

# The Local Limit Theorem: A Historical Perspective

D. R. McDonald

Department of Mathematics, University of Ottawa.  
(dmdsg@mathstat.uottawa.ca)

**Abstract.** The *local* limit theorem describes how the density of a sum of random variables follows the normal curve. However the local limit theorem is often seen as a curiosity of no particular importance when compared with the *central* limit theorem. Nevertheless the local limit theorem came first and is in fact associated with the foundation of probability theory by Blaise Pascal and Pierre de Fermat and was originally formalized by Jakob Bernoulli, Abraham DeMoivre and Pierre-Simon Laplace.

Here we describe the historical roots of the local limit theorem. We describe how it was supplanted by the central limit theorem in applications. Then we review the revival started by B. V. Gnedenko and we describe modern developments.

## 1 Introduction

When we say a histogram or probability density follows the normal curve we have in mind that the density resembles the density

---

Received: September 2005

*Key words and phrases:* Historical study, local limit theorem.

$\exp(-(z - \mu)^2/(2\sigma^2))/\sqrt{2\pi\sigma^2}$ ; i.e. is the normal density with mean  $\mu$  and standard deviation  $\sigma$ . One observes empirically that the histogram or empirical density of a sample of quality control measurements follows this normal density. Why is this?

A quality measurement (like the measurement of the width of ceramic tile) is in fact the sum of a large number of small errors (like variations in the mold, the temperature of the oven, the error in measuring length itself ...). The local limit theorem in the theory of probability states that the distribution of such sums tends to follow the normal curve. Consequently by the law of large numbers, the empirical histogram of equality measurements also tends to follow the normal. There are many proofs but in truth there is no *explanation* of this minor miracle.

To make this statement more precise consider a sequence of densities  $g_n$  with means  $a_n$  and standard deviations  $b_n$ . Any measurement  $x$  can be converted into standard units as follows,  $e_n(x) = (x - a_n)/b_n$ . We say the sequence of densities satisfies the local limit theorem if

$$b_n |g_n(x) - \frac{1}{b_n} \phi(e_n(x))| \rightarrow 0 \text{ where } \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

uniformly in  $x$  as  $n \rightarrow \infty$ .  $\phi(z)$  is called the standard normal density. Note that  $\phi(e_n(x))/b_n$  is a normal density with mean  $a_n$  and standard deviation  $b_n$  by a change of variable. The above local limit theorem says the density  $g_n$  may be approximated by a normal density with an error of size  $o(1/b_n)$  where  $o(\epsilon)$  is a function that satisfies  $\lim_{\epsilon \rightarrow 0} o(\epsilon)/\epsilon = 0$ . So the error gets small if  $b_n \rightarrow \infty$ .

Usually we are interested in the area under the density  $g_n$  between two values  $u_n$  and  $v_n$  say. Integrating the local limit theorem we have that

$$\left| \int_{u_n}^{v_n} g_n(x) dx - \int_{u_n}^{v_n} \frac{1}{b_n} \phi(e_n(x)) dx \right| = |v_n - u_n| o(1/b_n)$$

This means that if  $|v_n - u_n|/b_n$  is bounded then difference between the area from  $u_n$  to  $v_n$  under the density  $g_n$  and under the normal density with mean  $a_n$  and standard deviation  $b_n$  tends to zero. This means we can approximate the area under the density  $g_n$  by the area under the normal over intervals of length  $b_n$ .

Alternatively, define,

$$\Phi(t) = \int_{-\infty}^t \phi(z) dz.$$

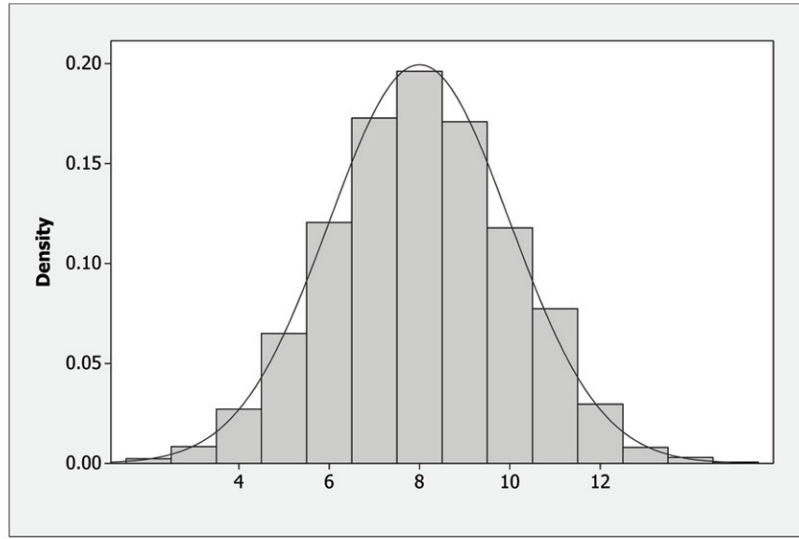


Figure 1: Normal approximation of a Binomial.

By a change of variables,

$$\int_{u_n}^{v_n} \frac{1}{b_n} \phi(e_n(x)) dx = \int_{e_n(u_n)}^{e_n(v_n)} \phi(z) dz = \Phi(e_n(v_n)) - \Phi(e_n(u_n))$$

Hence,

$$\left| \int_{u_n}^{v_n} g_n(x) dx - (\Phi(e_n(v_n)) - \Phi(e_n(u_n))) \right| = |v_n - u_n| o(1/b_n).$$

This approximation amounts to rescaling the density  $g_n$  into standard coordinates so  $a_n$  is rescaled to zero and  $u_n$  and  $v_n$  are rescaled to  $e_n(u_n)$  and  $e_n(v_n)$  respectively. Hence,

$$\begin{aligned} \int_{e_n(u_n)}^{e_n(v_n)} \frac{1}{b_n} g_n(e_n(x)) dx &= \int_{u_n}^{v_n} g_n(x) dx \\ &\approx \int_{u_n}^{v_n} \frac{1}{b_n} \phi(e_n(x)) dx \\ &= \int_{e_n(u_n)}^{e_n(v_n)} \phi(z) dz. \end{aligned}$$

Areas under this rescaled density are approximately equal to those under a superimposed standard normal density. This approximation is called the central limit theorem.

In Histogram 1 we approximate a Binomial density  $g_n$  with  $n = 16$  trials and  $p = 1/2$  by a Normal distribution with mean  $a_n = 8$  and standard deviation  $b_n = 2$ . The standardized values of  $e_n(6) = -1$ ,  $e_n(8) = 0$  and  $e_n(10) = 1$  are indicated. In standard units the area under the Binomial histogram may be approximated by the area under a Standard Normal.

Note that for integer valued densities  $g_n$  such as this one it is better to make a continuity correction. For instance if we seek the probability the Binomial is between 8 and 10 inclusively we really want the area in the boxes above 8, 9 and 10 and this is best approximated by the area under the Normal between 7.5 to 10.5. In standard units this means the area under the Standard Normal from 0.25 to 1.25. This continuity correction assumes the *density*  $g_n(x)$  is well approximated by  $g_n(e_n(x))/b_n$ ; i.e. it assumes a local limit theorem is valid.

The aim of this paper is to retrace the steps by Pascal, Bernoulli, de Moivre and Laplace leading first to the local limit theorem. We will see the local limit theorem was in some sense supplanted by the central limit theorem and essentially forgotten until its revival by Gnedenko forty years ago. We will also present a modern approach to local limit theorems which in some sense returns to the original ideas of the founders of probability theory. For simplicity we will restrict ourself to local limit theorems for integer valued random variables.

## 2 Historical antecedents

Notions of chance have existed since very ancient times. Julius Caesar was clearly using a popular metaphore when he uttered the famous words *Jacta alea est!* (let the dice fly) as he decided to cross the Rubicon with his army to overturn the Roman republic. Roman tesserae (dice) discovered in the ruins of Herculaneum are essentially the same as modern dice. The numbers one to six are engraved on the six sides of a die so that the sum of opposite faces is seven. We must assume that Romans had the intuitive notion that the chance of throwing a three is one in six.

Since Roman aleatores (gamblers) gambled with two dice, it would not be surprising that they knew intuitively that a throw of total of seven was more likely than a total of twelve. The question is whether they could have made a precise estimate. The probability of throwing a seven is equal to the number of ways of obtaining seven divided

by the total number of outcomes. The number of outcomes for the first die is six and for each of these outcomes there are six possible outcomes for the second die. Hence the total number of outcomes is 36. In other words the number of outcomes is the number of pairs  $(x, y)$  where  $x$  and  $y$  are chosen from one through six. To obtain a seven one has the the following six possibilities:  $(1,6)$ ,  $(6,1)$ ,  $(2,5)$ ,  $(5,2)$ ,  $(3,4)$ ,  $(4,3)$ . It follows that the probability of throwing a seven is  $1/6$ . Similarly the probability of throwing a twelve is  $1/36$ .

There is no evidence that Roman gamblers made this estimate. In fact it is unlikely because even the mathematical technology of fractions did not exist at that time. Nevertheless the idea that the calculation of probability reduces to counting is certainly ancient and is documented since the middle ages. For instance, Cardan (1576†) and Galileo (1642†) published counting arguments for games of chance [14]. However, the first systematic study of probability and counting is due to Pierre de Fermat and Blaise Pascal in an exchange of letters dating from 1654.

They were interested in the "problem of points". Suppose two players decide to play for a stake according to the rule the first to win 4 points takes all. A point could be decided by flipping heads or tails (or "navia aut capita" as an ancient Roman might say.) Now suppose the game has to be suspended when one player has 3 points while the other has only 1. How should the stake be split?

Fermat and Pascal found the solution by counting the possibilities. Pascal went further and calculated the fair split for a game where the player  $A$  needs  $m$  points to take all and player  $B$  needs  $k$  points. Let us follow the argument of James Bernoulli in Chapter Four of *Ars Conjectandi* published in 1713. The game will certainly end in  $n = k + m - 1$  tries. This means  $2^n$  possible sequences of winners; i.e. like  $(A, A, B, \dots)$  where  $A$  means player  $A$  wins and  $B$  means player  $B$  wins. Of these sequences player  $A$  will win if the sequence has zero  $B$ 's or one  $B$  or up to  $k - 1$   $B$ 's. The number of sequences with zero  $B$ 's is 1; i.e. all  $A$ 's. The number of sequences with one  $B$  is the number of ways of choosing one spot from  $n$  spots; i.e.  $n$ . The number of sequences with two  $B$ 's is the number of ways of choosing two spots from  $n$  spots; i.e.  $n(n - 1)/2$ . Continue in this way until we calculate the number of sequences with  $k - 1$   $B$ 's to be the number of ways of choosing  $k - 1$  spots from  $n$ ; i.e.  $n(n - 1) \dots (n - k + 2)/(k - 1)!$ . We conclude that the probability  $A$

will win is

$$\left(n + \frac{n(n-1)}{2} + \dots + \frac{n(n-1)\dots(n-k+2)}{(k-1)!}\right) \cdot 2^{-n}.$$

To complete his solution to the problem of points Pascal could rely on mathematical technology unavailable to the ancients. He could represent numbers as decimals. He could manipulate fractions and he had algebra at his disposition. Along the way he rediscovered the binomial coefficients; i.e. the number of ways of choosing  $m$  objects from  $n$  different objects is

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}.$$

This counting problem had already been solved by Chinese and Islamic scholars in medieval times. Pascal's triangle was called "Yang Hui's triangle" by the Chinese. Omar Khayyam (1131†) and Al-Karaji (1029†) had long since developed the binomial formula.

Pascal's contribution was the systematic application of counting methods to calculate probabilities. The construction of the abstract probability space with  $n+m-1$  games as above is thoroughly modern. It is abstract because a sequence with  $n$   $A$ 's in a row followed by  $m-1$   $B$ 's will never happen in practice because  $A$  wins after  $n$  games. This abstract space of sequences is a mathematical convenience and a huge intellectual step forward.

We can say Pascal was the first to consider a local limit theorem. If one tosses a coin  $n$  times then the probability of getting  $m$  heads is  $\binom{n}{m} 2^{-n}$ . This follows as above since there are  $2^n$  possible sequences and  $\binom{n}{m}$  of them lead to  $m$  heads. Let  $g_n$  be the density of the number of heads in  $n$  coin flips. Pascal's result shows  $g_n(m) = \binom{n}{m} 2^{-n}$ .

An additional step was taken by James Bernoulli (1705†). In *Ars Conjectandi* (published posthumously) he showed that if the probability of tossing a head is  $p$  then the density  $g_n$  of the number of heads in  $n$  tosses is

$$g_n(m) = \binom{n}{m} p^m (1-p)^{(n-m)}.$$

The first local limit theorem was proved by De Moivre (1754†) in *Approximatio ad Summam Terminorum Binomii (a + b)<sup>n</sup> in Seriem expansi* expanding on James Bernoulli’s work. For the case  $p = 1/2$  he proved;

**Theorem 2.1. (DeMoivre-Laplace local limit theorem)**

$$\lim_{n \rightarrow \infty} g_n(m) / \left( \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left(-\frac{(m - np)^2}{2np(1-p)}\right) \right) = 1$$

uniformly for  $m$  such that  $|\frac{m - np}{\sqrt{np(1-p)}}|$  remains bounded.

From Bernoulli’s work we know that the mean and standard deviation of the density  $g_n$  are  $a_n = np$  and  $b_n = \sqrt{np(1-p)}$  respectively. The above theorem can therefore be rewritten as

$$\lim_{n \rightarrow \infty} g_n(m) / \left( \frac{1}{b_n} \phi(e_n(m)) \right) = 1$$

uniformly for  $m$  such that  $|e_n(m)|$  remains bounded. Equivalently,

$$\lim_{n \rightarrow \infty} |g_n(m) - \frac{1}{b_n} \phi(e_n(m))| = o\left(\frac{1}{b_n} \phi(e_n(m))\right) = o\left(\frac{1}{b_n}\right)$$

uniformly for  $m$  such that  $|e_n(m)|$  remains bounded. This is of the form seen in the Introduction.

DeMoivre stated the local limit theorem for general  $p$  but the proof was provided by Laplace (1795) in *Théorie analytique des probabilités*. DeMoivre used the local limit theorem to add up the probabilities that  $S_n$  is in an interval of length of order  $\sqrt{n}$  to prove the first central limit theorem:

$$\sum_{m=np+a\sqrt{np(1-p)}}^{np+b\sqrt{np(1-p)}} g_n(m) \rightarrow \Phi(b) - \Phi(a).$$

DeMoivre proved only the case  $p = 1/2$  but Laplace extended these results to all  $p$ . For this reason we call the above theorem the DeMoivre-Laplace central limit theorem.

There was a gradual generalization of the scope of the central limit theorem to more general densities. The central limit theorem has been the object of research since that time and the local limit theorem languished. The Russian school (especially A. Khintchine)

used characteristic function methods to prove it thus avoiding the local behaviour of  $S_n$ . A. M. Lyapounov and then Jarl Waldemar Lindeberg (1932†) [6] extended the central limit theorem to independent but not necessarily random variables.

Consider a sequence of discrete densities  $f_k, k = 1, 2, \dots$  with support on the integers. Consider a sequence of independent random variables  $X_k, k = 1, 2, \dots$ , where  $X_k$  has density  $f_k$ ; that is  $P(X_k = m) = f_k(m)$ . Define  $S_n = X_1 + X_2 + \dots + X_n$  and let the density of  $S_n$  be  $g_n$ . Hence,

$$g_n(m) = f_1 * f_2 * \dots * f_n(m) \text{ where } * \text{ indicates convolution;}$$

that is

$$f * g(m) = \sum_{j=-\infty}^{\infty} f(j)g(m-j).$$

**Theorem 2.2. (Lindeberg)** *Suppose  $X_1, X_2, \dots$  are independent random variables where the mean  $\mu_m$  and the variance  $\sigma_m^2$  of  $X_m$  are both finite. Let the mean of  $S_n$  be  $a_n = \sum_{m=1}^n \mu_m$  and let the variance of  $S_n$  be  $b_n^2 = \sum_{m=1}^n \sigma_m^2$ . If Lindeberg's condition holds:*

$$\text{for all } \epsilon > 0, \lim_{n \rightarrow \infty} \frac{1}{b_n^2} \sum_{m=1}^n \sum_{|x| > \epsilon b_n} x^2 f_m(x) = 0$$

then uniformly in  $i, j$ ,

$$P(i \leq S_n \leq j) \rightarrow \Phi(e_n(j)) - \Phi(e_n(i)).$$

The local limit theorem was revived by B. V. Gnedenko [5] in 1948. This was a natural topic considering his illustrious teachers: Kolmogorov and Khintchine. To state the local limit theorem we will have to be a little bit careful because, if the summands  $X_1, X_2, \dots$  are all even valued there is no way the sum  $S_n$  can be odd so the local limit theorem will fail even when the central limit theorem holds.

**Definition 2.1.** *We say  $h$  is the maximal span of a density  $f$  if  $h$  is the largest integer such that the support of  $f$  is contained in the subgroup  $\{b + kh, k = \dots, -2, -1, 0, 1, \dots\}$  for some integer  $b$ .*

**Theorem 2.3. (Gnedenko)** *If  $X_1, X_2, \dots$  are independent random variables with identical density  $f$  with finite mean and variance and maximal span equal to 1 then*

$$|b_n P(S_n = k) - \phi(e_n(k))| \rightarrow 0$$

uniformly in  $k$  as  $n \rightarrow \infty$ .



The proof uses characteristic functions.

The extension of the local limit theorem to nonidentical random variables further complicates the problem of periodicity. Suppose  $f_m(0) = 1/2$ ,  $f_m(1) = 1/2^m$  and  $f_m(2) = 1/2 - 1/2^m$ . Then the support of this sequence of densities is essentially concentrated on the even integers so the maximal span is essentially 2 so Gnedenko's theorem can't hold. This problem was first addressed by N. G. Gamkrelidze [3] There are also many other generalizations [11, 12, 13, 9, 10]. A proof without using transform methods appeared in [7].

### 3 A local limit theorem without characteristic functions

Let  $X_1, X_2, \dots$  be independent, integer valued random variables  $X_m$  has density  $f_m$ . Let  $q(f_m) := \sum_k [f_m(k) \wedge f_m(k + 1)]$  and let  $Q_n := \sum_{m=1}^n q(f_m)$ .

**Theorem 3.1.** *Suppose there are numbers  $a_n$  and  $b_n$  such that*

$$\lim_{n \rightarrow \infty} P((S_n - a_n)/b_n \leq t) = \Phi(t), \text{ for } -\infty < t < \infty.$$

and such that

$$\limsup_{n \rightarrow \infty} b_n^2/Q_n < \infty.$$

Then the local limit theorem holds; that is as  $n \rightarrow \infty$ ,

$$\sup_k |b_n P(S_n = k) - \phi((k - a_n)/b_n)| \rightarrow 0.$$

The central limit theorem gives the probability that  $S_n$  falls in an interval of length of order  $b_n$ , say  $[k, k + \epsilon b_n]$ , as approximately

$$\int_k^{k+\epsilon b_n} g_n(x) dx \approx \int_k^{k+\epsilon b_n} \frac{1}{b_n} \phi(e_n(x)) dx. \tag{3.1}$$

Since  $e_n(x)$  is flat or constant on  $[k, k + \epsilon b_n]$  if  $\epsilon$  is small, it follows that the integral on the right hand side of (3.1) is approximately equal to the length of the interval times the value of  $g(e_n(x))$  at the left endpoint; i.e.  $\epsilon b_n \cdot \phi(e_n(k))/b_n$ . If we can show that  $g_n(x)$  is also fairly flat over  $[k, k + \epsilon b_n]$  for  $\epsilon$  small then

$$\int_k^{k+\epsilon b_n} g_n(x) dx \approx \epsilon b_n \cdot g_n(k) = \epsilon b_n \cdot P(S_n = k).$$

It would then follow that  $b_n \cdot P(S_n = k) \approx \phi(e_n(k))$ ; i.e. we would have a local limit theorem. We can formalize the idea of *flatness* by defining

$$s(X) = \sup_k |P(X = k + 1) - P(X = k)|.$$

**Lemma 3.1.** *If  $a_n$  and  $b_n$  are such that  $b_n \rightarrow \infty$  and  $\sup_n b_n^2 s(S_n) < \infty$  and  $(S_n - a_n)/b_n$  converges in distribution to the standard normal, then  $S_n$  satisfies the local limit theorem stated previously.*

The proof is given in [8]. Since Lindeberg's theorem gives sufficient conditions for convergence of  $(S_n - a_n)/b_n$  to the standard normal we see that the main step remaining to prove Theorem 3.1 is to give conditions ensuring that  $\sup_n b_n^2 s(S_n) < \infty$ . We do this by defining the Bernoulli part.

Suppose  $X$  has density  $f$ . Define  $\alpha(k) = f(k) \wedge f(k + 1)$ . Define

$$q \equiv q(f) = \sum_{k=-\infty}^{\infty} \alpha(k) = \sum_{k=-\infty}^{\infty} f(k) \wedge f(k + 1).$$

$q$  is called the amount of Bernoulli part inside of  $X$ . Define the density

$$f_T(k) = \frac{\alpha(k)}{q} = \frac{f(k) \wedge f(k + 1)}{q}.$$

Define the density

$$f_U(k) = \frac{1}{1 - q} \left( f(k) - \frac{(\alpha(k - 1) + \alpha(k))}{2} \right).$$

Define  $f_\epsilon(0) = 1 - q$  and  $f_\epsilon(1) = q$ . Define  $f_L(0) = 1/2$  and  $f_L(1) = 1/2$ .

Next we build the Bernoulli Decomposition: Construct a product space

$$\text{Integers} \times \text{Integers} \times \{0, 1\} \times \{0, 1\}.$$

with a product measure using the densities  $(f_T, f_U, f_\epsilon, f_L)$ . The projection maps  $(T, U, \epsilon, L)$  are independent random variables with densities  $(f_T, f_U, f_\epsilon, f_L)$ . Now check that

$$P((1 - \epsilon)U + \epsilon(T + L) = k) = f(k).$$

Define  $V = (1 - \epsilon)U + \epsilon T$  so  $X$  has the same density as  $V + \epsilon L$ .

We now arrive at the crux of the matter: Construct the Bernoulli Decomposition for each  $X_m$ :  $V_m + \epsilon_m L_m$ . Therefore  $S_n$  has the same

law as  $\sum_{m=1}^n [V_m + \epsilon_m L_m]$ . Define  $M_n = \sum_{m=1}^n \epsilon_m$ . Therefore  $S_n$  has the same law as

$$Z_n + \sum_{m=1}^{M_n} L_m^*$$

$Z_n = \sum_{m=1}^n V_m$  and  $L_1^*, L_2^*, \dots$  is another independent Bernoulli sequence.

We have found  $M_n$  independent Bernoulli random variables inside  $S_n$ . We know a lot about the Binomial distribution with  $p = 1/2$ . We can use this embedded Binomial to prove flatness of  $g_n$ .

**Lemma 3.2.** *Let*

$$b(k, n) = \binom{n}{k} \left(\frac{1}{2}\right)^n.$$

*Then*

$$|b(k + 1, n) - b(k, n)| \leq 32/n.$$

The proof is given in [8].

Now we can calculate

$$\begin{aligned} s(S_n) &= \sup_k |P(S_n = k) - P(S_n = k + 1)| \\ &= \sup_k |P(Z_n + \sum_{m=1}^{M_n} L_m^* = k) - P(Z_n + \sum_{m=1}^{M_n} L_m^* = k + 1)| \\ &\leq \sup_k \left( \sum_{m=0}^n P(M_n = m) |P(Z_n + \sum_{j=1}^m L_j^* = k | M_n = m) \right. \\ &\quad \left. - P(Z_n + \sum_{j=1}^m L_j^* = k + 1 | M_n = m) \right) \\ &\leq \sum_{m=0}^n P(M_n = m) \frac{64}{m + 1} \\ &= 64E \left[ \frac{1}{M_n + 1} \right]. \end{aligned}$$

since  $32/n \leq 64/(n + 1)$ .

Now,  $EM_n = Q_n$  so it isn't hard to show the above is of order  $1/Q_n$ . Hence  $Q_n \cdot s(S_n)$  is bounded so if we assume  $\limsup_{n \rightarrow \infty} b_n^2/Q_n < \infty$  as in Theorem 3.1 it follows that  $\limsup_{n \rightarrow \infty} b_n^2 \cdot s(S_n) < \infty$  and by Lemma 3.1 this proves Theorem 3.1.

## 4 Summary and future work

We have gone back to the beginning in a sense using the Bernoulli Part and the flatness of the Binomial to derive the local limit theorem from the CLT. There are other recent results on local limit theorems [4, 10] but to date there is no satisfactory theory when  $S_n$  is the sum of dependent random variables in spite of an attempt like [2]. Moreover proving flatness over a distance  $b_n$  isn't of much use for heavy tailed distributions but there are some results in this direction [1].

It is interesting to consider a Markov chain on the integers with a transition kernel  $K_{ij}$  defined by

$$K_{ij} = \begin{array}{c|c|c|c} j & i-1 & i & i+1 \\ \hline & 1/3 & 1/3 & 1/3 \end{array} \quad \text{if } i \text{ is even}$$

and

$$K_{ij} = \begin{array}{c|c|c|c} j & i-1 & i & i+1 \\ \hline & 3/8 & 1/4 & 3/8 \end{array} \quad \text{if } i \text{ is odd.}$$

Starting from 0, the state  $S_n$  at time  $n$  clearly satisfies the central limit theorem with mean zero and standard deviation  $b_n$ . Define the function  $\alpha(i) = 18/17$  if  $i$  is even and  $\alpha(i) = 16/17$  if  $i$  is odd. It is easy to see  $\alpha$  is an invariant measure for  $K$ ; i.e.  $\alpha K = \alpha$ . Moreover the average mass per integer is one. The ratio limit theorem implies that  $P(S_n = i)/\alpha(i)$  is fairly flat. Roughly speaking the walk spends a fraction  $9/17$  on the even integers and  $8/17$  on the odd integers. Consequently it is not unreasonable that there exist a local limit theorem

$$|b_n P(S_n = k) - \alpha(k) \phi(e_n(k))| \rightarrow 0$$

uniformly in  $k$  as  $n \rightarrow \infty$ .

To date there don't exist any published local limit theorems with nonhomogeneous local character as above. This is clearly another avenue for future work.

## Acknowledgements

David McDonald would like to thank Professors Nasser Arghami and Ahmad Parsian and Foroozan KheradPazhuh for their kind hospitality during his visit to the Mathematics House in Isfahan and to

the Mathematics and Statistics Department at Ferdowsi University Mashad. Talks given there formed the basis of this paper. Finally I would like to thank Professor Mohamed Mamaghani for encouraging me to write this historical review of the local limit theorem.

## References

- [1] Asmussen, S., Kalashnikov, V., Konstantinides, D., Klüppelberg C., and Tsitsiashvili, G. (2002), A local limit theorem for random walk maxima with heavy tails. *Statist. Probab. Letters*.
- [2] Dabrowski, A. and McDonald, D. (1996), An application of the Bernoulli part to local limit theorems for stationary sequences. *Canadian Journal of Statistics*, **24**(3), 293-305.
- [3] Gamkrelidze, N. G. (1964), On the local limit theorem for lattice random variables. *Theory of Prob. Appl.*, **9**, 662-665.
- [4] Gamkrelidze, N. G. (1988), On the application of a smoothness function in proving a local limit theorem. *Theory Prob. Applic.*, **33**, 352-354.
- [5] Gnedenko, B. V. (1948), On the local limit theorem in the theory of probability. *Uspekhi Mat. Nauk*, **3**, 187-194.
- [6] Lindeberg, J. W. (1922), Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.*, **15**, 211-225.
- [7] McDonald, D. (1979), On local limit theorems for integer valued random variables. *Theory of Prob. and Its Appl.*, **24**, 613-619.
- [8] Davis, B. and McDonald, D. (1995), An Elementary Proof of the Local Central Limit Theorem. *J. of Theoretical Probability*, **8**(3).
- [9] Mukhin, A. (1984), Local limit theorems for distributions of sums of independent random vectors. *Theory Probab Appl.*, **29**, 369-375.
- [10] Mukhin, A. B. (1991), Local limit theorems for lattice random variables. *Theory Prob. Appl.*, **35**, 698-713.

- [11] Prokhorov, Yu. (1954), On a local limit theorem for lattice distributions. Dokl. Akad. Nauk SSSR (N.S.), 535-538 (In Russian).
- [12] Rozanov, Yu. (1957), On a local limit theorem for lattice distributions. Theory Probab. Appl., **2** 260-265.
- [13] Statulyavichus, V. (1965), Limit theorems for densities and asymptotic expansions for distributions of sums of independent random variables. Theory Probab. Appl., **10**, 582-595.
- [14] Todhunter, I. (1965), History of the Theory of Probability. New York: Chelsea Publishing Co., 624 p.p.