Stochastic-Process Limits

An Introduction to Stochastic-Process Limits And their Application to Queues

Ward Whitt

AT&T Labs - Research
The Shannon Laboratory
Florham Park, New Jersey

Draft
June 13, 2001

Preface

0.1. What Is This Book About?

This book is about *stochastic-process limits*, i.e., limits in which a sequence of stochastic processes converges to another stochastic process. Since the converging stochastic processes are constructed from initial stochastic processes by appropriately scaling time and space, the stochastic-process limits provide a macroscopic view of uncertainty. The stochastic-process limits are interesting and important because they generate simple approximations for complicated stochastic processes and because they help explain the statistical regularity associated with a macroscopic view of uncertainty.

This book emphasizes the continuous-mapping approach to obtain new stochastic-process limits from previously established stochastic-process limits. The continuous-mapping approach is applied to obtain stochastic-process limits for *queues*, i.e., probability models of service systems or waiting lines. These limits for queues are called *heavy-traffic limits*, because they involve a sequence of models in which the offered loads are allowed to increase towards the critical value for stability. These heavy-traffic limits generate simple approximations for complicated queueing processes under normal loading and reveal the impact of variability upon queueing performance. By focusing on the application of stochastic-process limits to queues, this book also provides an introduction to heavy-traffic stochastic-process limits for queues.

0.2. In More Detail

More generally, this is a book about *probability theory* – a subject which has applications to every branch of science and engineering. Probability theory can help manage a portfolio and it can help engineer a communication

Contents

Chapter 1

Experiencing Statistical Regularity

1.1. A Simple Game of Chance

A good way to experience statistical regularity is to repeatedly play a game of chance. So let us consider a simple game of chance using a spinner. To attract attention, it helps to have interesting outcomes, such as falling into an alligator pit or winning a dash for cash (e.g., you receive the opportunity to run into a bank vault and drag out as many money bags as you can within thirty seconds). However, to focus on statistical regularity, rather than fear or greed, we consider repeated plays with a simple outcome.

In our game, the payoff in each of several repeated plays is determined by spinning the spinner. We pay a fee for each play of the game and then receive the payoff indicated by the spinner. Let the payoff on the spinner be uniformly distributed around the circle; i.e., if the angle after the spin is θ, then we receive $\theta/2\pi$ dollars. Thus our payoff on one play is U dollars, where U is a uniform random number taking values in the interval $[0,1]$.

We have yet to specify the fee to play the game, but first let us simulate the game to see what cumulative payoffs we might receive, not counting the fees, if we play the game repeatedly. We perform the simulation using our favorite random number generator, by generating n uniform random numbers $U_1, \ldots, U_n$, each taking values in the interval $[0,1]$, and then forming associated partial sums by setting

$$S_k \equiv U_1 + \cdots + U_k, \quad 1 \leq k \leq n \ ,$$

and $S_0 \equiv 0$, where $\equiv$ denotes equality by definition. The n^{th} partial sum S_n is the total payoff after n plays of the game (not counting the fees to play the game). The successive partial sums form a *random walk*, with U_n being the n^{th} step and S_n being the position after n steps.

1.1.1. Plotting Random Walks

Now, using our favorite plotting routine, let us plot the random walk, i.e., the $n+1$ partial sums S_k, $0 \leq k \leq n$, for a range of n values, e.g., for $n = 10^j$ for several values of j. This simulation experiment is very easy to perform. For example, it can be performed almost instantaneously with the statistical package *S* (or *S-Plus*), see Becker, Chambers and Wilks (1988) or Venables and Ripley (1994), using the function

```
walk <- function(j) {
uniforms <- runif(10^j)              # generate random numbers
firstsums <- cumsum(uniforms)        # form the partial sums
sums <- c(0, firstsums)              # include a 0th sum
index <- order(sums) -1              # adjust the index
plot(index, sums) }                  # do the plotting
```

Plots of the random walk with $n = 10^j$ for $j = 1, \ldots, 4$ are shown in Figure 1.1. For small n, e.g., for $n = 10$, we see irregularly spaced (vertically) points increasing to the right, but as n increases, the spacing between the points becomes blurred and regularity emerges: The plots approach a straight line with slope equal to $1/2$, the mean of a single step U_k. If we look at the pictures in successive plots, ignoring the units on the axes, we see that the plots become independent of n as n increases. Looking at the plot for large n produces a macroscopic view of uncertainty.

The plotter automatically plots the random walk $\{S_k : 0 \leq k \leq n\}$ in the available space. Ignoring the units on the axes is equivalent to regarding the plot as a display in the unit square. By "unit square" we do not mean that the rectangle containing the plot is necessarily a square, but that new units can range from 0 to 1 on both axes, independent of the original units. The plotter automatically plots the random walk in the available space by

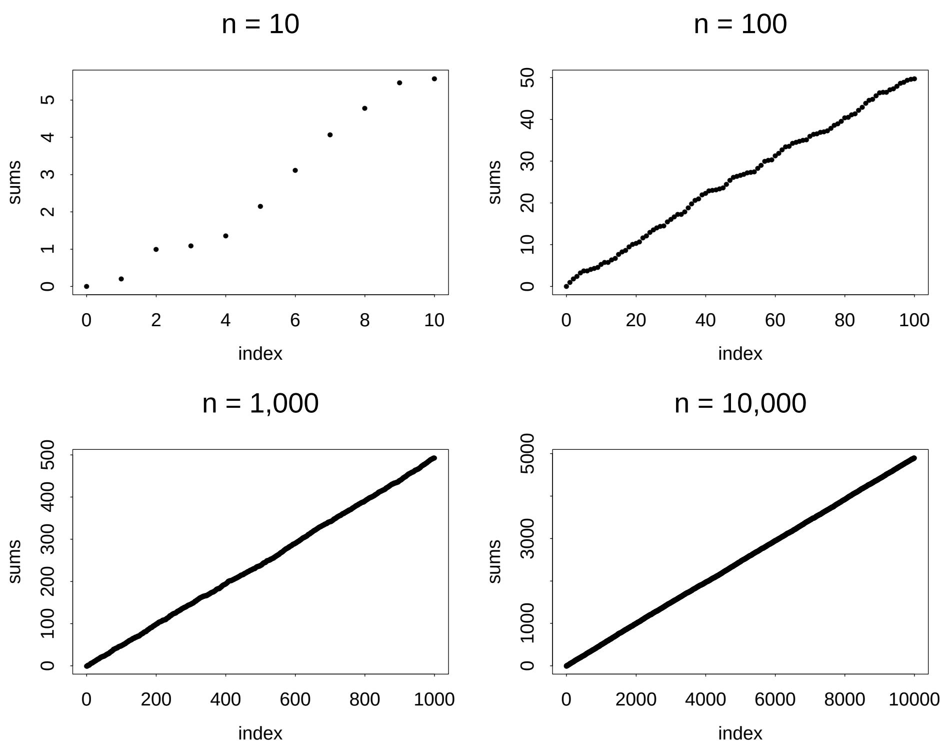

Figure 1.1: Possible realizations of the first 10^j steps of the random walk $\{S_k : k \geq 0\}$ with steps uniformly distributed in the interval $[0, 1]$ for $j = 1, \ldots, 4$.

scaling time and space (the horizontal and vertical dimensions). Time is scaled by placing the $n+1$ points $1/n$ apart horizontally. Space is scaled by subtracting the minimum and dividing by the range (assuming that the range is not zero); i.e., we interpret the plot as

$$plot(\{S_k : 0 \le k \le n\}) \equiv plot(\{(S_k - min)/range : 0 \le k \le n\}) ,$$

where

$$min \equiv min(\{S_k : 0 \le k \le n\})$$

and

$$range \equiv max(\{S_k : 0 \le k \le n\}) - min(\{S_k : 0 \le k \le n\}) .$$

Combining these two forms of scaling, the plotter displays the ordered pairs $(k/n, (S_k - min)/range)$ for $0 \le k \le n$. With that scaling, the ordered pairs do indeed fall in the unit square. Also note that $(S_k - min)/range$ must assume (approximately) the values 0 and 1 for at least one argument. That occurs because, without the rescaling, the plotting makes the units on the ordinate (y axis) range from the minimum value to the maximum value (approximately).

To confirm the regularity we see in Figure 1.1, we should repeat the experiment. When we repeat the experiment with different random number seeds (new uniform random numbers), the outcome for small n changes somewhat from experiment to experiment, but we always see essentially the same picture for large n. Thus the plots show regularity associated with both large n and repeated experiments.

1.1.2. When the Game is Fair

Now let us see what happens when the game is fair. Since the expected payoff is 1/2 dollar each play of the game, the game is fair if the fee to play is 1/2 dollar. To examine the consequences of making the game fair, we consider a minor modification of the simulation experiment above: We repeat the experiment after subtracting the mean 1/2 from each step of the random walk; i.e., we plot the *centered random walk* (i.e., the centered partial sums $S_k - k/2$ for $0 \le k \le n$) for the same values of n as before.

If we consider the case $n = 10^4$, it is natural to expect to see a horizontal line instead of the line with slope 1/2 in Figure 1.1. However, what we see is very different! Instead of a horizontal line, for $n = 10^4$ we see an irregular path, as shown in Figure 1.2.

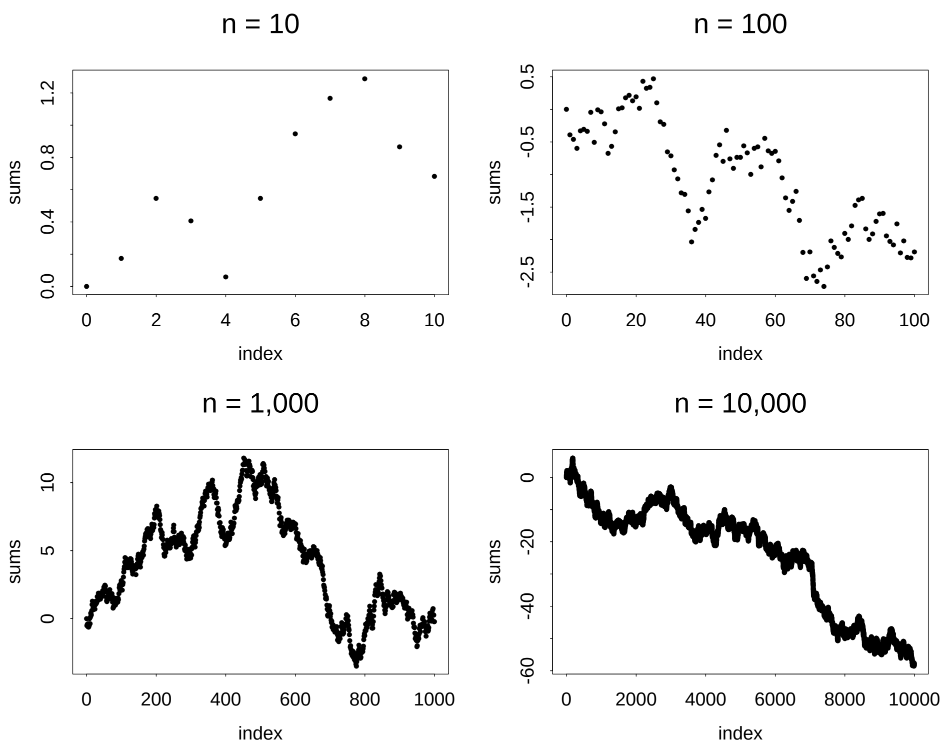

Figure 1.2: Possible realizations of the first 10^j steps of the centered random walk $\{S_k - k/2 : k \geq 0\}$ with steps uniformly distributed in the interval $[0, 1]$ for $j = 1, \ldots, 4$.

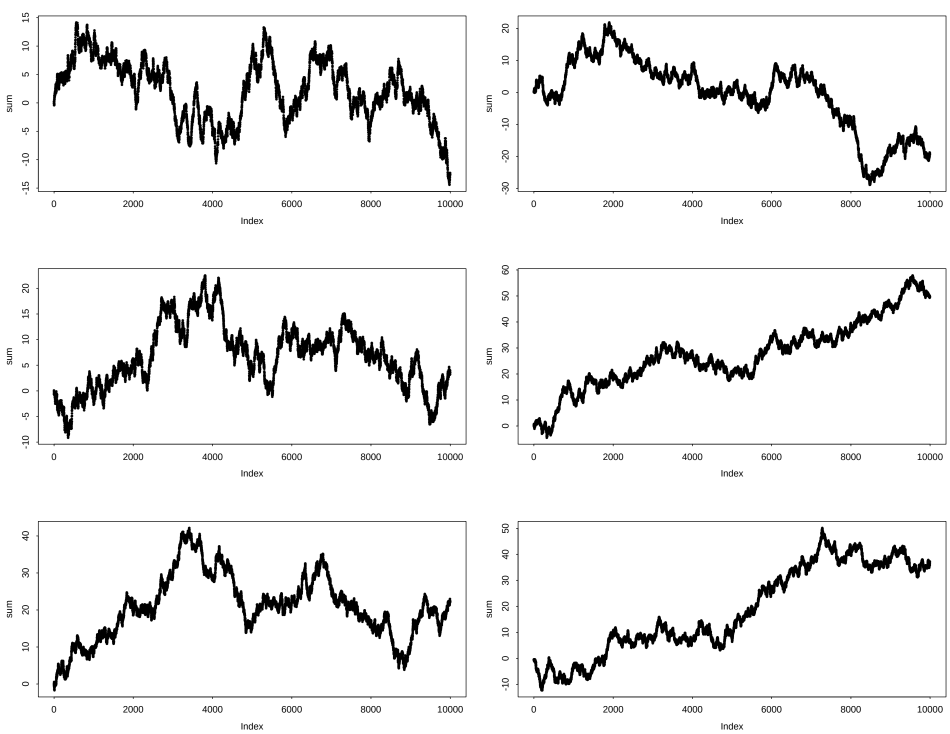

Figure 1.3: Six independent realizations of the first 10^4 steps of the centered random walk $\{S_k - k/2 : k \geq 0\}$ associated with steps uniformly distributed in the interval $[0, 1]$.

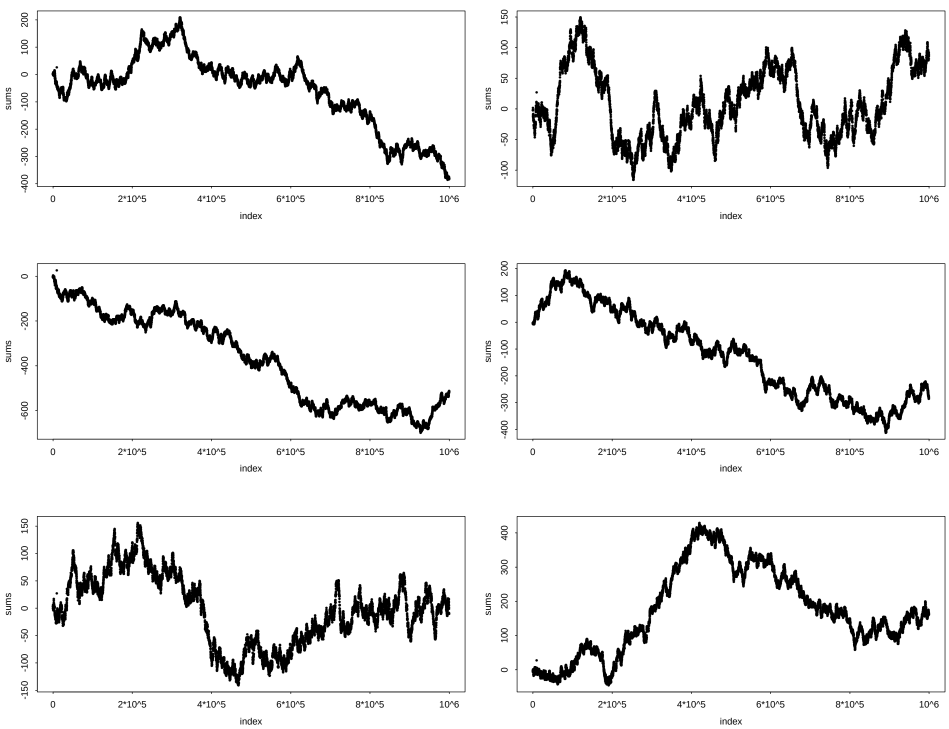

Figure 1.4: Six independent realizations of the first 10^6 steps of the centered random walk $\{S_k - k/2 : k \geq 0\}$ associated with steps uniformly distributed in the interval $[0, 1]$.

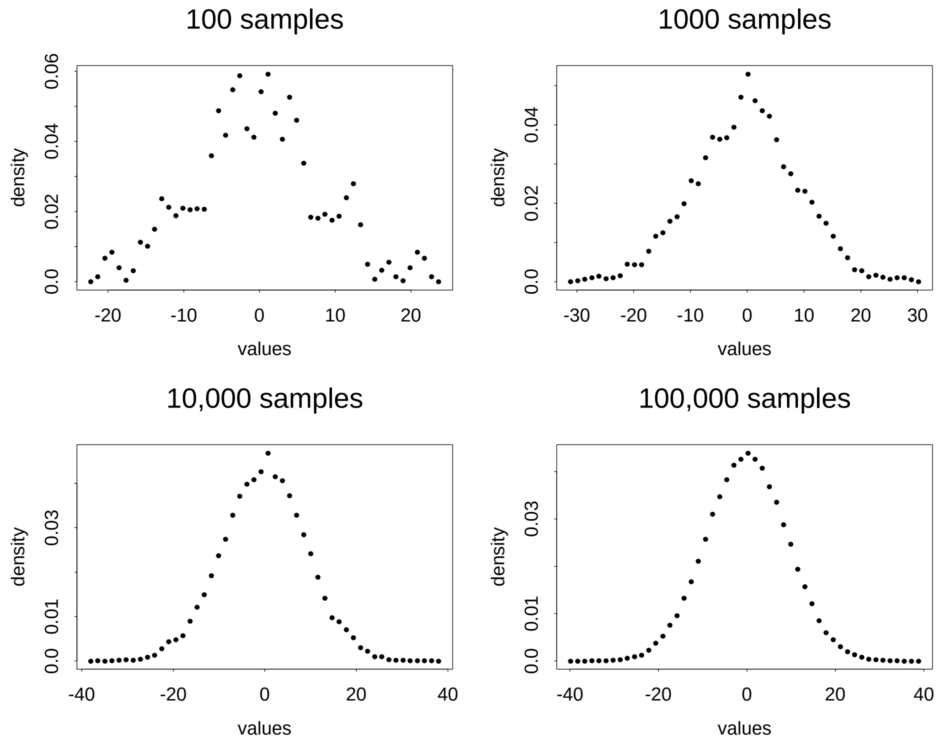

Figure 1.7: Estimates of the probability density of the final position of the random walk, obtained from 10^j independent samples of the centered partial sum $S_{1000} - 500$ for $j = 2, \ldots, 5$, for the case in which the steps U_k are uniformly distributed in the interval $[0, 1]$, based on the nonparametric density estimator *density* from *S*.

parameter settings). Essentially the same plots are obtained for independent samples from normal distributions. From Figure 1.7, it is evident that the density estimates converge to a normal pdf as $n \to \infty$. For more on density estimation, see Devroye (1987).

It is not our purpose to delve deeply into statistical issues, but it is worth remarking that we obtain new interesting plots, like the random walk plots, when we do. Our brief examination of the distribution of the final position of the random walk suggests looking for a more precise statistical test to determine whether or not the final position of the random walk is indeed approximately normally distributed. To evaluate whether some data can be regarded as an independent sample any specified probability distribution, it

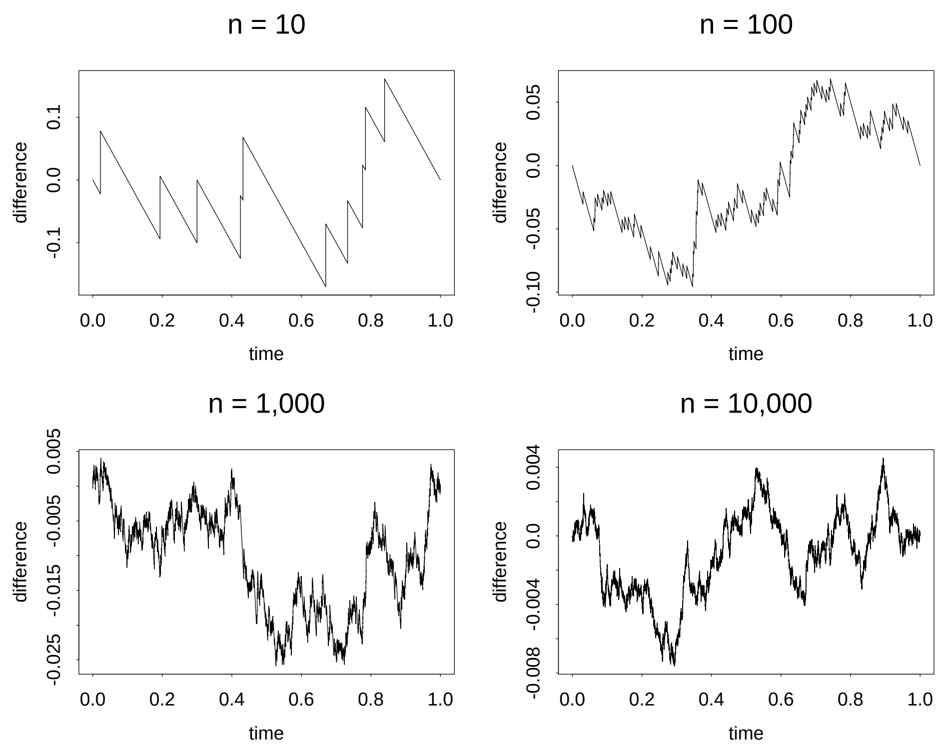

Figure 1.8: The difference between the empirical cdf and the actual cdf for samples of size 10^j from the uniform distribution over the interval $[0, 1]$ for $j = 1, \ldots, 4$.

position. That makes sense as well, because both the empirical cdf and the actual cdf must assume the common value 1 at the right endpoint.

It turns out that there is statistical regularity in the empirical cdf's just like there is in the random walks. As before, the plots look the same for all sufficiently large n. Moreover, except for having the final position be 0, the plots look just like the random-walk plots. More generally, this example illustrates that statistical analysis is an important source of motivation for stochastic-process limits. We discuss this example further in Section 2.2. There we show how to develop a statistical test applicable to any continuous cdf, including the normal cdf that is of interest for the final position of the random walk.

1.1.4. Making an Interesting Game

We have digressed from our original game of chance to consider the statistical regularity observed in the plots, which of course really is our main interest. But now let us return for a moment to the game of chance.

A gambling house cannot afford to make the game fair. The gambling house needs to charge a fee greater than the expected payoff in order to make a profit. What would be a good fee for the gambling house to charge?

From the perspective of the gambling house, one might think the larger the fee the better, but the players presumably have the choice of whether or not to play. If the gambling house charges too much, few players will want to play. The fee should be large enough for the gambling house to make money, but small enough so that potential players will want to play. We take that to mean that the individual players should have a good chance of winning.

One might think that those objectives are inconsistent, but they are not. The key to achieving those objectives is the realization that *the player and the gambling house experience the game in different time scales.* An individual player might contemplate playing the game 100 times on a single day, while the gambling house might offer the game to hundreds or thousands of players on each of many consecutive days.

Thus, the player might evaluate his experience by the possible outcomes from about 100 plays of the game, while the gambling house might evaluate its experience by the possible outcomes from something like $10^4 - 10^6$ plays of the game. What we need, then, is a fee close enough to \$0.50 that the player has a good chance of winning in 100 plays, while the gambling house receives a good reliable return over $10^4 - 10^6$ games.

A reasonable fee might be \$0.51, giving the gambling house a 1 cent or 2% advantage on each play. (Gambling houses actually tend to take more, which shows the appeal of gambling despite the odds.) To see how the \$0.51 fee works, let us consider the possible experiences of the player and the gambling house. In Figure 1.9 we plot six independent realizations of a player's position during 100 plays of the game when there is a fee of \$0.51 for each play. The game looks pretty interesting for the player from Figure 1.9. The player has a reasonable chance of winning. Indeed, the player wins in plots 3 and 5, and finishes about even in plot 2. How do things look for the gambling house?

To see how the gambling house fares, we should look at the net payoffs over a much larger number of games. Hence, in Figures 1.10 and 1.11 we plot

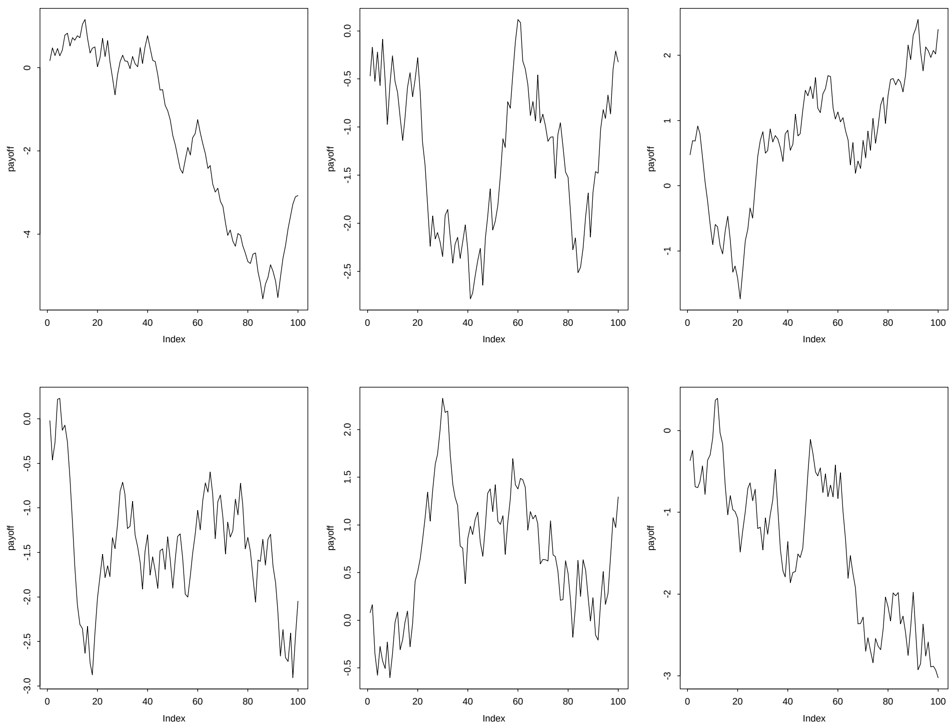

Figure 1.9: Six possible realizations of the first 100 net payoffs, positions of the random walk $\{S_k - 0.51k : k \geq 0\}$, with steps U_k uniformly distributed in the interval $[0, 1]$ and a fee of \$0.51.

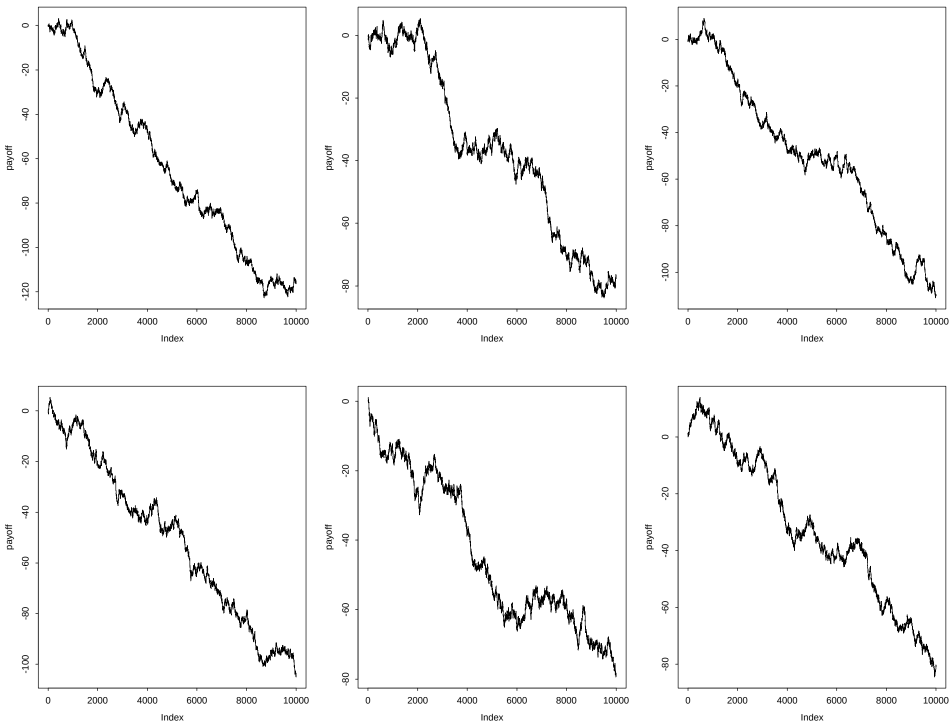

Figure 1.10: Possible realizations of the first 10^4 net payoffs (steps of the random walk $\{S_k - 0.51k : k \geq 0\}$ with steps U_k uniformly distributed in the interval $[0, 1]$.

six independent realizations of a player's position during 10^4 and 10^6 plays of the game. As before, we let the plotter automatically do the scaling, so that the units on the vertical axes change from plot to plot. But that does not alter the conclusions. In these larger time scales, we see that the player consistently loses money, so that a profit for the gambling house becomes essentially a sure thing. When we increase the number of plays to 10^6, there is little randomness left. That is shown in Figure 1.11. Further repetitions of the experiment confirm these observations. We again see the regularity associated with a macroscopic view of uncertainty.

Above we picked a candidate fee out of the air. We could instead be more systematic. For example, we might seek the largest fee such that the player satisfies some criteria indicating a good experience. Letting the fee for each game be f, we might want to constrain the probability p that a

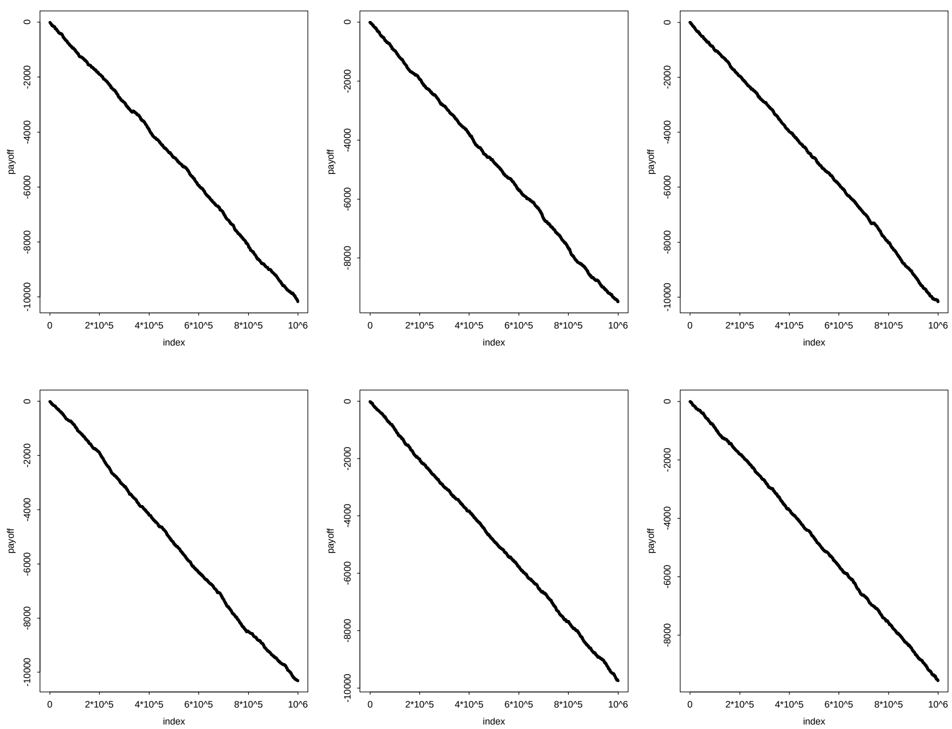

Figure 1.11: Possible realizations of the first 10^6 net payoffs (steps of the random walk $\{S_k - 0.51k : k \geq 0\}$ with steps U_k uniformly distributed in the interval $[0, 1]$.

player wins at least a certain amount w, i.e., by requiring that

$$P(S_{100} - f(100) \geq w) \geq p \ .$$

Given such a formulation, we can determine the optimal fee f, i.e., the maximum fee f such that the constraint is satisfied, which is attained when the probability just equals p.

As noted at the outset, when we consider making the game interesting, we might well conclude that a uniform payoff distribution for each play is boring. We might want to have the possibility of much larger positive and/or negative payoffs on one play. It is easy to devise more interesting games with different payoff distributions, but the statistical regularity associated with large numbers observed above tends to be the same. Readers are invited to make their own games and look at the net payoffs for 10^j plays for various values of j.

An extreme case that is often attractive is to have, like a lottery, some small chance of a very large payoff. However, with independent trials, as determined by successive spins of the spinner, the gambling house faces the danger of having to make too many large payoffs. Such large losses are avoided in lotteries by not letting the game be based on independent trials. In a lottery only a few prizes are awarded (and possibly shared) so that the people running the lottery are guaranteed a positive return. However, an insurance company cannot control the outcomes so tightly, so that careful analysis of the possible outcomes is necessary; e.g., see Embrechts, Klüppelberg and Mikosch (1997). We too will be interested in the possibility of exceptionally large values in random events.

1.2. Stochastic-Process Limits

The plots we have looked at indicate that there is statistical regularity associated with large n, i.e., with large sample sizes. We now want to understand *why* we see what we see, and what we will see in other related situations. For that purpose, we turn to probability theory; see Ross (1993) and Feller (1968) for introductions.

1.2.1. A Probability Model

We can use probability theory to explain what we have seen in the random walk plots. The first step is to introduce an appropriate mathematical model: Assuming that our random number generator is working properly (an important issue, which we will not address, e.g., see p. 123 of Venables and Ripley (1994), L'Ecuyer(1998a,b) and references cited there), the observed values U_k, $1 \leq k \leq n$, should be distributed approximately as the first n values from a sequence of *independent and identically distributed* (IID) random variables uniformly distributed on $[0, 1]$ (defined on an underlying probability space). Indeed, the model fit is usually so good that there is a tendency to identify the mathematical model with the physical experiment (a mistake), but since the model fit is so good, we need not doubt that the mathematical conclusions are applicable.

Remark 1.2.1. *Mathematics and the physical world.* It is important to realize that a physical phenomenon, a mathematical model of that physical phenomenon and a simulation of that mathematical model are three different things. But, if the mathematical model is well chosen, the three may be closely related. In particular, a mathematical model, whether simulated or analyzed, may provide useful desciptions of the physical phenomenon.

We are interested in mathematical queueing models because of their ability to explain queueing phenomena, but we should not expect a perfect match. For example, mathematical models often succeed by exploiting the infinite, even though the physical phenomenon is finite. Random numbers generated on a computer are inherently finite, and yet simulations based on random numbers can be well described by mathematical models exploiting the infinite.

Here, we perform stochastic simulations to reveal statistical regularity, and we introduce and analyze mathematical models to explain that statistical regularity. We expect to capture key features, but we do not expect a perfect fit. We want the the mathematics to explain key features observed in the simulations, and we want the simulations to confirm key features predicted by the mathematics. ■

With that attitude, let us consider the probability model consisting of a sequence of IID uniform random numbers. Within the context of that probability model, we want to formulate stochastic-process limits suggested by the plots. First, we see that as n increases the plotted random walk ceases to look discrete. For all sufficiently large n, the plotted random walk looks like a function of a continuous variable. Thus it is natural to seek a continuous-time representation of the original discrete-time random walk. We can do that by considering the associated continuous-time process $\{S_{\lfloor t \rfloor} : t \geq 0\}$, where $\lfloor \cdot \rfloor$ is the *floor function*, i.e., $\lfloor t \rfloor$ denotes the greatest integer less than or equal to t. If we also want to introduce centering, centered random walks in Figure 1.16 with $n = 10^4$, what we see is again approximately a plot of Brownian motion. ■

We can easily construct many other examples of random walks with dependent steps. For instance, we could consider a *random walk in a random environment*. A simple example has a two-state Markov-chain environment process with transition probabilities $P_{1,2} = 1 - P_{1,1} = p$ and $P_{2,1} = 1 - P_{2,2} = q$ for $0 < p < 1$ and $0 < q < 1$. We then let the k^{th} step X_k have one distribution if the Markov chain is in state 1 at the k^{th} step, and another distribution if the Markov chain is in state 2 then. We first run the Markov chain. Then, conditional on the realized states of the Markov chain, the random variables X_k are mutually independent with the appropriate distributions (depending upon the state of the Markov chain). If we consider a stationary version of the Markov chain, then the sequence $\{X_k : k \geq 1\}$ is stationary. Regardless of the initial conditions, we again see the same statistical regularity in the associated partial sums when n is sufficiently large. We invite the reader to consider such examples.

1.3.3. Different Step Distributions

Now let us return to random walks with IID steps and consider different possible step distributions. We now repeat the experiments above with various functions of the uniform random numbers, i.e., for $X_k \equiv f(U_k)$, $1 \leq k \leq n$, for different real-valued functions f. In particular, consider the following three cases:

$$\begin{aligned}
\text{(i)} \quad X_k &\equiv -m\log(1-U_k) \quad \text{for} \quad m = 1,\, 10 \\
\text{(ii)} \quad X_k &\equiv U_k^p \quad \text{for} \quad p = 1/2,\, 3/2 \\
\text{(iii)} \quad X_k &\equiv U_k^{-1/p} \quad \text{for} \quad p = 1/2,\, 3/2\;. \qquad (3.5)
\end{aligned}$$

As before, we form partial sums associated with the new summands X_k, just as in (3.4).

Before actually considering the plots, we observe that what we are doing covers the general IID case. Given the sequence of IID random variables $\{U_k : k \geq 1\}$, by the method above we can create an associated sequence of IID random variables $\{X_k : k \geq 1\}$ where X_k has an arbitrary cdf F. Letting the left-continuous inverse of F be

$$F^{\leftarrow}(t) \equiv \inf\{s : F(s) \geq t\}, \quad 0 < t < 1\;,$$

we can obtain the desired random variables X_k with cdf F by letting

$$X_k \equiv F^{\leftarrow}(U_k), \quad k \geq 1 \ . \tag{3.6}$$

Since

$$F^{\leftarrow}(s) \leq t \quad \text{if and only if} \quad F(t) \geq s \ , \tag{3.7}$$

we obtain

$$P(F^{\leftarrow}(U) \leq t) = P(U \leq F(t)) = F(t) \ ,$$

where U is a random variable uniformly distributed on $[0,1]$, which implies that $F^{\leftarrow}(U)$ has cdf F for any cdf F when U is uniformly distributed on $[0,1]$. For example, we see that X_k has an exponential distribution with mean m in case (i) of (3.5): If $F(t) = e^{-t/m}$, then $F^{\leftarrow}(t) = -m\log(1-t)$ and

$$P(X_k > t) = P(-m\log(1-U_k) > t) = P(1-U_k < e^{-t/m}) = e^{-t/m} \ .$$

Incidentally, we could also work with the right-continuous inverse of F, defined by

$$F^{-1}(t) \equiv \inf\{s : F(s) > t\} = F^{\leftarrow}(t+), \quad 0 < t < 1 \ ,$$

where $F^{\leftarrow}(t+)$ is the right limit at t, because

$$P(F^{-1}(U) = F^{\leftarrow}(U)) = 1 \ ,$$

since $F^{\leftarrow}$ and F^{-1} differ at, at most, countably many points.

Moreover, $F^{\leftarrow}(U_k)$, $k \geq 1$, are IID when U_k, $k \geq 1$, are IID. Of course, there also are other ways to generate IID random variables with specified distributions, but what we are doing is often a natural way.

So let us plot the uncentered and centered random walks with the step sizes in (3.5). When we do so for cases (i) and (ii), we see essentially the same pictures as before. For example, plots of the first 10^4 steps of the centered random walks in the four cases in (i) and (ii) of (3.5) are shown in Figure 1.17.

Again the plots look like plots of Brownian motion, indistinguishable from the plots for the uniform steps in Figure 1.3. Note that the units on the y axis change from plot to plot, but the plots themselves tend to have a common distribution.

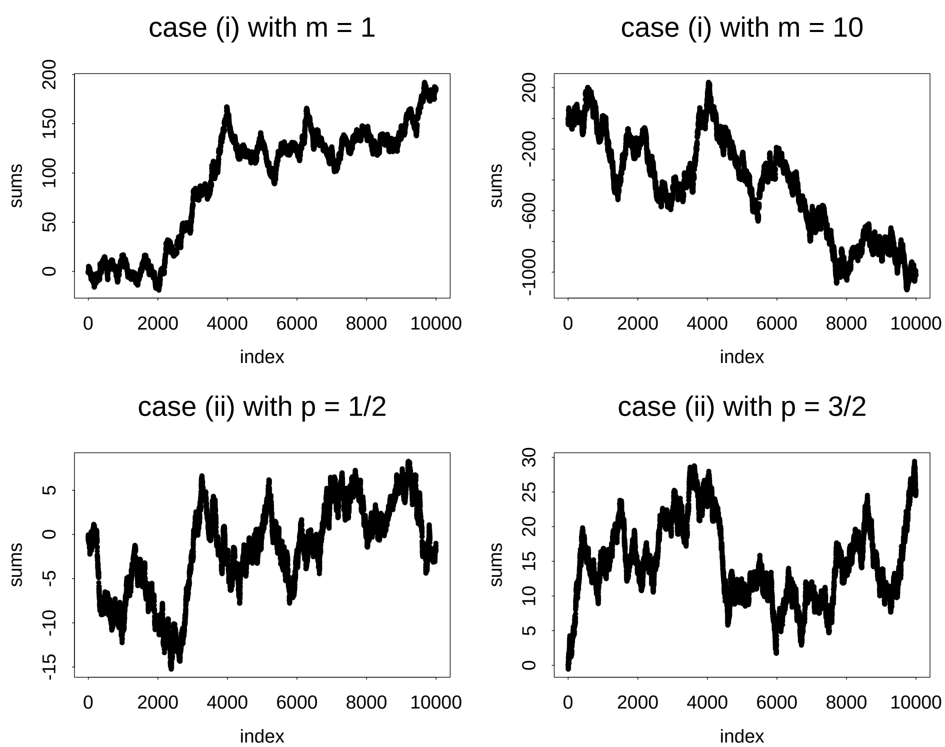

Figure 1.17: Possible realizations of the first 10^4 steps of the random walk $\{S_k - mk : k \geq 0\}$ with steps distributed as X_k in cases (i) and (ii) of (3.5).

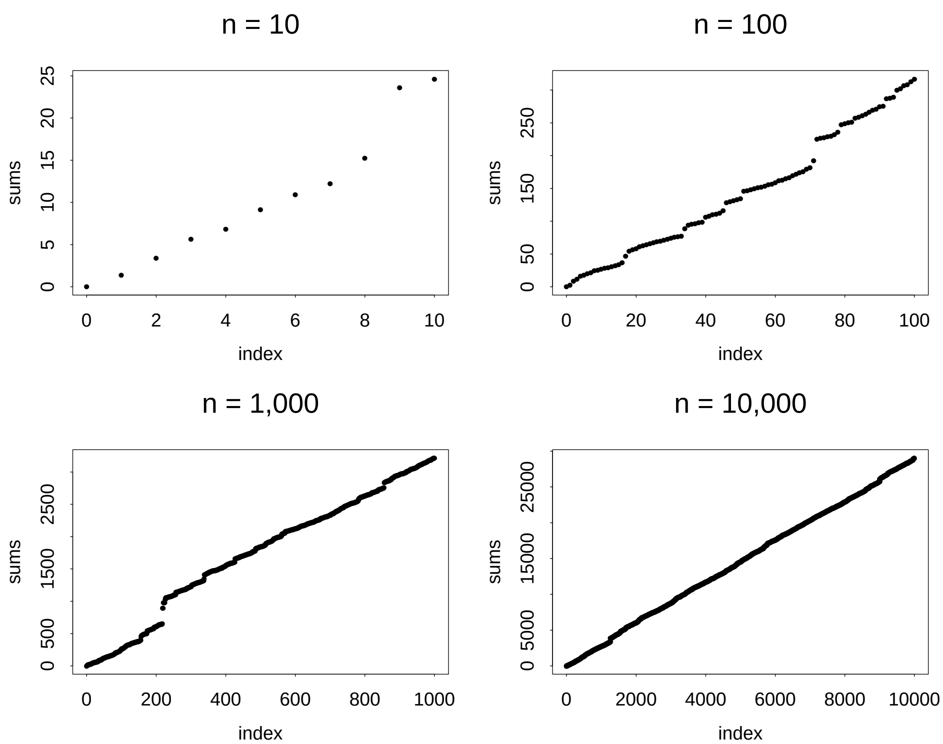

Figure 1.18: Possible realizations of the first 10^j steps of the uncentered random walk $\{S_k : k \geq 0\}$ with steps distributed as $U_k^{-1/p}$ in case (iii) of (3.5) for $p = 3/2$ and $j = 1, \dots, 4$.

1.4. The Exception Makes the Rule

Just when boredom has begun to set in, after seeing the same thing in cases (i) and (ii) in (3.5), we should be ready to appreciate the startlingly different large-n pictures in case (iii). Plots of the uncentered random walks are plotted in Figures 1.18 and 1.19.

In the case $p = 3/2$ in Figure 1.18, the plot of the uncentered random walk is again approaching a line as $n \to \infty$, but not as rapidly as before. (Again we ignore the units on the axes when we look at the plots.) However, in the case $p = 1/2$ in Figure 1.19 we something radically different: For large n, the plots have *jumps*!

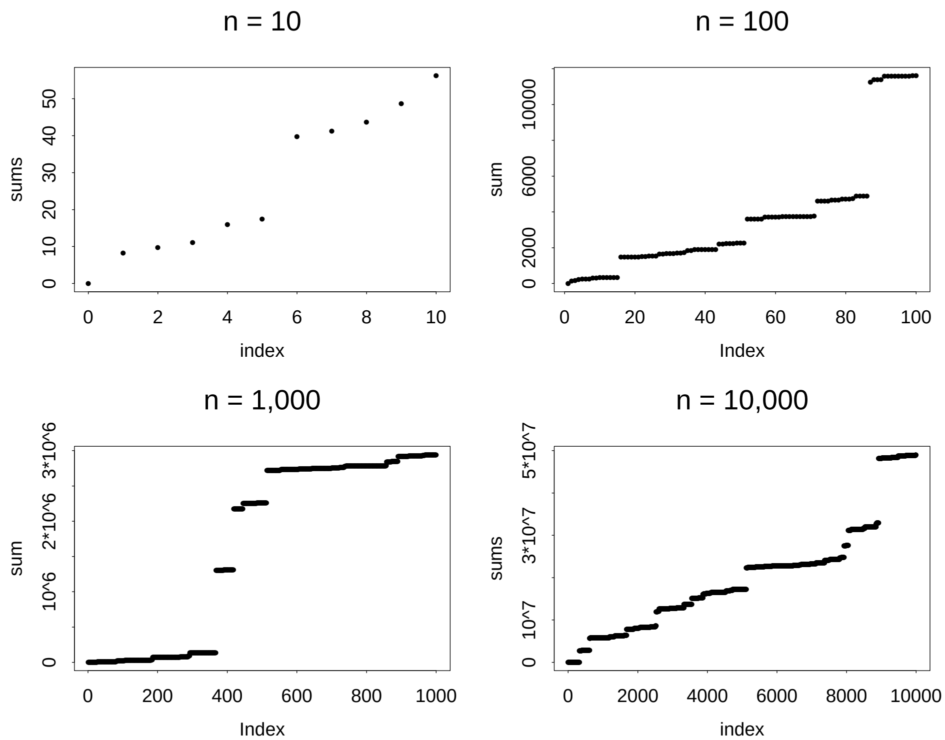

Figure 1.19: Possible realizations of the first 10^j steps of the uncentered random walk $\{S_k : k \geq 0\}$ with steps distributed as $U_k^{-1/p}$ in case (iii) of (3.5) for $p = 1/2$ and $j = 1, \ldots, 4$.

1.4.1. Explaining the Irregularity

Fortunately, probability theory again provides an explanation for the *irregularity* that we now see: The SLLN states, under the prevailing IID assumptions, that scaled partial sums $n^{-1}S_n$ will approach the mean EX_1 w.p.1 as $n \to \infty$, regardless of other properties of the probability distribution of X_1, *provided that a finite mean exists.* Knowing the SLLN, we should expect to see lines when $n = 10^4$ in all experiments except possibly in case (iii).

We might initially be fooled in case (iii), but we should anticipate occasional large steps because $U^{-1/p}$ involves *dividing* by very small values when U is small. Upon more careful examination, we see that $U^{-1/p}$ has a *Pareto distribution* with parameter p, which we refer to as Pareto(p), when U is uniformly distributed on $[0,1]$, i.e.,

$$P(U^{-1/p} > t) = P(U < t^{-p}) = t^{-p}, \quad t \geq 1\,, \tag{4.1}$$

with mean

$$E(U^{-1/p}) = \int_0^\infty P(U^{-1/p} > t)dt = 1 + \int_1^\infty t^{-p}dt\ , \tag{4.2}$$

which is finite, and equal to $1 + (p-1)^{-1}$, if and only if $p > 1$; see Chapter 19 of Johnson and Kotz (1970) for background on the Pareto distribution and Lemma 1 on p. 150 of Feller (1971) for the integral representation of the mean.

Thus the SLLN tells us not to expect the same behavior observed in the previous experiments in case (iii) when $p \leq 1$. Thus, unlike all previous random walks considered, the conditions of the SLLN are *not satisfied* in case (iii) with $p = 1/2$.

Now let us consider the random walk with Pareto(p) steps for $p = 3/2$ in (3.5) (iii). Consistent with the SLLN, Figure 1.18 shows that the plots are approaching a straight line as $n \to \infty$ in this case. But what happens when we center?

1.4.2. The Centered Random Walk with $p = 3/2$

So now let us consider the centered random walk in case (iii) with $p = 3/2$. (Since the mean is infinite when $p = 1/2$, we cannot center when $p = 1/2$. We will return to the case $p = 1/2$ later.) We center by subtracting the mean, which in the case $p = 3/2$ is $1+(p-1)^{-1} = 3$. Plots of the centered random walk with $p = 3/2$ for $n = 10^j$ with $j = 1, 2, 3, 4$ are shown in Figure

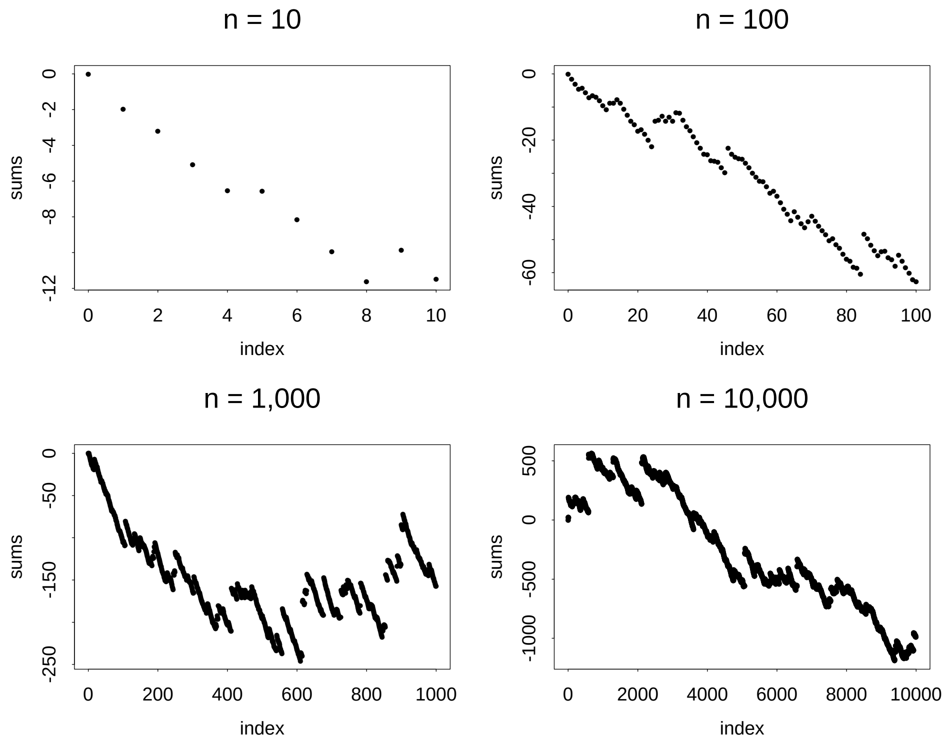

Figure 1.20: Possible realizations of the first 10^j steps of the centered random walk $\{S_k - 3k : k \geq 0\}$ associated with the Pareto steps $U_k^{-1/p}$ for $p = 3/2$, having mean 3 and infinite variance, for the cases $j = 1, \ldots, 4$.

1.20. As before, the centering causes the plotter to automatically blow up the picture. However, now the slight departures from linearity for large n in Figure 1.18 are magnified. Now, just as in Figure 1.19, we see jumps in the plot!

Once again, probability theory offers an explanation. Just as the SLLN ceases to apply when the IID summands have infinite mean, so does the (classical) CLT cease to apply when the IID summands have finite mean but infinite variance. Such a case occurs with the Pareto(p) summands in case (iii) in (3.5) when $1 < p \leq 2$. Thus, consistent with what we see in Figure 1.18, the SLLN holds, but the CLT does not, for the Pareto(p) random variable $U^{-1/p}$ in case (iii) when $p = 3/2$.

We have arrived at another critical point, where an important intellectual step is needed. We need to recognize that, *even though the sample paths are*

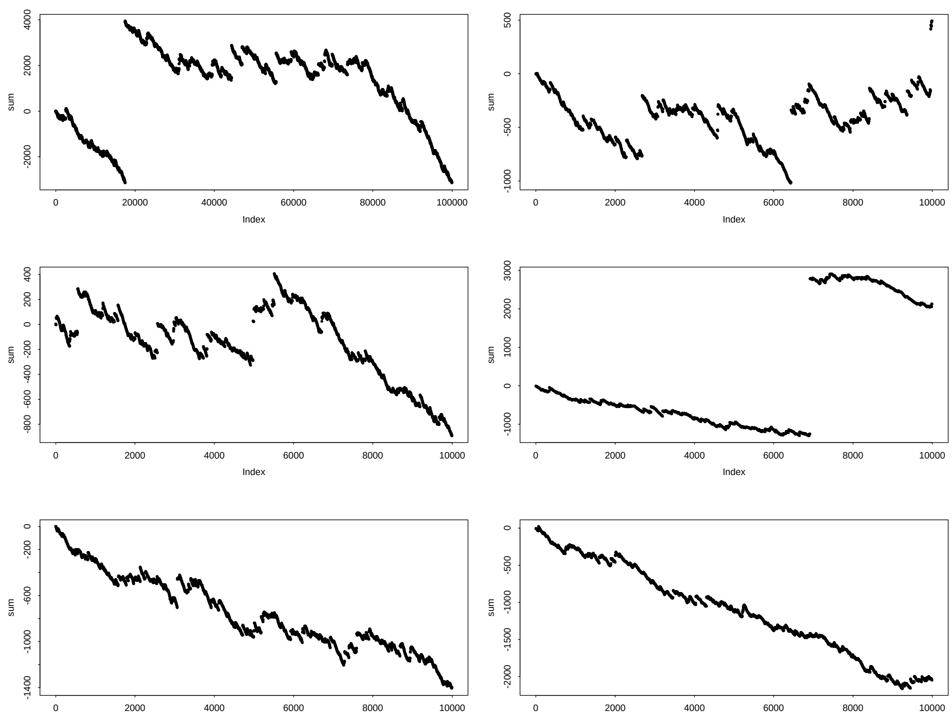

Figure 1.21: Six independent realizations of the first 10^4 steps of the centered random walk $\{S_k - 3k : k \geq 0\}$ associated with the Pareto steps $U_k^{-1/p}$ for $p = 3/2$, having mean 3 and infinite variance.

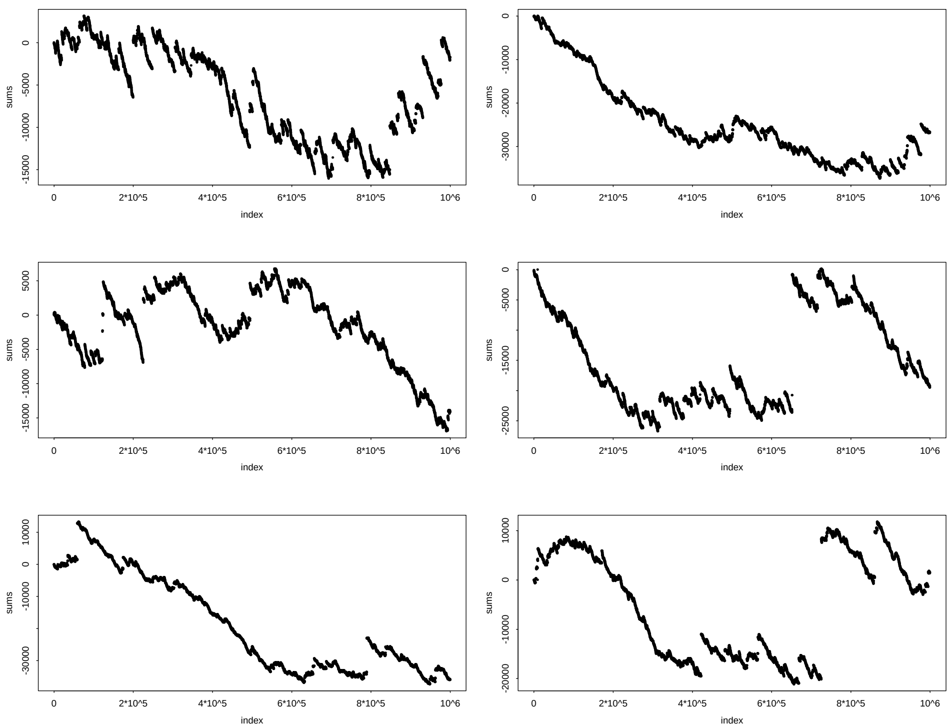

Figure 1.22: Six independent realizations of the first 10^6 steps of the centered random walk $\{S_k - 3k : k \geq 0\}$ associated with the Pareto steps $U_k^{-1/p}$ for $p = 3/2$, having mean 3 and infinite variance.

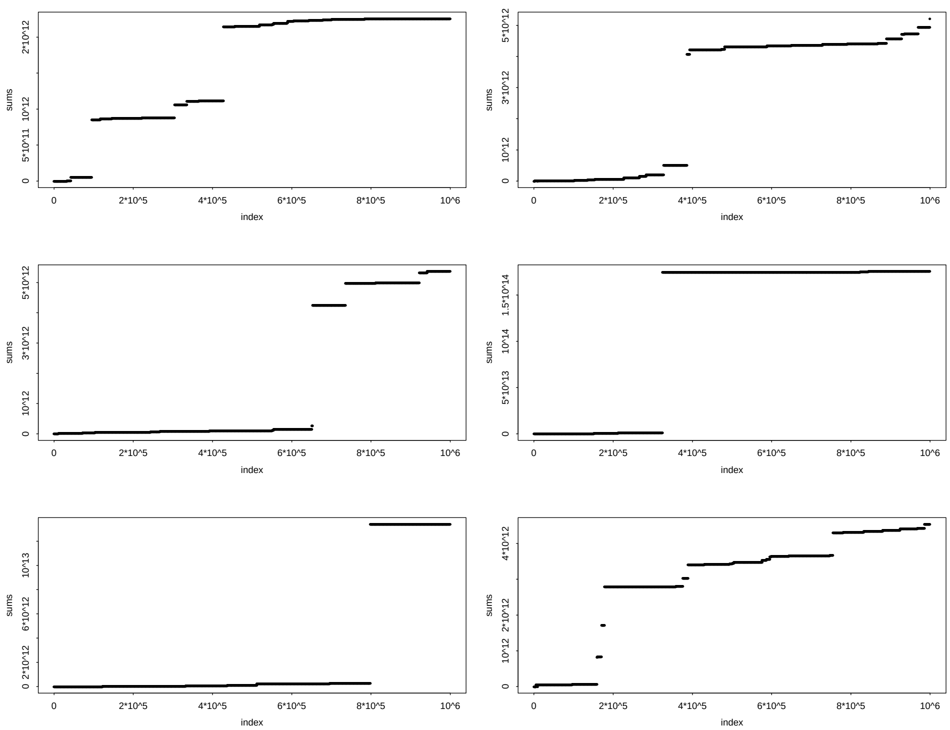

Figure 1.26: Six independent possible realizations of the first 10^6 steps of the uncentered random walk $\{S_k : k \geq 0\}$ with steps distributed as $U_k^{-1/p}$ in case (iii) of (3.5) for $p = 1/2$.

Paralleling Figures 1.4 and 1.22, we confirm what we see in Figure 1.25 by plotting six independent samples of the uncentered random walk in case (iii) with $p = 1/2$ for $n = 10^6$ in Figure 1.26. Even though the plots of the uncentered random walks with Pareto(0.5) steps in Figures 1.19 – 1.26 are radically different from the previous plots of centered and uncentered random walks, we see remarkable statistical regularity in the new plots. As before, the plots tend to be independent of n for all n sufficiently large, provided we ignore the units on the axes. Thus we see self-similarity, just as in the plots of the centered random walks before. *From the random-walk plots, we see that statistical regularity can occur in many different forms.*

Given what we have just done, it is natural to again look for statistical regularity in the final positions. Thus we consider the final positions S_n (without centering) for $n = 1000$ and perform 10,000 independent replica-