2y_{kn}^2}.$$

In Equation (2.31), we divided the sum in three parts. Now the middle part for

$k = \frac{n}{2}$ disappears because n is odd. We therefore can rewrite J_n as

$$J_n = 2 \cdot \frac{1}{\sqrt{n}} \sum_{k > \frac{n}{2} : y_{kn} \leq n^{\beta - \frac{1}{2}}} e^{-2y_{kn}^2}.$$

The properties stated in Inequalities (2.33) also hold for odd n but this time for $k > \frac{n}{2}$. The Inequality (2.34) was stated for $k \geq \frac{n}{2}$. Here it is necessary to look at the inequality with $k \geq \frac{n+1}{2} = \lceil \frac{n}{2} \rceil$.

Inequalities (2.35)-(2.37) thus become

$$\sum_{k > \lceil \frac{n}{2} \rceil : y_{kn} \leq n^{\beta - \frac{1}{2}}} \int_{y_{k-1,n}}^{y_{kn}} e^{-2y^2} dy \geq \frac{J_n}{2} \geq \sum_{k > \lceil \frac{n}{2} \rceil : y_{kn} \leq n^{\beta - \frac{1}{2}}} \int_{y_{kn}}^{y_{k+1,n}} e^{-2y^2} dy$$

$$\int_{y_{\lceil \frac{n}{2} \rceil, n}}^{\bar{y}_n} e^{-2y^2} dy \geq \frac{J_n}{2} \geq \int_{y_{\lceil \frac{n}{2} \rceil + 1, n}}^{\bar{y}_n} e^{-2y^2} dy$$

$$\int_0^{\bar{y}_n} e^{-2y^2} dy - \int_0^{y_{\lceil \frac{n}{2} \rceil, n}} e^{-2y^2} dy \geq \frac{J_n}{2} \geq \int_0^{\bar{y}_n} e^{-2y^2} dy - \int_0^{y_{\lceil \frac{n}{2} \rceil + 1, n}} e^{-2y^2} dy$$

where $\bar{y}_n$ is the largest y_{kn} such that $y_{kn} \leq n^{\beta - \frac{1}{2}}$.

Because of

$$y_{\lceil \frac{n}{2} \rceil, n} = \frac{\frac{n+1}{2} - \frac{n}{2}}{\sqrt{n}} = \frac{\frac{1}{2}}{\sqrt{n}} = \frac{1}{2\sqrt{n}} \xrightarrow[n \to \infty]{} 0$$

and

$$y_{\lceil \frac{n}{2} \rceil + 1, n} = \frac{\frac{n+1}{2} + 1 - \frac{n}{2}}{\sqrt{n}} = \frac{\frac{3}{2}}{\sqrt{n}} = \frac{3}{2\sqrt{n}} \xrightarrow[n \to \infty]{} 0,$$

we get the same result for J_n as for the even case. Hence, it holds

$$\lim_{n \to \infty} J_n = 2 \int_0^{\infty} e^{-2y^2} dy = \sqrt{\frac{\pi}{2}}.$$

Now, we could continue as in Part 2 and get the same result.

A different treatment of even and odd n for Part 3 (2.2.3) and Part 4 (2.2.4) is not necessary.

Appendix

A.1 Gaussian Integral

We will look at two different ways to compute the Gaussian integral. The Gaussian integral is

$$A = \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}z^2\right) dz.$$

Due to the symmetry of the Gaussian function, it suffices to compute

$$I = \int_{0}^{\infty} \exp\left(-\frac{1}{2}z^2\right) dz$$

such that $A = 2I$.

1. Double integration

This method goes back to Laplace and was published in [Lap12]. First, it is important to note that

$$I = \int_{0}^{\infty} \exp\left(-\frac{1}{2}z^2\right) dz = \int_{0}^{\infty} \exp\left(-\frac{1}{2}(xy)^2\right) y\, dx$$

with the substitution $z = xy$. By renaming y in z it follows that

$$I = \int_{0}^{\infty} \exp\left(-\frac{1}{2}(xz)^2\right) z\, dx,$$

such that

$$\begin{aligned}
I^2 &= \left(\int_0^\infty \exp\left(-\frac{1}{2}z^2\right)dz\right)\left(\int_0^\infty \exp\left(-\frac{1}{2}(xz)^2\right)z\,dx\right) \\
&= \int_0^\infty \int_0^\infty \exp\left(-\frac{1}{2}\left(x^2+1\right)z^2\right)z\,dz\,dx.
\end{aligned}$$

Set $\left(x^2+1\right)z^2 = 2t$ such that $z\,dz = \frac{dt}{x^2+1}$ to get

$$\begin{aligned}
I^2 &= \int_0^\infty \int_0^\infty \exp\left(-t\right)\frac{dt}{x^2+1}dx = \left(\int_0^\infty \exp\left(-t\right)dt\right)\left(\int_0^\infty \frac{dx}{x^2+1}\right) \\
&= \left[-\exp(-t)\right]_0^\infty \left[\arctan(x)\right]_0^\infty \\
&= \frac{\pi}{2}
\end{aligned}$$

It follows that $I = \sqrt{\frac{\pi}{2}}$ such that

$$A = 2\sqrt{\frac{\pi}{2}} = \sqrt{2\pi}.$$

2. Double integration in the Cartesian coordinate system

This is the „usual“ double integration method. It was used by Siméon Denis Poisson and popularized by Sturm in [SP59]. As above, it is I^2 which is computed. So,

$$\begin{aligned}
I^2 &= \left(\int_0^\infty \exp\left(-\frac{1}{2}x^2\right)dx\right)\left(\int_0^\infty \exp\left(-\frac{1}{2}y^2\right)dy\right) \\
&= \int_0^\infty \exp\left(-\frac{1}{2}(x^2+y^2)\right)dx\,dy.
\end{aligned}$$

Now, the integral is computed with polar coordinates (r, θ) in which $dx\,dy = r\,dr\,d\theta$ such that

$$
\begin{aligned}
I^2 &= \int_0^{\frac{\pi}{2}} \int_0^{\infty} \exp\left(-\frac{1}{2}r^2\right) r\,dr\,d\theta \\
&= \left(\int_0^{\frac{\pi}{2}} d\theta\right) \left(\int_0^{\infty} \exp\left(-\frac{1}{2}r^2\right) r\,dr\right) \\
&= [\theta]_0^{\frac{\pi}{2}} \left[-exp\left(-\frac{1}{2}r^2\right)\right]_0^{\infty} \\
&= \frac{\pi}{2}
\end{aligned}
$$

As above, it follows that $I = \sqrt{\frac{\pi}{2}}$ and $A = 2\sqrt{\frac{\pi}{2}} = \sqrt{2\pi}$.

A.2 Motivation with Normalized Random Variable

Here, we carry out the same calculation as in Section 2.1, but without many comments due to similarity. We try to derive the proper density function of the normal distribution.

So, in contrast to Section 2.1, the random variable S_n is normalized to Z_n which means that $E(Z_n) = 0$ and $Var(Z_n) = 1$.

Thus,

$$Z_n = \frac{S_n - \mu_n}{\sigma_n},$$

where μ_n is the expectation value and σ_n^2 the variance of S_n.

To prove the normality of Z_n only a few calculation steps are necessary:

$$\begin{aligned} E(Z_n) &= E\left(\frac{S_n - \mu_n}{\sigma_n}\right) \\ &= \frac{1}{\sigma_n} E(S_n - \mu_n) \\ &= \frac{1}{\sigma_n} \left(E(S_n) - E(\mu_n)\right) \\ &= \frac{1}{\sigma_n} (\mu_n - \mu_n) = 0 \end{aligned}$$

$$\begin{aligned} Var(Z_n) &= Var\left(\frac{S_n - \mu_n}{\sigma_n}\right) \\ &= \frac{1}{\sigma_n^2} Var(S_n - \mu_n) \\ &= \frac{1}{\sigma_n^2} Var(S_n) \\ &= \frac{1}{\sigma_n^2} \sigma_n^2 = 1 \end{aligned}$$

The variables are specified similarly as in Section 2.1:

$$\begin{aligned}
y_{kn} &= \frac{k-\mu_n}{\sigma_n}, \qquad k = 0, 1, \ldots, n, \\
y_{k+1,n} &= \frac{k+1-\mu_n}{\sigma_n} \\
&= \frac{k-\mu_n}{\sigma_n} + \frac{1}{\sigma_n} \\
&= y_{kn} + \frac{1}{\sigma_n}, \\
k &= \sigma_n y_{kn} + \mu_n, \\
f_n(y_{kn}) &= P(S_n = k) = P\left(\frac{S_n - \mu_n}{\sigma_n} = y_{kn}\right)
\end{aligned}$$

Substituting these definitions in Equation (2.6) leads to

$$\begin{aligned}
&& f_n(y_{kn}) &= f_n(y_{k+1,n})\,\frac{\sigma_n y_{kn} + \mu_n + 1}{n - (\sigma_n y_{kn} + \mu_n)} \\
\Leftrightarrow && f_n\left(y_{kn} + \frac{1}{\sigma_n}\right) &= f_n(y_{kn})\,\frac{n - \sigma_n y_{kn} - \mu_n}{\sigma_n y_{kn} + \mu_n + 1} \\
\Leftrightarrow && f_n\left(y_{kn} + \frac{1}{\sigma_n}\right) - f_n(y_{kn}) &= f_n(y_{kn})\left(\frac{n - \sigma_n y_{kn} - \mu_n}{\sigma_n y_{kn} + \mu_n + 1} - 1\right) \\
\Leftrightarrow && f_n\left(y_{kn} + \frac{1}{\sigma_n}\right) - f_n(y_{kn}) &= f_n(y_{kn})\,\frac{n - 2\sigma_n y_{kn} - 2\mu_n - 1}{\sigma_n y_{kn} + \mu_n + 1}.
\end{aligned}$$

Because of $\phi_n(y_{kn}) := \sqrt{n} f_n(y_{kn}) \rightarrow \phi(y)$ as n $\rightarrow \infty$ and $y_{kn} \rightarrow y$, the equation becomes

$$\phi\left(y + \frac{1}{\sigma_n}\right) - \phi(y) = \phi(y)\,\frac{n - 2\sigma_n y - 2\mu_n - 1}{\sigma_n y + \mu_n + 1}.$$

Dividing both sides by $\frac{1}{\sigma_n}$ leads to

$$\frac{\phi\left(y + \frac{1}{\sigma_n}\right) - \phi(y)}{\frac{1}{\sigma_n}} = \phi(y)\,\frac{n\sigma_n - 2\sigma_n^2 y - 2\mu_n\sigma_n - \sigma_n}{\sigma_n y + \mu_n + 1}.$$

Now, it is time to put some concrete values in this equation. These values are $\mu_n = \frac{n}{2}$ and $\sigma_n^2 = \frac{n}{4}$ $\left(\sigma_n = \frac{\sqrt{n}}{2}\right)$, so that they let the equation look like

$$\frac{\phi\left(y + \frac{2}{\sqrt{n}}\right) - \phi(y)}{\frac{2}{\sqrt{n}}} = -\phi(y)\,\frac{\frac{n}{2}y + \frac{\sqrt{n}}{2}}{\frac{\sqrt{n}}{2}y + \frac{n}{2} + 1}.$$

With this, it is easy to get the ODE. We just have to let $n \to \infty$ on both sides which gives

$$\phi'(y) = -y\,\phi(y).$$

Without any problems, by using the integrating factors method again we can find the general solution of this ODE which is

$$\phi(y) = c\,e^{-\frac{1}{2}y^2}.$$

And here you can see the proper density function of the normal distribution. This is exactly what we wanted to show for the normalized random variable Z_n.

A.3 Short Proof of Chebyshev's Inequality

Theorem (Chebyshev's Inequality)**.** *Let X be an integrable random variable. Then for every $\epsilon > 0$*

$$P\left(|X - E(X)| \geq \epsilon\right) \leq \frac{Var(X)}{\epsilon^2}.$$

Proof. Obviously it is

$$\begin{aligned}|X - E(X)|^2 &\geq |X - E(X)|^2 1_{\{\omega:|X(\omega)-E(X)|\geq\epsilon\}}\\ &\geq \epsilon^2 1_{\{\omega:|X(\omega)-E(X)|\geq\epsilon\}}.\end{aligned}$$

Formation of the expectation value leads to

$$\begin{aligned}Var(X) &= E(|X - E(X)|^2)\\ &\geq E(\epsilon^2 1_{\{\omega:|X(\omega)-E(X)|\geq\epsilon\}})\\ &= \epsilon^2 P(|X - E(X)| \geq \epsilon).\end{aligned}$$

□

This proof is from [Irl05, p.132].

Bibliography

[Irl05] Albrecht Irle. *Wahrscheinlichkeitstheorie und Statistik: Grundlagen - Resultate - Anwendungen.* Teubner, Wiesbaden, 2 edition, 2005. 3, 11, VII

[Kry] Nicolai V. Krylov. *An Undergraduate Lecture on the Central Limit Theorem.* i

[Lap12] Pierre-Simon Laplace. Théorie analytiques des probabilités, 1812. I

[SP59] Charles Sturm and E. Prouhet. *Cours d'analyse de l'Ecole polytechnique.* Mallet-Bachelier, Paris, 1857-1859. II

[Tij04] Henc Tijms. *Understanding probability: Chance rules in everyday life.* Cambridge Univ. Press, New York and NY, 2004. 9

Teixeira Parente, Mario
(Nachname, Vorname)

IC 6/SS 2013
(Studiengruppe/Semester)

München, 21. August 2013
(Ort, Datum)

Erklärung

Hiermit erkläre ich, dass ich die Bachelorarbeit selbstständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

(Unterschrift)