Topics in
Probability Theory and Stochastic Processes
Steven R. Dunbar

Local Limit Theorems

Rating
Mathematicians Only: prolonged scenes of intense rigor.
Section Starter Question

Consider a binomial probability value for a large value of the binomial parameter $n$. How would you approximate this probability value? Would you expect this approximation to be uniformly good for all values of $k$? Explain why or why not?

---

Key Concepts

1. This section gives two estimates of binomial probability values for a large value of the binomial parameter $n$ in terms of the normal density. Each estimate is uniform in $k$. These estimates are two forms of the *Local Limit Theorem*.

---

Vocabulary

1. A **local limit theorem** describes how the probability mass function of a sum of independent discrete random variables at a particular value approaches the normal density.
Mathematical Ideas

Introduction

A local limit theorem describes how the probability mass function of a sum of independent discrete random variables approaches the normal density. We observe that the histogram of a sum of independent Bernoulli random variables resembles the normal density. From the Central Limit Theorem, we see that in standard units the area under the one bar of a binomial histogram may be approximated by the area under a standard normal. Theorems which compare the probability of a discrete random variable in terms of the area under a bar of a histogram to the area under a normal density are often called local limit theorems. This is illustrated in Figure 1.

In a way, Pascal laid the foundation for the local limit theorem when he formulated the binomial probability distribution for a Bernoulli random variable with $p = 1/2 = q$. James Bernoulli generalized the distribution to the case where $p \neq 1/2$. De Moivre proved the first real local limit theorem for the case $p = 1/2 = q$ in essentially the form of Lemma 9 in The de Moivre-Laplace Central Limit Theorem. Laplace provided a correct proof for the case with $p \neq 1/2$. De Moivre then 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 Central Limit Theorem. See Lemma 10 and following in The de Moivre-Laplace Central Limit Theorem. Khintchine, Lyapunov, and Lindeberg proved much more general versions of the Central Limit Theorem using characteristic functions and Fourier transform methods. Historically, the original Local Limit Theorem was overshadowed by the Central Limit Theorem, and forgotten until its revival by Gnedenko in 1948, [1].

First Form of Local Limit Theorem

Recall that $X_k$ is a Bernoulli random variable taking on the value 1 or 0 with probability $p$ or $1 - p$ respectively. Then

$$S_n = \sum_{k=1}^{n} X_i$$

is a binomial random variable indicating the number of successes in a composite experiment.
Figure 1: Comparison of the binomial distribution with $n = 20$, $p = 6/10$ with the normal distribution with mean $np$ and variance $np(1 - p)$. 

Figure 1
In this section, we study the size of \( P_n[S_n = k] \) as \( n \) approaches infinity. We will give an estimate uniform in \( k \) that is a form of the local limit theorem.

**Theorem 1** (Local Limit Theorem, First Form).

\[
P_n[S_n = k] = \frac{1}{\sqrt{2\pi p(1 - p)n}} \exp \left( \frac{-(k - np)^2}{2np(1 - p)} \right) + o(1).
\]

uniformly for all \( k \in \mathbb{Z} \).

**Remark.** Note that this is stronger than Lemma 9, the de Moivre-Laplace Binomial Point Mass Limit in the section on the de Moivre-Laplace Central Limit Theorem. It is stronger in that here the estimate is uniform for all \( k \in \mathbb{Z} \) instead of just an interval of order \( \sqrt{n} \) around the mean.

**Proof.**

1. Take \( t = \frac{7}{12} \). Note that \( \frac{1}{2} < t < \frac{2}{3} \).

(a) Proposition 1 (the Optimization Extension of the Central Limit Theorem) in **The Moderate Deviations Result** says

\[
P_n[S_n = k] = \frac{1}{\sqrt{2\pi p(1 - p)n}} \exp \left( \frac{-(k - np)^2}{2np(1 - p)} \right) (1 + \delta_n(k)),
\]

where

\[
\lim_{n \to \infty} \max_{|k - np| < n^{7/12}} |\delta_n(k)| = 0.
\]

(b) Rewriting this, we have

\[
P_n[S_n = k] = \frac{1}{\sqrt{2\pi p(1 - p)n}} \exp \left( \frac{-(k - np)^2}{2np(1 - p)} \right) \\
+ \frac{1}{\sqrt{2\pi p(1 - p)}} \exp \left( \frac{-(k - np)^2}{2np(1 - p)} \right) \frac{\delta_n(k)}{\sqrt{n}},
\]

where the last term is \( o_u(n^{-1/2}) \) uniformly in \( k \) when \( |k - np| < n^{7/12} \). Since

\[
\frac{1}{\sqrt{2\pi p(1 - p)}} \exp \left( \frac{-(k - np)^2}{2np(1 - p)} \right) \leq 1
\]

we can simplify this to

\[
P_n[S_n = k] = \frac{1}{\sqrt{2\pi p(1 - p)n}} \exp \left( \frac{-(k - np)^2}{2np(1 - p)} \right) + \frac{\delta_n(k)}{\sqrt{n}}.
\]
2. Theorem 8.1 (the Moderate Deviations Theorem) says that if we take $a_n = \frac{n^{1/12}}{\sqrt{p(1-p)}}$, then $\lim_{n \to \infty} a_n = +\infty$ and $\lim_{n \to \infty} a_n n^{-1/6} = 0$. Thus,

$$\mathbb{P}_n \left[ S_n \geq np + n^{7/12} \right] = \mathbb{P}_n \left[ \frac{S_n}{n} - p \geq n^{-5/12} \right]$$

$$= \mathbb{P}_n \left[ \frac{S_n}{n} - p \geq \sqrt{p(1-p)} \cdot \frac{n^{1/12}}{\sqrt{p(1-p)}} \cdot \frac{1}{n^{1/2}} \right]$$

$$\sim \frac{\sqrt{p(1-p)}}{n^{1/2} \sqrt{2\pi}} \cdot \exp \left( \frac{-n^{1/6}}{2p(1-p)} \right)$$

where the last asymptotic limit comes from the Moderate Deviations Theorem on $|k - np| < n^{7/12}$.

3. The asymptotic limit implies a weaker estimate:

$$\mathbb{P}_n \left[ S_n \geq np + n^{7/12} \right] = o \left( n^{-1/2} \right).$$

To see this, note that

$$\mathbb{P}_n \left[ S_n \geq np + n^{7/12} \right] = \mathbb{P}_n \left[ S_n \geq np + n^{7/12} \right] \cdot \left( \frac{\sqrt{p(1-p)}}{n^{1/2} \sqrt{2\pi}} \cdot \exp \left( \frac{-n^{1/6}}{2p(1-p)} \right) \right).$$

Note that the first factor goes to 1 by step 2. So consider $n^{5/12} \exp \left( \frac{-n^{1/6}}{2p(1-p)} \right) \to 0$. This estimate is uniform in $k$ for $k - np \geq n^{7/12}$.

4. Note that

$$\exp \left( \frac{-(k - np)^2}{2np(1-p)} \right) \geq \exp \left( \frac{-n^{14/12}}{2np(1-p)} \right) = \exp \left( \frac{-n^{1/6}}{2p(1-p)} \right)$$

so $\exp \left( \frac{-(k - np)^2}{2np(1-p)} \right) = o(1)$.

5. Using step 4, we see that

$$\frac{1}{\sqrt{2\pi p(1-p)n}} \exp \left( \frac{-(k - np)^2}{2np(1-p)} \right) = o(n^{-1/2})$$

uniformly in $k$ with $k - np \geq n^{7/12}$. 
6. Make step 3 look the same as step 1b by putting step 5 into step 3 without disturbing the estimate:

\[ P_n [ S_n = k ] = \frac{1}{\sqrt{2\pi p(1-p)n}} \exp \left( \frac{- (k - np)^2}{2np(1-p)} \right) + o(n^{-1/2}). \]

Finally, factor out \( \frac{1}{\sqrt{2\pi p(1-p)n}} \) and the proof is finished.

\[ \square \]

**The Second Form of the Local Limit Theorem**

Recall that \( Y_i \) is a sequence of independent random variables which take values 1 with probability \( \frac{1}{2} \) and \(-1\) with probability \( \frac{1}{2} \). This is a mathematical model of a fair coin flip game where a 1 results from “heads” on the \( i \)th coin toss and a \(-1\) results from “tails”. Let \( H_n \) and \( L_n \) be the number of heads and tails respectively in \( n \) flips. Then \( T_n = \sum_{i=1}^{n} Y_i = H_n - L_n = 2S_n - n \) counts the difference between the number of heads and tails, an excess of heads if positive. The second form of the local limit theorem is useful for estimating the probability that \( T_n \) takes a value close to its average \((2p-1)n\).

**Remark.** The following version of Stirling’s Formula follows from the statement of the First Form of the Local Limit Theorem. However, note that the proof of the Local Limit Theorem uses the Moderate Deviations Theorem. The proof of the Moderate Deviations Theorem uses the Optimization Extension of the de Moivre-Laplace Central Limit Theorem. Step 1 of the proof of the Optimization Extension uses Stirling’s Formula. So this is not a new proof of Stirling’s Formula for binomial coefficients. It is a long way to derive the asymptotics for binomial coefficients from the usual Stirling Formula.

**Corollary 1.**

\[ \binom{n}{k} = \sqrt{\frac{2}{\pi}} \frac{2^n}{\sqrt{n}} \exp \left( - \frac{2}{n} (k - n/2)^2 + o(1) \right) \]

uniformly for \( k \in \mathbb{Z} \).

**Proof.** Use \( p = 1/2 \) in the First Form of the Local Limit Theorem. \[ \square \]

**Theorem 2** *(Local Limit Theorem, Second Form).* Set \( T_n = 2S_n - n \), then

\[ P_n [ T_n = k ] = \sqrt{\frac{2}{\pi}} \left( \frac{p}{1-p} \right)^{k/2} \left( \frac{2\sqrt{p(1-p)}}{\sqrt{n}} \right)^n \left( \exp \left( \frac{-k^2}{2n} \right) + o(1) \right) \]

7
uniformly for $k \in \mathbb{Z}$ such that $n + k$ is even.

Proof.

$$
P_n [T_n = k] = P_n [2S_n - n = k] = P_n \left[ S_n = \frac{n + k}{2} \right] = \left( \frac{n}{n+k} \right) p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}}
$$

Using the statement of the Corollary

$$
= \sqrt{\frac{2}{\pi}} p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}} \frac{2^n}{\sqrt{n}} \left( \exp \left( -\frac{2}{n} \left( \frac{n}{2} + \frac{k}{2} - \frac{n}{2} \right)^2 \right) + o(1) \right)
$$

$$
= \sqrt{\frac{2}{\pi}} \left( \frac{p}{1-p} \right)^{k/2} \left( \frac{2\sqrt{p(1-p)}}{\sqrt{n}} \right)^n \left( \exp \left( -\frac{k^2}{2n} \right) + o(1) \right).
$$

Remark. Note that the First Form and the Second Form of the Local Limit Theorem say the same thing in the symmetric case $p = 1/2$.

**Corollary 2.** Let $K$ be a fixed finite subset of $\mathbb{Z}$ that contains at least one even number and one odd number. Then

$$
\lim_{n \to \infty} \frac{1}{n} \ln \left( P_n [T_n \in K] \right) = \ln \left( 2\sqrt{p(1-p)} \right).
$$

Proof. The Second Form of the Local Limit Theorem says that the probability that $T_n$ is in a fixed finite subset of $\mathbb{Z}$ decreases exponentially as $n$ approaches infinity. 

**Discussion and Extensions**

As the Second Form of the Local Limit Theorem indicates, to state a Local Limit Theorem for sums of random variables more general than Bernoulli random variables will take some care. For example, if the summands $X_1, X_2, \ldots$ are all even valued, there is no way the sum $X_1 + \cdots + X_n$ can be odd so a Local Limit Theorem will require at least an additional hypothesis while the Central Limit Theorem will still hold.
Definition. We say that $h$ is the maximal span of a density $f$ if $h$ is the largest integer such that the support of $f$ is contained in the affine subset \( \{b + kh : k = \ldots, -2, -1, 0, 1, 2, \ldots\} \) of $\mathbb{Z}$ for some $b$.

Let $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ be the standard normal density.

**Theorem 3** (Gnedenko’s Local Limit Theorem). Let $X_1, X_2, \ldots$ be independent random variables with identical density $f$ with finite mean $\mu$ and finite variance $\sigma^2$ and maximal span equal to 1. Let $S_n = X_1 + \cdots + X_n$. Then

\[
(n\sigma^2) \left| \mathbb{P}[S_n = k] - \frac{1}{(n\sigma^2)} \phi((x - n\mu)/(n\sigma^2)) \right| \to 0
\]

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

**Proof.** This was proved by B. V. Gnedenko in 1948, [1], using characteristic functions. \qed

McDonald, [3], has more technical local limit theorems and further references to other generalizations.

The following example shows that any extension of the Local Limit Theorem to nonidentical random variables illustrates is complicated by the same problem of periodicity that already appears in the Second Form of the Local Limit Theorem above.

**Example.** Suppose the densities $f_m$ are $f_m(0) = 1/2$, $f_m(1) = 1/2^m$, $f_m(2) = 1/2 - 1/2^m$. The first random variable will be either 0 or 1, each with probability 1/2. For $m > 1$, the random variables will “essentially” differ by 2. That is, the maximal span is “essentially” 2, but Gnedenko’s theorem will not hold. In fact, consider the histogram of the convolution of the distributions $f_1, \ldots f_{20}$ in Figure 2. This histogram is the distribution of the sum of the first 20 of the random variables.

Note that the distribution overall resembles the normal distribution, as expected from the Lindeberg Central Limit Theorem. However closer inspection show that the value of the distribution at symmetric distances around the mean 20 differ. That is the value of the distribution at 21 is greater than the value of the distribution at 19. This suggests that at least the uniform approach to the normal distribution at integer values will not hold.
Figure 2: Distribution of the sum of the first 20 of the random variables with nonidentical independent distributions \((1/2, 1/2^m, 1/2 - 1/2^m)\).

Sources

This section is adapted from: *Heads or Tails*, by Emmanuel Lesigne, Student Mathematical Library Volume 28, American Mathematical Society, Providence, 2005, Chapter 9, [2]. The historical remarks and the generalizations of the Local Limit Theorem are adapted from [3].

Algorithms, Scripts, Simulations

Algorithm

The experiment is flipping a coin \(n\) times, and repeat the experiment \(k\) times. Then check the probability of a specific value and compare to the
normal probability density. Also compare the logarithmic rate of growth to the predicted rate.

Scripts

R

R script for Local Limit Theorem.

```r
p <- 0.5
n <- 10000
k <- 1000
upperintvalue <- 10
lowerintvalue <- 0
winLose <- 2 * (array( 0+(runif(n*k) <= p), dim=c(n,k))) - 1
# 0+ coerces Boolean to numeric
totals <- colSums(winLose)
# -n..2..n binomial rv sample, size k

stddev <- sqrt(p*(1-p)*n)
prob <- sum( 0+(totals == upperintvalue) )/k

f1 <- (sqrt(2/pi))
f2 <- (p/(1-p))^(upperintvalue/2)
f3 <- ((2*sqrt(p*(1-p)))^n)/sqrt(n)
f4 <- exp( -(upperintvalue^2/(2*n)) )
theoretical <- f1*f2*f3*f4
cat(sprintf("Empirical probability:\n", prob ))
cat(sprintf("Local Limit Theorem \ estimate:\n", theoretical))

set <- (totals >= lowerintvalue) & (totals <= upperintvalue)
logintervalprobrate <- (1/n)*log( sum( 0+(set))/k )
theoreticallograte <- log( 2*sqrt(p*(1-p)) )
cat(sprintf("Empirical log rate:\n", logintervalprobrate ))
cat(sprintf("Theoretical log rate:\n", theoreticallograte))
```

Octave

Octave script for Local Limit Theorem.

```octave
p = 0.5;
```
n = 10000;
k = 1000;

upperintvalue = 10;
lowerintvalue = 0;

winLose = 2 * (rand(n,k) <= p) - 1;
headsTotal = sum(winLose);
# 0..n binomial rv sample, size k

stddev = sqrt(p*(1-p)*n);
prob = sum(headsTotal == upperintvalue)/k;

f1 = (sqrt(2/pi));
f2 = (p/(1-p))^(upperintvalue/2);
f3 = ((2*sqrt(p*(1-p)))^n)/sqrt(n);
f4 = exp(-(upperintvalue^2/(2*n)));
theoretical = f1*f2*f3*f4;

disp("Empirical_probability:")
disp( prob )
disp("Local_Limit_Theorem_estimate:")
disp( theoretical )

set = (headsTotal >= lowerintvalue) & (headsTotal <= upperintvalue)
logintervalprobrate = (1/n)*log( sum(set)/k );
theoreticallograte = log( 2*sqrt(p*(1-p)) );

disp("Empirical_log_rate:")
disp( logintervalprobrate )
disp("Theoretical_log_rate:")
disp( theoreticallograte )

Perl PDL script for Local Limit Theorem.

use PDL::NiceSlice;
use PDL::Constants qw(PI);

$p = 0.5;
$n = 10000;
$k = 1000;
$upperintvalue = 10;
$lowerintvalue = 0;

#note order of dims!!
$winLose = 2 * ( random( $k, $n ) <= $p ) - 1;

# -n..2..n binomial r.v. sample, size k
#note transpose, PDL likes x (row) direction for implicitly threaded operation
$totals = $winLose->transpose->sumover;

$stddev = sqrt( $p * ( 1 - $p ) * $n );

$prob = ( ( $totals == $upperintvalue )->sumover ) / $k;

$f1 = ( sqrt( 2 / PI ) );
$f2 = ( $p / ( 1 - $p ) )**( $upperintvalue / 2 );
$f3 = ( ( 2 * sqrt( $p * ( 1 - $p )) )**$n ) / sqrt($n);
$f4 = exp( -( $upperintvalue**2 / ( 2 * $n ) ) );

$theoretical = $f1 * $f2 * $f3 * $f4;
print "Empirical probability: ", $prob, "\n";
print "Local Limit Theorem estimate: ", $theoretical, "\n";

$set = ( $totals >= $lowerintvalue ) & ( $totals <= $upperintvalue
$logintervalprobrate = ( 1 / $n ) * log( sum($set) / $k );
$theoreticalallograte = log( 2 * sqrt( $p * ( 1 - $p ) ) );
print "Empirical log probability rate:", $logintervalprobrate, "\n"
print "Theoretical log rate: ", $theoreticalallograte, "\n";

SciPy Scientific Python script for Local Limit Theorem.

import scipy

p = 0.5
n = 10000
k = 1000
upperintvalue = 10
lowerintvalue = 0

winLose = 2 * (scipy.random.random((n,k))<= p) - 1
# Note Booleans True for Heads and False for Tails
totals = scipy.sum(winLose, axis = 0)
# -n..2..n binomial r.v. sample, size k
# Note how Booleans act as 0 (False) and 1 (True)

stddev = scipy.sqrt( p * (1 - p) * n )

prob = (scipy.sum( totals == upperintvalue )).astype('float')/k
# Note the casting of integer type to float to get float

f1 = ( scipy.sqrt( 2 / scipy.pi ) );
f2 = ( p / ( 1 - p ) )** ( upperintvalue / 2 );
f3 = ( ( 2 * scipy.sqrt( p * ( 1 - p ) ) )**n ) / scipy.sqrt( n )
f4 = scipy.exp( -( upperintvalue**2 / ( 2.0 * n ) ) );
# Note the use of 2.0 to force integer arith to floating point
theoretical = f1 * f2 * f3 * f4;

print "Empirical probability: ", prob
print "Moderate Deviations Theorem estimate: ", theoretical

set = ( totals >= lowerintvalue ) & ( totals <= upperintvalue );
logintervalprobrate = ( 1 / n ) * scipy.log((scipy.sum(set))).astype('float')
theoreticallograte = scipy.log( 2 * scipy.sqrt( p * ( 1 - p ) ) );
print "Empirical log probability rate: ", logintervalprobrate
print "Theoretical log rate: ", theoreticallograte
Problems to Work for Understanding

1. Modify the value of $p$ in a script, say to $p = 0.51$, and investigate the convergence and the rate of convergence to the normal density function.

2. Modify the value of $n$ in a script to a sequence of values and investigate the rate of convergence of the logarithmic rate.

---

Reading Suggestion:

References


---

Outside Readings and Links:


3. Encyclopedia of Mathematics, Local Limit Theorems

---

I check all the information on each page for correctness and typographical errors. Nevertheless, some errors may occur and I would be grateful if you would alert me to such errors. I make every reasonable effort to present current and accurate information for public use, however I do not guarantee the accuracy or timeliness of information on this website. Your use of the information from this website is strictly voluntary and at your risk.

I have checked the links to external sites for usefulness. Links to external websites are provided as a convenience. I do not endorse, control, monitor, or guarantee the information contained in any external website. I don’t guarantee that the links are active at all times. Use the links here with the same caution as you would all information on the Internet. This website reflects the thoughts, interests and opinions of its author. They do not explicitly represent official positions or policies of my employer.

Information on this website is subject to change without notice.

Steve Dunbar’s Home Page, http://www.math.unl.edu/~sdunbar1
Email to Steve Dunbar, sdunbar1 at unl dot edu
Last modified: Processed from \TeX source on April 26, 2015