Steven R. Dunbar
Department of Mathematics
203 Avery Hall
University of Nebraska-Lincoln
Lincoln, NE 68588-0130
http://www.math.unl.edu
Voice: 402-472-3731
Fax: 402-472-8466

Topics in Probability Theory and Stochastic Processes Steven R. Dunbar

---

The de Moivre-Laplace Central Limit Theorem

Rating

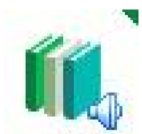
Mathematicians Only: prolonged scenes of intense rigor.

Section Starter Question

What is the most important probability distribution? Why do you choose that distribution as most important?

Key Concepts

1. The statement, proof and meaning of the de Moivre-Laplace Central Limit Theorem.



Vocabulary

1. The **standard Gaussian density** is
$$\phi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}.$$

2. The **complementary cumulative distribution function** $\Phi(y)$ of the Gaussian density is $\Phi^c(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty \mathrm{e}^{-x^2/2} \, \mathrm{d}x$. $\Phi^c(y)$ measures the area under the upper tail of the Gaussian density.

3. The **complementary error function** $\operatorname{erfc} x$ is defined by $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \mathrm{e}^{-x^2} \, \mathrm{d}x$.

4. The **de Moivre-Laplace Central Limit Theorem** is the statement that for $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a < b$, then
$$\lim_{n\to\infty} \mathbb{P}_n \left[a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b \right] = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x$$

Mathematical Ideas

History of the de Moivre-Laplace Central Limit Theorem

The first statement of what we now call the de Moivre-Laplace Central Limit Theorem occurs in *The Doctrine of Chances* by Abraham de Moivre in 1738. He proved the result for $p = 1/2$. This finding was far ahead of its time, and was nearly forgotten until the famous French mathematician Pierre-Simon Laplace rediscovered it. Laplace generalized the theorem to $p \neq 1/2$ in *Théorie Analytique des Probabilités* published in 1812. Gauss also contributed to the statement and proof of the general form of the theorem.

Laplace also discovered the more general form of the Central Limit Theorem but his proof was not rigorous. As with De Moivre, Laplace's finding received little attention in his own time. It was not until the end of nineteenth century that the generality of the central limit theorem was realized. The Russian mathematician Aleksandr Liapunov gave the first rigorous proof of the general Central Limit Theorem in 1901-1902. As a result a general version of the Central Limit Theorem is occasionally referred to as Liapunov's theorem. A theorem with weaker hypotheses but with equally strong conclusion is Lindeberg's Theorem of 1922. It says that the sequence of random variables need not be identically distributed, instead the random variables only need zero means with individual variances small compared to their sum.

George Pólya first used the name "Central Limit Theorem" (in German: "zentraler Grenzwertsatz") in 1920 in the title of a paper, [4].

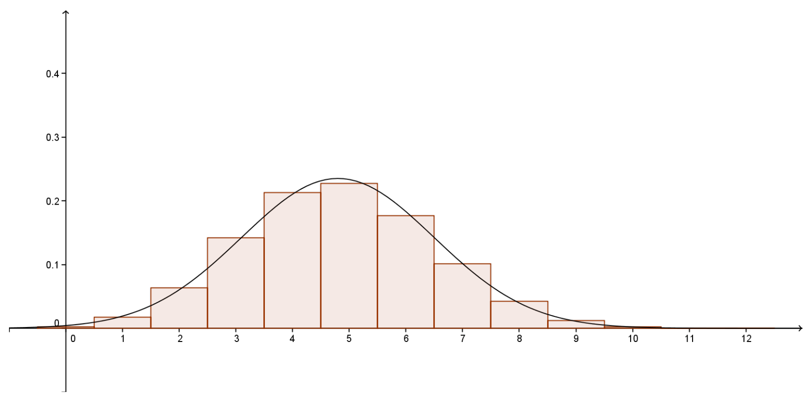


Figure 1: Comparison of the binomial distribution with $n = 12$, $p = 4/10$ with the normal distribution with mean np and variance $np(1-p)$.

Interpretation of the de Moivre-Laplace Central Limit Theorem

Theorem 1 (de Moivre-Laplace Central Limit Theorem)**.** *Let $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a < b$. Then*

$$\lim_{n\to\infty} \mathbb{P}_n \left[a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b \right] = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x$$

and the convergence is uniform in a and b.

If $y \in \mathbb{R}$, $n \in \mathbb{Z}$ and $0 < p < 1$ define

$$k(y) = \lfloor np + y\sqrt{np(1-p)} \rfloor.$$

Then

$$\lim_{n\to\infty} \sum_{j=0}^{k(y)} \binom{n}{j} p^j (1-p)^{n-j} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{np+y\sqrt{np(1-p)}} \mathrm{e}^{\frac{-(x-np)^2}{2np(1-p)}} \, \mathrm{d}x \, .$$

This is illustrated in Figure 1

Heuristic Proof of the de Moivre-Laplace Central Limit Theorem

We are interested in the natural random variation of S_n around its mean. From the Weak Law of Large Numbers, we know that $\mathbb{P}_n \left[|\frac{S_n}{n} - p| > \epsilon \right] \to 0$.

From the Large Deviations result we also know that $\mathbb{P}_n\left[|\frac{S_n}{n} - p| > \epsilon\right] \leq \mathrm{e}^{-nh_+(\epsilon)} + \mathrm{e}^{-nh_-(\epsilon)}$. Equivalently, we can say that S_n will fall outside the range $np(1 \pm \epsilon)$ with probability near 0. Finally, note that $\mathbb{E}\left[(S_n - np)^2\right] = np(1-p)$. We ask, "How large a fluctuation or deviation of S_n from np should be surprising?". We want a function $\psi(n)$ with

$$\lim_{n\to\infty} \mathbb{P}_n\left[|S_n - np| > \psi(n)\right] = \alpha, \text{ for } 0 < \alpha < 1. \tag{1}$$

To measure the surprise of a fluctuation, we specify α, then ask what is the order of $\psi(n)$ as a function of n? Small but fixed values of α would indicate large surprise, i.e. unlikely deviations, and so we expect $\psi(n)$ to grow but more slowly than ϵn.

Take $p = 1/2$ to simplify the calculations for the discovery oriented proof in this subsection. We can make some useful guesses about $\psi(n)$. Interpret the probability on the left in as the area in the histogram for the binomial distribution of S_n. From the expression of Wallis' Formula for the central term in the binomial distribution, the maximum height of the histogram bars is of the order $\frac{1}{\sqrt{n\pi}}$, see Wallis' Formula. That means that to get a fixed area α around that central term requires an interval of width at least a multiple of $\sqrt{n}$. If we take $\psi(n) = x_n\sqrt{n}/2$ (with the factor $1/2$ put in to make variances cancel nicely), then we are looking for a sequence x_n which will make

$$\lim_{n\to\infty} \mathbb{P}_n\left[|S_n - n/2| > x_n\sqrt{n}/2\right] = \alpha \text{as } n \to \infty$$

true for $0 < \alpha < 1$. By Chebyshev's Inequality, we can estimate this probability as

$$\mathbb{P}_n\left[|S_n/n - 1/2| > x_n/(2\sqrt{n})\right] \leq 1/x_n^2.$$

If $\limsup_{n\to\infty} x_n = \infty$, we could only obtain $\alpha = 0$, so x_n is bounded above. If $\liminf_{n\to\infty} x_n = 0$ then for a fixed $\epsilon > 0$ and some subsequence n_m such that for sufficiently large m

$$\mathbb{P}_{n_m}\left[|S_{n_m}/n_m - 1/2| > \epsilon > x_{n_m}/(2\sqrt{n_m})\right] \to 0.$$

which is also contradiction to the assumption $\alpha > 0$. Hence x_n is bounded below by a positive value. We guess that $x_n = x$ so $\psi(n) = x\sqrt{n}/2$ for all values of n.

Breiman's Proof of the de Moivre-Laplace Central Limit Theorem

To simplify the calculations, take the number of trials to be even and $p = 1/2$. Then the expression we want to evaluate and estimate is

$$\mathbb{P}_{2n}\left[|S_{2n} - n| < x\sqrt{2n}/2\right].$$

This is evaluated as

$$\sum_{|k-n|<x\sqrt{n/2}} 2^{-2n}\binom{2n}{k} = \sum_{|j|<x\sqrt{n/2}} 2^{-2n}\binom{2n}{n+j}.$$

Let $P_n = 2^{-2n}\binom{2n}{n}$ be the central binomial term and then write each binomial probability in terms of this central probability P_n, specifically

$$2^{-2n}\binom{2n}{n+j} = P_n \cdot \frac{n(n-1)\cdots(n-j+1)}{(n+j)\cdots(n+1)}.$$

Name the fractional factor above as $D_{j,n}$ and rewrite it as

$$D_{j,n} = \frac{1}{(1+j/n)(1+j/(n-1))\cdots(1+j/(n-j+1))}$$

and then

$$\log(D_{j,n}) = -\sum_{k=0}^{j-1}\log(1+j/(n-k)).$$

Now use the common two-term asymptotic expansion for the logarithm function $\log(1+x) = x(1+\epsilon_1(x))$. Note that

$$\epsilon_1(x) = \frac{\log(1+x)}{x} - 1 = \sum_{k=2}^{n}\frac{(-1)^{k+1}x^k}{k}$$

so $-x/2 < \epsilon_1(x) < 0$ and $\lim_{x\to 0}\epsilon_1(x) = 0$.

$$\log(D_{j,n}) = -\sum_{k=0}^{j-1}\frac{j}{n-k}\left(1+\epsilon_1\left(\frac{j}{n-k}\right)\right).$$

Let

$$\epsilon_{1,j,n} \sum_{k=0}^{j-1} \frac{j}{n-k} = \sum_{k=0}^{j-1} \frac{j}{n-k} \epsilon_1 \left(\frac{j}{n-k} \right).$$

Then we can write

$$\log(D_{j,n}) = -(1 + \epsilon_{1,j,n}) \sum_{k=0}^{j-1} \frac{j}{n-k}.$$

Note that j is restricted to the range $|j| < x\sqrt{n/2}$ so

$$\frac{j}{n-k} < \frac{x\sqrt{n/2}}{n - x\sqrt{n/2}} = \frac{x}{\sqrt{2n} - x}$$

and then

$$\epsilon_{1,j,n} = \max_{|j|<x\sqrt{n/2}} \epsilon_1 \left(\frac{j}{n-k} \right) \to 0 \text{ as } n \to \infty \ .$$

Write

$$\frac{j}{n-k} = \frac{j}{n} \cdot \frac{1}{1 - k/n}$$

and then expand $\frac{1}{1-k/n} = 1 + \epsilon_2(k/n)$ where $\epsilon_2(x) = 1/(1-x) - 1 = \sum_{k=1}^{\infty} x^k$ so $\epsilon_2(x) \to 0$ as $x \to 0$. Once again k is restricted to the range $|k| \le |j| < x\sqrt{n/2}$ so

$$\frac{k}{n} < \frac{x\sqrt{n/2}}{n} = \frac{x}{\sqrt{2n}}$$

so that

$$\epsilon_{2,j,n} = \max_{|k|<x\sqrt{n/2}} \epsilon_2 \left(\frac{k}{n} \right) \to 0 \text{ as } n \to \infty \ .$$

Then we can write

$$\log(D_{j,n}) = -(1 + \epsilon_{1,j,n})(1 + \epsilon_{2,j,n}) \sum_{k=0}^{j-1} \frac{j}{n}.$$

Simplify this to

$$\log(D_{j,n}) = -(1 + \epsilon_{3,j,n}) \sum_{k=0}^{j-1} \frac{j}{n} = -(1 + \epsilon_{3,j,n}) \frac{j^2}{n}$$

where $\epsilon_{3,j,n} = \epsilon_{1,j,n} + \epsilon_{2,j,n} + \epsilon_{1,j,n} \cdot \epsilon_{2,j,n}$. Therefore $\epsilon_{3,j,n} \to 0$ as $n \to \infty$ uniformly in j. Exponentiating

$$D_{j,n} = \mathrm{e}^{-j^2/n}(1 + \Delta_{j,n})$$

where $\Delta_{j,n} \to 0$ as $n \to 0$ uniformly in j.

Using Stirling's Formula,

$$P_n = 2^{-2n} \frac{(2n)!}{n! \cdot n!} = \frac{1}{\sqrt{n\pi}}(1 + \delta_n).$$

Summarizing

$$\begin{aligned}
\mathbb{P}_{2n}\left[|S_{2n} - n| < x\sqrt{2n}/2\right] &= \sum_{|j|<x\sqrt{n/2}} 2^{-2n}\binom{2n}{n+j} \\
&= \sum_{|j|<x\sqrt{n/2}} P_n \cdot D_{j,n} \\
&= \sum_{|j|<x\sqrt{n/2}} P_n \cdot \mathrm{e}^{-j^2/n}(1 + \Delta_{j,n}) \\
&= (1 + \delta_n) \sum_{|j|<x\sqrt{n/2}} \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-j^2/n} \sqrt{\frac{2}{n}}
\end{aligned}$$

Make the change of variables $t_j = j\sqrt{2/n}$, $\Delta t = t_{j+1} - t_j = \sqrt{2/n}$ so the summation is over the range $-x < t_j < x$. Then

$$\mathbb{P}_{2n}\left[|S_{2n} - n| < x\sqrt{2n}/2\right] = (1 + \delta_n) \sum_{-x<t_j<x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-t_j^2} 2\Delta t.$$

The factor on the right is the approximating sum for the integral of the standard normal density over the interval $[-x, x]$. Therefore,

$$\lim_{n\to\infty} \mathbb{P}_{2n}\left[|S_{2n} - n| < x\sqrt{2n}/2\right] = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} \mathrm{e}^{-t^2/2} \, \mathrm{d}t\,.$$

Formal Proof of the de Moivre-Laplace Theorem

Theorem 2 (de Moivre-Laplace Central Limit Theorem)**.** *Let* $a, b \in \mathbb{R} \cup \{\pm\infty\}$ *with* $a < b$. *Then*

$$\lim_{n\to\infty} \mathbb{P}_n \left[a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b \right] = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x$$

and the convergence is uniform in a *and* b.

For the proof of the de Moivre-Laplace Central Limit Theorem, we need several lemmas.

Lemma 3.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x = 1.$$

See the proofs in Gaussian Density.

Definition. Define the **complement of the cumulative distribution function** (ccdf) as $\Phi^c(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty \mathrm{e}^{-x^2/2} \, \mathrm{d}x$. The ccdf $\Phi^c(y)$ measures the area under the upper tail of the standard Gaussian density while $\Phi(y)$ is the cumulative density function of the standard Gaussian density.

Definition. Some mathematicians and engineers use alternative accumulation functions for a scaled Gaussian density called the **error function** $\operatorname{erf}(x)$ and the **complementary error function** $\operatorname{erfc} x$. Define $\operatorname{erfc}(y) := \frac{2}{\sqrt{\pi}} \int_y^\infty \mathrm{e}^{-x^2} \, \mathrm{d}x$ and $\operatorname{erf}(y) := 1 - \operatorname{erfc}(y)$. Note that $\Phi^c(y) = \frac{1}{2} \operatorname{erfc}(y/\sqrt{2})$.

Lemma 4.

$$\frac{1}{\sqrt{2\pi}(y+1)} (\mathrm{e}^{-y^2/2} - \mathrm{e}^{-(y+1)^2/2}) \le \Phi^c(y) \le \frac{1}{\sqrt{2\pi} y} \mathrm{e}^{-y^2/2}.$$

Proof. For the lower bound on $\Phi^c(y)$:

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\int_y^\infty \mathrm{e}^{-\frac{t^2}{2}}\ \mathrm{d}t &\geq \frac{1}{\sqrt{2\pi}}\int_y^{y+1} \mathrm{e}^{-\frac{t^2}{2}}\ \mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}}\int_y^{y+1} \frac{t\mathrm{e}^{-\frac{t^2}{2}}}{t}\ \mathrm{d}t \\ &\geq \frac{1}{\sqrt{2\pi}}\int_y^{y+1} \frac{t\mathrm{e}^{-\frac{t^2}{2}}}{y+1}\ \mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}(y+1)}\left[-\mathrm{e}^{-\frac{t^2}{2}}\right]_y^{y+1} \\ &= \frac{1}{\sqrt{2\pi}(y+1)}(\mathrm{e}^{-y^2/2} - \mathrm{e}^{-(y+1)^2/2}).\end{aligned}$$

For the upper bound on $\Phi^c(y)$:

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\int_y^\infty \mathrm{e}^{-\frac{t^2}{2}}\ \mathrm{d}t &= \frac{1}{\sqrt{2\pi}}\int_y^\infty \frac{t\mathrm{e}^{-\frac{t^2}{2}}}{t}\ \mathrm{d}t \\ &\leq \frac{1}{\sqrt{2\pi}}\int_y^\infty \frac{t\mathrm{e}^{-\frac{t^2}{2}}}{y}\ \mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}y}\left[-\mathrm{e}^{-\frac{t^2}{2}}\right]_y^\infty \\ &= \frac{1}{\sqrt{2\pi}y}\mathrm{e}^{-y^2/2}.\end{aligned}$$

□

Lemma 5. $\Phi^c(y) \sim \frac{1}{\sqrt{2\pi}y}\mathrm{e}^{-y^2/2}$ *as* $y \to \infty$.

Proof. Using Lemma 4

$$\frac{y}{(y+1)}(1 - \frac{\mathrm{e}^{-(y+1)^2/2}}{\mathrm{e}^{-y^2/2}}) \leq \frac{\Phi^c(y)}{(1/\sqrt{2\pi}y)\mathrm{e}^{-y^2/2}} \leq 1.$$

Since

$$\frac{\mathrm{e}^{-(y+1)^2/2}}{\mathrm{e}^{-y^2/2}} = \mathrm{e}^{\frac{-2y-1}{2}}$$

then

$$\Phi^c(y) \sim \frac{1}{\sqrt{2\pi}y} \mathrm{e}^{-y^2/2}.$$

□

Alternative Proof.

$$\begin{aligned}
\Phi^c(y) &= \frac{1}{\sqrt{2\pi}} \int_y^\infty \mathrm{e}^{-x^2/2} \, \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \int_y^\infty \frac{x\mathrm{e}^{-x^2/2}}{x} \, \mathrm{d}x
\end{aligned}$$

Integration by parts gives

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \left[\frac{\mathrm{e}^{-y^2/2}}{y} - \int_y^\infty \frac{\mathrm{e}^{-x^2/2}}{x^2} \, \mathrm{d}x \right] \\
&= \frac{1}{\sqrt{2\pi}} \left[\frac{\mathrm{e}^{-y^2/2}}{y} - \int_y^\infty \frac{x\mathrm{e}^{-x^2/2}}{x^3} \, \mathrm{d}x \right]
\end{aligned}$$

Again integrating by parts

$$= \frac{1}{\sqrt{2\pi}} \left[\frac{\mathrm{e}^{-y^2/2}}{y} - \frac{\mathrm{e}^{-y^2/2}}{y^3} + \int_y^\infty \frac{3\mathrm{e}^{-x^2/2}}{x^4} \, \mathrm{d}x \right].$$

Continuing in this way, we obtain the asymptotic series

$$\Phi^c y = \frac{\mathrm{e}^{-y^2/2}}{\sqrt{2\pi}y} \left[1 - \frac{1}{y^2} + \frac{3}{y^4} - \frac{15}{y^6} + \frac{105}{y^8} - \dots \right].$$

□

Lemma 6. *If*

1. $\{f_n\}$ *is a sequence of monotone functions* $f_n : \mathbb{R} \to [0, 1]$
2. $\{f_n\}$ *converges pointwise to* f
3. $f : \mathbb{R} \to \mathbb{R}$ *with* $f(\mathbb{R}) \supseteq (0, 1)$

4. f is continuous.

Then this convergence is uniform.

Proof. 1. Let $\epsilon > 0$ be given.

2. Without loss of generality, f_n is monotone increasing and $\lim_{x\to-\infty} f_n(x) = 0$ and $\lim_{x\to\infty} f_n(x) = 1$. This means that $\lim_{x\to-\infty} f(x) = 0$ and $\lim_{x\to\infty} f(x) = 1$.

3. There exists x so that for all $z \geq x$, $|f(z) - 1| < \epsilon$. Let $N_1 \in \mathbb{N}$ be so that $n \geq N_1$ means that $|f_n(x) - f(x)| < \epsilon$ and so $|f_n(x) - 1| < 2\epsilon$. Since f_n is monotone for all $z \geq x$, $|f_n(z) - 1| \leq |f_n(x) - 1| < 2\epsilon$. Thus, for all $z \geq x$ and $n \geq N_1$,

$$|f_n(z) - f(z)| \leq |f_n(z) - 1| + |f(z) - 1| < 3\epsilon.$$

4. Similarly, there exists y and N_2 so that for all $z \geq y$ and all $n \geq N_2$

$$|f_n(z) - f(z)| \leq |f_n(z)| + |f(z)| < 3\epsilon.$$

5. Since $[y, x]$ is compact, $f(t)$ is uniformly continuous on $[y, x]$. There exists $\delta > 0$ for continuity on $[y, x]$. Choose R so large that $(x-y)/R < \delta$. Define $a_j = y + j((x - y)/R)$. and $j = 1, \ldots, R$ partition $[y, x]$ into subintervals $[a_{j-1}, a_j]$ each of diameter less than δ.

6. There exists m_j so that for all $n \geq m_j$, $|f_n(a_{j-1}) - f(a_{j-1})| < \epsilon$. There exists $\hat{m_j}$ so that for all $n \geq \hat{m_j}$, $|f_n(a_j) - f_j(a_j)| < \epsilon$.

7. Choose $n \geq M_j := \max\{m_j, \hat{m_j}\}$ and pick $z \in [a_{j-1}, a_j]$. Note that

$$f_n(a_{j-1}) \leq f_n(z) \leq f_n(a_j)\text{implies } f(a_{j-1}) - \epsilon \leq f_n(z) \leq f(a_j) + \epsilon.$$

Since $|f(a_j) - f(a_{j-1})| < \epsilon$ implies $f(a_j) - 2\epsilon \leq f_n(z) \leq f_n(a_j) + \epsilon$. Thus, we have

$$\begin{aligned} |f(z) - f_n(z)| &\leq |f(z) - f(a_j)| + |f_n(z) - f(a_j)| \\ &< \epsilon + 2\epsilon \\ &= 3\epsilon. \end{aligned}$$

8. Choose $N := \max\{N_1, N_2, M_1, \dots, M_R\}$. Then for $n > N$ and for all z, $|f(z) - f_n(z)| \le 3\epsilon$.

□

Lemma 7. *The convergence in the de Moivre-Laplace Central Limit Theorem is uniform in both a and b.*

Proof. Let $f_n(b) = \mathbb{P}_n\left[-\infty \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right]$ and $f(b) = \frac{1}{\sqrt{2\pi}}\int_\infty^b \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x$. Then f_n is monotone increasing and $f_n : \mathbb{R} \to [0,1]$. By the de Moivre-Laplace Central Limit Theorem to be proved below, $\{f_n\}$ converges pointwise to f. Clearly by Lemma 3, $f : \mathbb{R} \to \mathbb{R}$ with $f(\mathbb{R}) = [0,1]$. Finally $f(b)$ is continuous. Therefore, by Lemma 6. □

We need the following statement of Stirling's Formula:

Lemma 8 (Stirling's Formula)**.** *For each $n > 0$, set*

$$n! = \sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n}(1 + \epsilon_n).$$

There exists a real constant A so that $|\epsilon_n| < \frac{A}{n}$.

Proof. From Theorem 1 in Stirling's Formula from the Sum of Average Differences we know that

$$n! = \sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n}(1 + \epsilon_n).$$

There exists a real constant A so that $|\epsilon_n| < \frac{A}{n}$.

From Theorem 13 in Stirling's Formula Derived from the Gamma Function we know that for $n \ge 2$,

$$\left| \frac{n!}{\sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n}} - 1 - \frac{1}{12n} \right| \le \frac{1}{288n^2} + \frac{1}{9940n^3}.$$

Therefore, we can take

$$\epsilon_n < \frac{1}{n}\left(\frac{1}{12} + \frac{1}{288n} + \frac{1}{9940n^2}\right)$$

and we can take A can be taken as an upper bound on $\frac{1}{12} + \frac{1}{288n} + \frac{1}{9940n^2}$.

From the last conclusion in we know that

$$\sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n} < n! < \sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n+1/(12n)}.$$

From Corollary 1 in Stirling's Formula by Euler-Maclaurin Summation we know that

$$\sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n} < n! < \sqrt{2\pi} n^{n+1/2} \mathrm{e}^{-n+1/(12(n-1/2))}.$$

Any of these error estimates on Stirling's Formula is sufficient to establish the conclusion of the Lemma. □

Definition. Let $(s_{n,k})_{n>0,k\in I_n}$ and $(t_n)_{n>0}$ for $t_n > 0$ be two sets of real numbers. Then we say $s_{n,k} = O_u(t_n)$ if $|s_{n,k}| \le c t_n$ for all $k \in I_n$. Here, O_u means big-O uniformly.

Lemma 9 (de Moivre-Laplace Binomial Point Mass Limit)**.**

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{1}{\sqrt{2\pi \cdot np(1-p)}} \mathrm{e}^{\left(-\frac{(k-np)^2}{2np(1-p)}\right)} \cdot (1 + \delta_n(k))$$

where for $a > 0$,

$$\lim_{n\to\infty} \max_{|k-np|<a\sqrt{n}} |\delta_n(k)| = 0.$$

Proof. 1. Set $I_n := \{k : np - a\sqrt{n} < k < np + a\sqrt{n}\}$. Then the max in the statement of the lemma is taken over I_n.

2. Using Stirling's Formula

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$
$$= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{np}{k}\right)^k \left(\frac{n(1-p)}{(n-k)}\right)^{n-k} \left(\frac{1+\epsilon_n}{(1+\epsilon_k)(1+\epsilon_{n-k})}\right).$$

3. For $k \in I_n$, we have

$$\frac{n}{(np + a\sqrt{n})(n(1-p) + a\sqrt{n})} \le \frac{n}{k(n-k)} \le \frac{n}{(np - a\sqrt{n})(n(1-p) - a\sqrt{n})}. \quad (2)$$

This inequality is established in an exercise at the end of the section.

4. Then

$$\frac{1}{np(1-p)} \cdot \left(1+\frac{a}{p}\frac{1}{\sqrt{n}}\right)^{-1}\left(1+\frac{a}{1-p}\frac{1}{\sqrt{n}}\right)^{-1} \le \frac{n}{k(n-k)}$$
$$\le \frac{1}{np(1-p)}\left(1-\frac{a}{p}\frac{1}{\sqrt{n}}\right)^{-1}\left(1-\frac{a}{1-p}\frac{1}{\sqrt{n}}\right)^{-1}$$

so

$$\frac{1}{np(1-p)} \cdot \left(1-\frac{a}{p}\frac{1}{\sqrt{n}}+O\left(\frac{1}{n}\right)\right)\left(1-\frac{a}{1-p}\frac{1}{\sqrt{n}}+O\left(\frac{1}{n}\right)\right) \le \frac{n}{k(n-k)}$$
$$\le \frac{1}{np(1-p)}\left(1+\frac{a}{p}\frac{1}{\sqrt{n}}+O\left(\frac{1}{n}\right)\right)\left(1+\frac{a}{1-p}\frac{1}{\sqrt{n}}+O\left(\frac{1}{n}\right)\right).$$

Thus,

$$\frac{1}{np(1-p)} \cdot \left(1-\frac{c}{\sqrt{n}}+O\left(\frac{1}{n}\right)\right) \le \frac{n}{k(n-k)}$$
$$\le \frac{1}{np(1-p)}\left(1+\frac{c}{\sqrt{n}}+O\left(\frac{1}{n}\right)\right).$$

5. Summarizing

$$\sqrt{\frac{n}{k(n-k)}} = \frac{1}{\sqrt{np(1-p)}}\left(1+O_u\left(\frac{1}{\sqrt{n}}\right)\right),$$

because $\sqrt{1+h} = 1+\frac{1}{2}h+\dots$.

6. This establishes the square root factor in the de Moivre-Laplace Binomial Point Mass Limit. Next is the approximation of the $p^k(1-p)^{n-k}$ factors as an exponential.

7. Recall the series expansions and inequalities

   (a) $\ln(1+t) = t - \frac{t^2}{2} + O\left(t^3\right)$,

   (b) For $k \in I_n$,

$$\frac{-a\sqrt{n}}{np-a\sqrt{n}} \le \frac{k-np}{k} \le \frac{a\sqrt{n}}{np+a\sqrt{n}}.$$

(c) For $k \in I_n$, $\frac{k-np}{k} = O_u\left(n^{-1/2}\right)$,

(d) For $k \in I_n$, $\frac{k-np}{n-k} = O_u\left(n^{-1/2}\right)$,

These expansions are established in the exercises at the end of the section.

8. For $k \in I_n$

$$\ln\left[\left(\frac{n}{k}p\right)^k\left(\frac{n}{n-k}(1-p)\right)^{n-k}\right]$$
$$= k\ln\left(1-\frac{k-np}{k}\right)+(n-k)\ln\left(1+\frac{k-np}{n-k}\right)$$
$$= -\frac{1}{2}(k-np)^2\left(\frac{1}{k}+\frac{1}{n-k}\right)+kO_u(n^{-3/2})+(n-k)O_u(n^{-3/2})$$
$$= -\frac{1}{2}(k-np)^2\frac{1}{np(1-p)}+O_u(n^{-1/2}).$$

9. Therefore,

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

10. Recall that for $k \in I_n$

(a) $\epsilon_n < A/n$,

(b) $\epsilon_k < A/k$,

(c) $\epsilon_{n-k} < A/(n-k)$,

(d)
$$\frac{1}{k} = O_u\left(\frac{1}{n}\right),$$

(e)
$$\frac{1}{n-k} = O_u\left(\frac{1}{n}\right)$$

so that

$$\frac{1+\epsilon_n}{(1+\epsilon_k)(1+\epsilon_{n-k})} = \left(1 + O_u\left(\frac{1}{n}\right)\right)$$

11. Now combining parts (a), (b) and (c) above, we get

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

□

Lemma 10. *Let $[a,b]$ be an interval of $\mathbb{R}$ and let f be a function defined on $\mathbb{R}$ that is zero outside of $[a,b]$ and continuous on $[a,b]$. Then for any t*

$$\lim_{h\to 0, h\geq 0} h \sum_{k=\infty}^{\infty} f(t+kh) = \int_a^b f(x)\,\mathrm{d}x$$

Remark. This lemma is a generalization and extension of the definition of Riemann integration.

Proof. The result follows from the uniform continuity of the function f on the interval $[a,b]$. Let $\epsilon > 0$ be given, and then choose h small enough that $|f(x) - f(y)| < \epsilon$ whenever $x, y \in [a,b]$ and $|x-y| < h$. Let

$$\{k \in \mathbb{Z} | a \leq t + kh \leq b\} = \{i, i+1, i+2, \ldots, j\}$$

and $M = \sup_{a\leq x\leq b} |f(x)|$. Then with the Triangle Inequality we have that

$$\begin{aligned} &\left| h \sum_{k: t+kh\in[a,b]} f(t+kh) - \int_a^b f(x)\,\mathrm{d}x \right| \\ &\quad \leq hM + \sum_{k=i}^{j-1} \left| hf(t+kh) - \int_{t+kh}^{t+(k+1)h} f(x)\,\mathrm{d}x \right| + 2hM \\ &\quad 3hM + (j-1)h\epsilon \leq 3hM + (b-a)\epsilon. \end{aligned}$$

The term hM at the beginning comes from the term $hf(t+jh)$ which does not occur in the sum. The term $2hM$ at the end comes from the two leftover integral portions $\int_a^{t+ih} f(x)\,\mathrm{d}x$ and $\int_{t+jh}^b f(x)\,\mathrm{d}x$. □

Now we can complete the proof of the de Moivre-Laplace Central Limit Theorem.

Proof. Completion of Proof of the Central Limit Theorem The proof shows that a sum of binomial point masses can be expressed in the form of Lemma 10

Case 1, a and b are finite real numbers

Let K_n be the interval

$$[a\sqrt{np(1-p)}, b\sqrt{np(1-p)}]$$

1. We start with

$$\mathbb{P}_n\left[S_n - np \in K_n\right] = \sum_{k=0}^{n} 1_{K_n}(k-np) \cdot \mathbb{P}_n\left[S_n = k\right]$$
$$= \frac{1}{\sqrt{2\pi np(1-p)}} \sum_{k=0}^{n} \left[1_{K_n}(k-np) \exp\left(-\frac{(k-np)^2}{2np(1-p)}\right) \cdot (1+\delta_n(k))\right]$$

   where $\lim_{n\to\infty} \max_k |\delta_n(k)| = 0$ by the de Moivre-Laplace Binomial Point Mass Limit.

2. Then

$$\mathbb{P}_n\left[S_n - np \in K_n\right]$$
$$= \frac{1+\delta_n}{\sqrt{2\pi np(1-p)}} \sum_{k=0}^{n} 1_{K_n}(k-np) \cdot \exp\left(-\frac{(k-np)^2}{2np(1-p)}\right)$$

   where $\lim_{n\to\infty} \delta_n = 0$.

3. When n is large enough, the expression in the previous paragraph as a sum $\sum_{k=0}^{n}$ becomes a sum $\sum_{k\in\mathbb{Z}}$ so that

$$\frac{1}{\sqrt{2\pi}} \frac{1+\delta_n}{\sqrt{np(1-p)}}$$
$$\times \sum_{k\in\mathbb{Z}} \left[1_{[a,b]}\left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{\frac{np}{1-p}}\right)\right.$$
$$\left. \times \exp\left(-\frac{1}{2}\left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{np}1-p\right)^2\right)\right] \quad (3)$$

4. Expanding over the numerator of the fraction $\frac{1+\delta_n}{\sqrt{np(1-p)}}$ there is a term

$$\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{np(1-p)}} \\ \times \sum_{k \in \mathbb{Z}} \left[1_{[a,b]} \left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{\frac{np}{1-p}} \right) \right. \\ \left. \times \exp\left(-\frac{1}{2} \left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{np} 1 - p \right)^2 \right) \right]$$

which converges to a finite value as shown in the next paragraph. Then the term

$$\frac{1}{\sqrt{2\pi}} \frac{\delta_n}{\sqrt{np(1-p)}} \\ \times \sum_{k \in \mathbb{Z}} \left[1_{[a,b]} \left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{\frac{np}{1-p}} \right) \right. \\ \left. \times \exp\left(-\frac{1}{2} \left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{\frac{np}{1-p}} \right)^2 \right) \right]$$

converges to 0 and can be dropped.

5. Set $h = 1/\sqrt{np(1-p)}$ and $f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$. Then the expression in the equation (3) has the form of the limit in Lemma 10. Therefore, the expression in the equation (3) approaches

$$\frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-x^2/2} \, \mathrm{d}x$$

proving the de Moivre-Laplace Central Limit Theorem for finite values of a and b.

Case 2, $a = -\infty$ and b is a finite real number The proof must show that

$$\mathbb{P}_n \left[S_n - np \leq b\sqrt{np(1-p)} \right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b \mathrm{e}^{-x^2/2} \, \mathrm{d}x \,.$$

1. Let $b \in \mathbb{R}$ and $\epsilon > 0$. Fix $c > \max(0, b)$ so that
$$\frac{1}{\sqrt{2\pi}} \int_c^\infty \mathrm{e}^{-x^2/2}\ \mathrm{d}x < \epsilon.$$
2. Then
$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-c} \mathrm{e}^{-x^2/2}\ \mathrm{d}x < \epsilon$$
and
$$\frac{1}{\sqrt{2\pi}} \int_{-c}^{c} \mathrm{e}^{-x^2/2}\ \mathrm{d}x > 1 - 2\epsilon.$$
3. Write
$$\left| \mathbb{P}_n \left[S_n - np \le b\sqrt{np(1-p)} \right] - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b} \mathrm{e}^{-x^2/2}\ \mathrm{d}x \right| \le A_n + B_n + C$$
where
$$\begin{aligned} A_n &= \mathbb{P}_n \left[S_n - np \le -c\sqrt{np(1-p)} \right] \\ B_n &= \left| \mathbb{P}_n \left[-c \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b \right] - \frac{1}{\sqrt{2\pi}} \int_{-c}^{b} \mathrm{e}^{-x^2/2}\ \mathrm{d}x \right| \\ C &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-c} \mathrm{e}^{-x^2/2}\ \mathrm{d}x \end{aligned}$$
4. We have that
$$0 \le A_n \le 1 - \mathbb{P}_n \left[-c \le \frac{S_n - np}{\sqrt{np(1-p)}} \le c \right]$$
5. As in Part 1,
$$\lim_{n\to\infty} \mathbb{P}_n \left[-c \le \frac{S_n - np}{\sqrt{np(1-p)}} \le c \right] = \frac{1}{\sqrt{2\pi}} \int_{-c}^{c} \mathrm{e}^{-x^2/2}\ \mathrm{d}x > 1 - 2\epsilon$$

6. This shows that $A_n \leq 2\epsilon$ for large enough n. Similarly, $\lim_{n\to\infty} B_n = 0$ and $C < \epsilon$. This finishes Case 2.

Case 3, a is a finite real number and $b = -\infty$ This case is similar to Case 2.

□

Practical Applications

Weak Law

Corollary 1. *The Weak Law of Large Numbers is a direct consequence of the Central Limit Theorem. That is, we get directly that*

$$\lim_{n\to\infty} \mathbb{P}_n \left[\left| \frac{S_n}{n} - p \right| \geq \epsilon \right] \to 0.$$

Actually a stronger statement is possible:

Corollary 2. *Let u_n be a sequence such that $\lim_{n\to\infty} \frac{u_n}{\sqrt{n}}$. Then =*

$$\lim_{n\to\infty} \mathbb{P}_n \left[u_n \left| \frac{S_n}{n} - p \right| \geq \epsilon \right] \to 0.$$

Rules for Validity of the Approximation

Rules for deciding when to use this approximation: (according to Feller, volume I)

$$np(1-p) > 18.$$

Application

Example. Tony Gwyn's batting average in 1995 was 197 hits out of 535, (about .368). His lifetime average was .338. The question is whether Tony Gwyn was a "lucky" .300 hitter in 1995? We assume yes and that hits are independently distributed random variables. We want to know $\mathbb{P}_n\left[S_n \geq 197\right] = \mathbb{P}_n \left[\frac{S_{535} - 535 \cdot .3}{\sqrt{535}\sqrt{(.3)(.7)}} \geq \frac{197 - 160.5}{\sqrt{112}} \right] \approx \Phi^c(3.44) \approx .00115.$

That is, the probability of this many hits occurring "by chance" if Gwynn actually was a .300 hitter are small, about 1%, so at least under the stringent assumptions of the approximation, Gwynn actually improved in 1995.

Another question is was whether his actual "ability" was .338. Here, $p = .338$ and $\mathbb{P}_n[S_n \geq 197] = \cdots \approx \Phi^c(1.48) \approx .0694$. This is at least a believably large probability, so we admit that it may be possible.

Sources

This section is adapted from: *Heads or Tails*, by Emmanuel Lesigne, Student Mathematical Library Volume 28, American Mathematical Society, Providence, 2005, Chapter 7, [3]. See also the proofs in [1] and [2].

Problems to Work for Understanding

1. For equation (2) show that for $k \in I_n$, we have

$$\frac{n}{(np + a\sqrt{n})\,(n(1-p) + a\sqrt{n})} \leq \frac{n}{k(n-k)} \leq \frac{n}{(np - a\sqrt{n})}\left(n(1-p) - a\sqrt{n}\right)$$

2. Show that $\ln(1+t) = t - \frac{t^2}{2} + O\left(t^3\right)$.

3. Show that $\frac{k-np}{k} = O_u\left(n^{-1/2}\right)$.

4. Show that $\frac{k-np}{n-k} = O_u\left(n^{-1/2}\right)$.

5. Show that

$$\frac{-a\sqrt{n}}{np - a\sqrt{n}} \leq \frac{k - np}{k} \leq \frac{a\sqrt{n}}{np + a\sqrt{n}}.$$

6. Write out in detail Case 3 in the completion of the proof of the de Moivre-Laplace Central Limit Theorem.

Reading Suggestion:

References

[1] Leo Breiman. *Probability*. SIAM, 1992.

[2] William Feller. *An Introduction to Probability Theory and Its Applications, Volume I*, volume I. John Wiley and Sons, third edition, 1973. QA 273 F3712.

[3] Emmanuel Lesigne. *Heads or Tails: An Introduction to Limit Theorems in Probability*, volume 28 of *Student Mathematical Library*. American Mathematical Society, 2005.

[4] George Pólya. Über den zentralen grenzwertsatz der wahrscheinlichkeitsrechnung und das momentenproblem. *Mathematische Zeitschrift*, 8:171–181, 1920.

Outside Readings and Links:

1.
2.
3.
4.

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 LaTeX source on December 9, 2011