

# Experiencing Statistical Regularity

**Ward Whitt**

**Columbia University**

**October 29, 2003**

**Heavy-Tailed Distributions**

**Long-Range Dependence**

**Self-Similarity**

# Simulation Experiments

# Plotting Random Walks

---

$X_1, X_2, \dots$  **IID random variables**

$$S_n = X_1 + \dots + X_n, \quad n \geq 1,$$

with  $S_0 = 0$  **partial sums**

**Plot**  $S_0, S_1, \dots, S_n$  for  $n = 10^j$  with  $j = 1, 2, \dots$

To start:  $X_i = U_i$  **uniformly distributed** on  $[0, 1]$ .

# Program in Splus

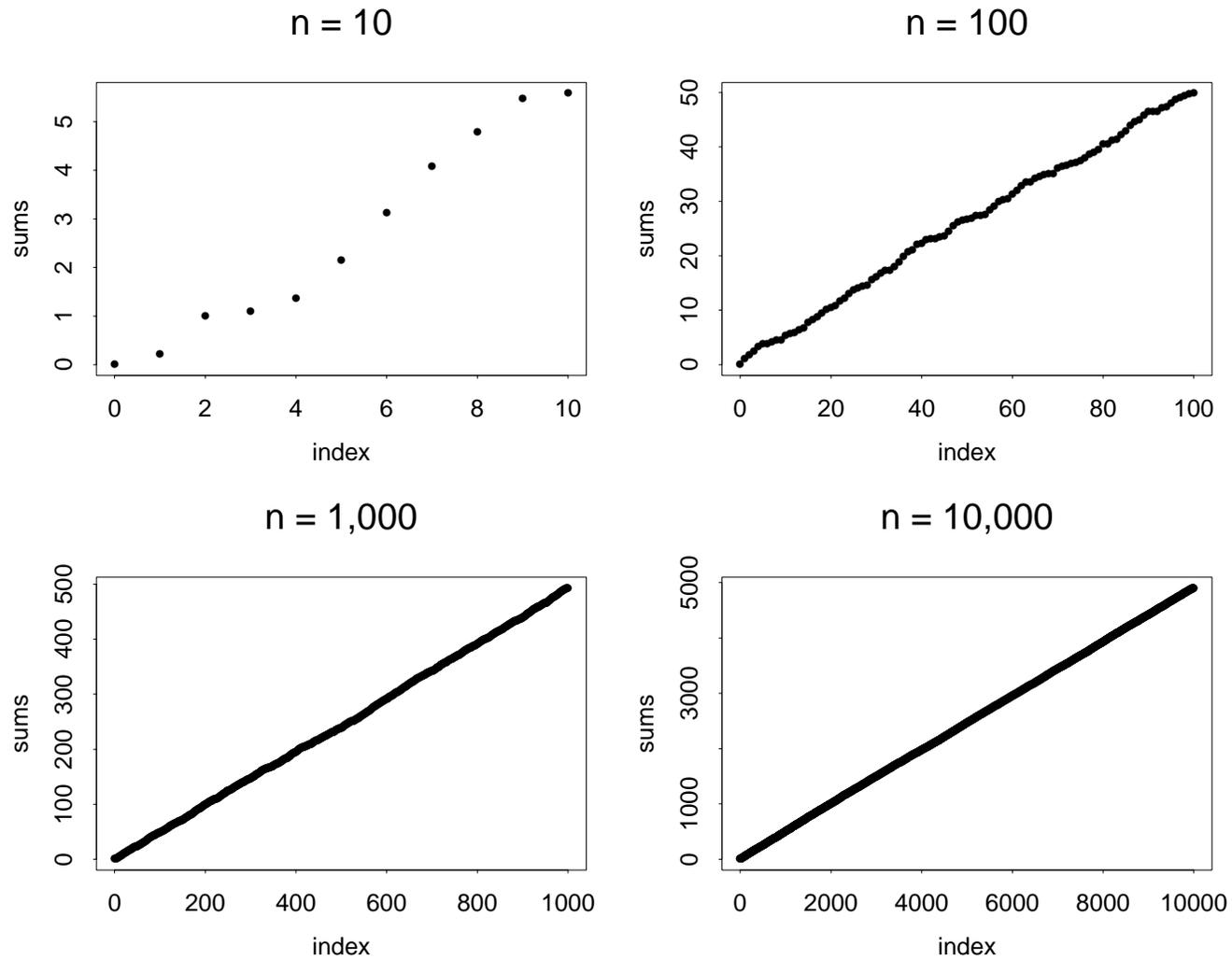
---

```
walk <- function(j) {  
  uniforms <- runif(10j)  
  firstsums <- cumsum(uniforms)  
  sums <- c(0, firstsums)  
  index <- order(sums) - 1  
  plot(index, sums) }
```

**What should we see?**

Looking at  $S_1, \dots, S_n$   
when  $n = 10^j$  for  $j = 1, 2, 3, 4$

**a stochastic-process limit**



**Plots for  $n = 10^j$  with  $j = 1, 2, 3, 4$**

**How does plotting work?**

# Map Into the Unit Square

(if we ignore the units on the axes)

# The Plot Function: Step 1

---

Fit horizontally: create a function on  $[0, 1]$ .

For  $y_0, y_1, \dots, y_n$  given, let  $x : [0, 1] \rightarrow \mathbb{R}$   
be defined by

$$x(t) \equiv y_{\lfloor nt \rfloor}, \quad 0 \leq t \leq 1,$$

where  $\lfloor z \rfloor$  is the greatest integer less than  
or equal to  $z$ .

## The Plot Function: Step 2

---

**Fit vertically.**

**Place between infimum and supremum.**

## Fit Vertically

---

For  $x : [0, 1] \rightarrow \mathbb{R}$  given,

$$\mathbf{plot}(x) \equiv (x - \mathbf{inf}(x)) / \mathbf{range}(x),$$

where

$$\mathbf{inf}(x) \equiv \inf\{x(t) : 0 \leq t \leq 1\}$$

$$\mathbf{sup}(x) \equiv \sup\{x(t) : 0 \leq t \leq 1\}$$

$$\mathbf{range}(x) \equiv \mathbf{sup}(x) - \mathbf{inf}(x)$$

When you plot a random walk,

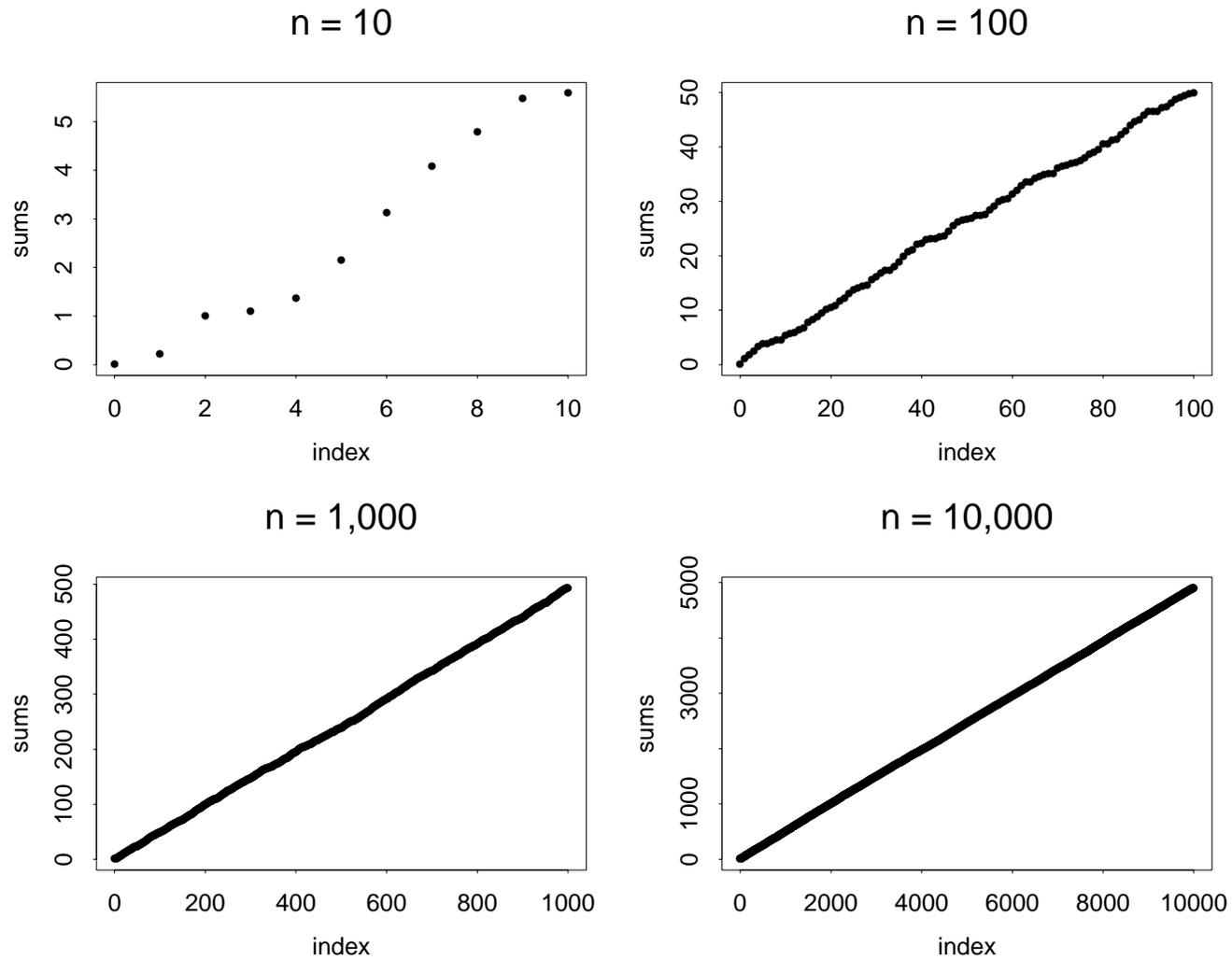
---

You get a random plot.

You get a random function.

a random function mapping  $[0, 1]$  into  $[0, 1]$ .

You get a stochastic process.



**Plots for  $n = 10^j$  with  $j = 1, 2, 3, 4$**

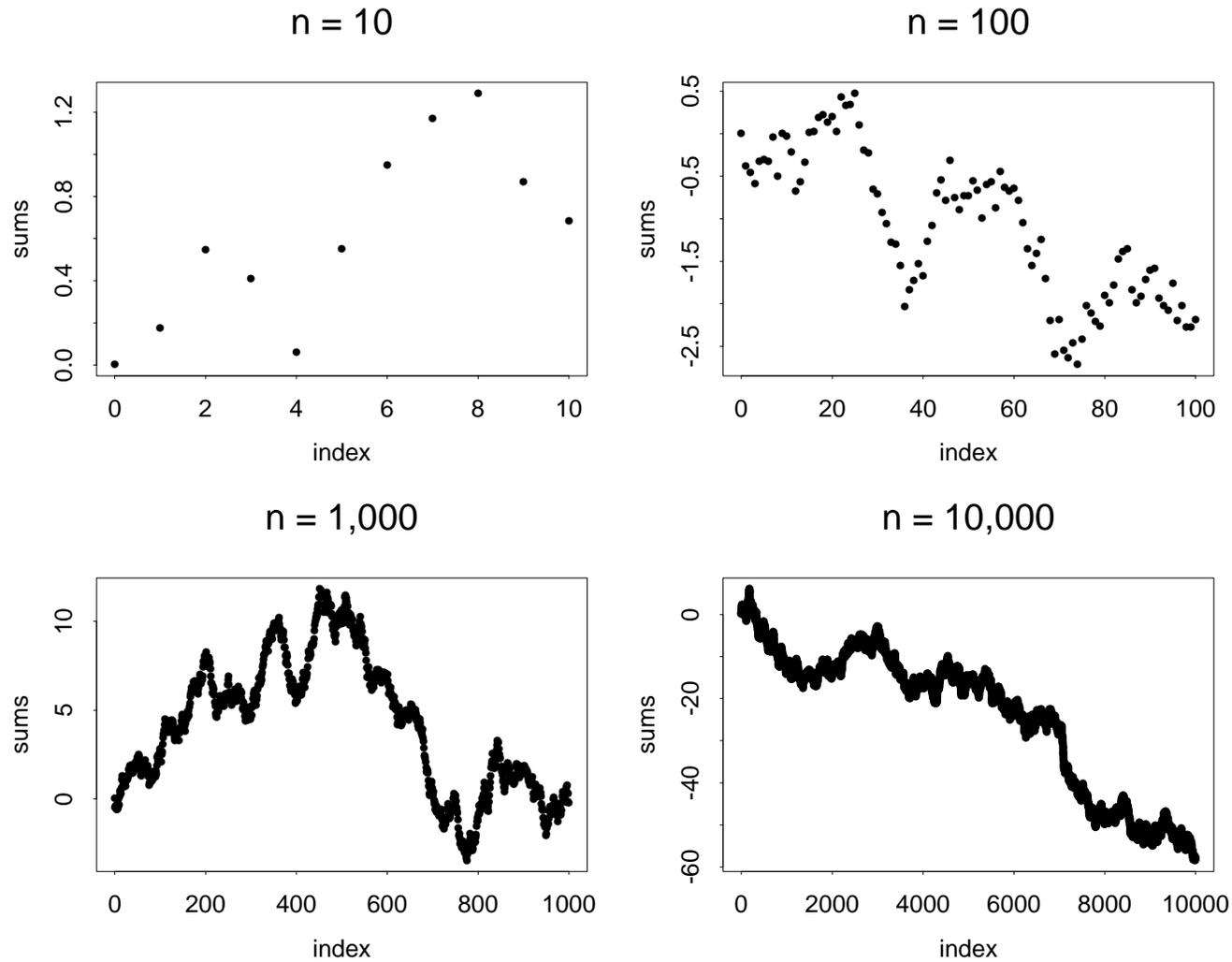
# Modified Experiment

---

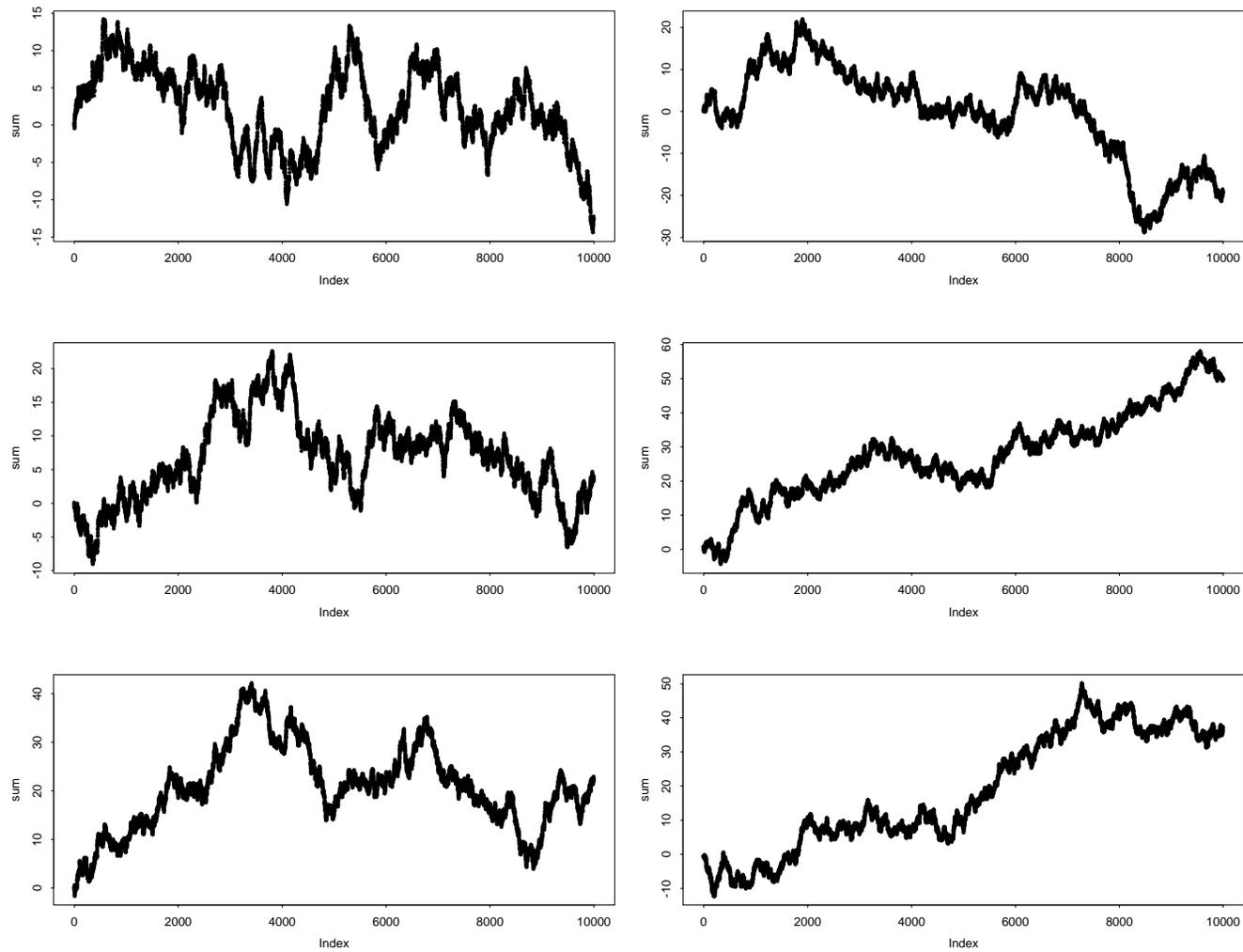
Let  $X_i = U_i - 0.5$ .

Construct **centered random walk**.

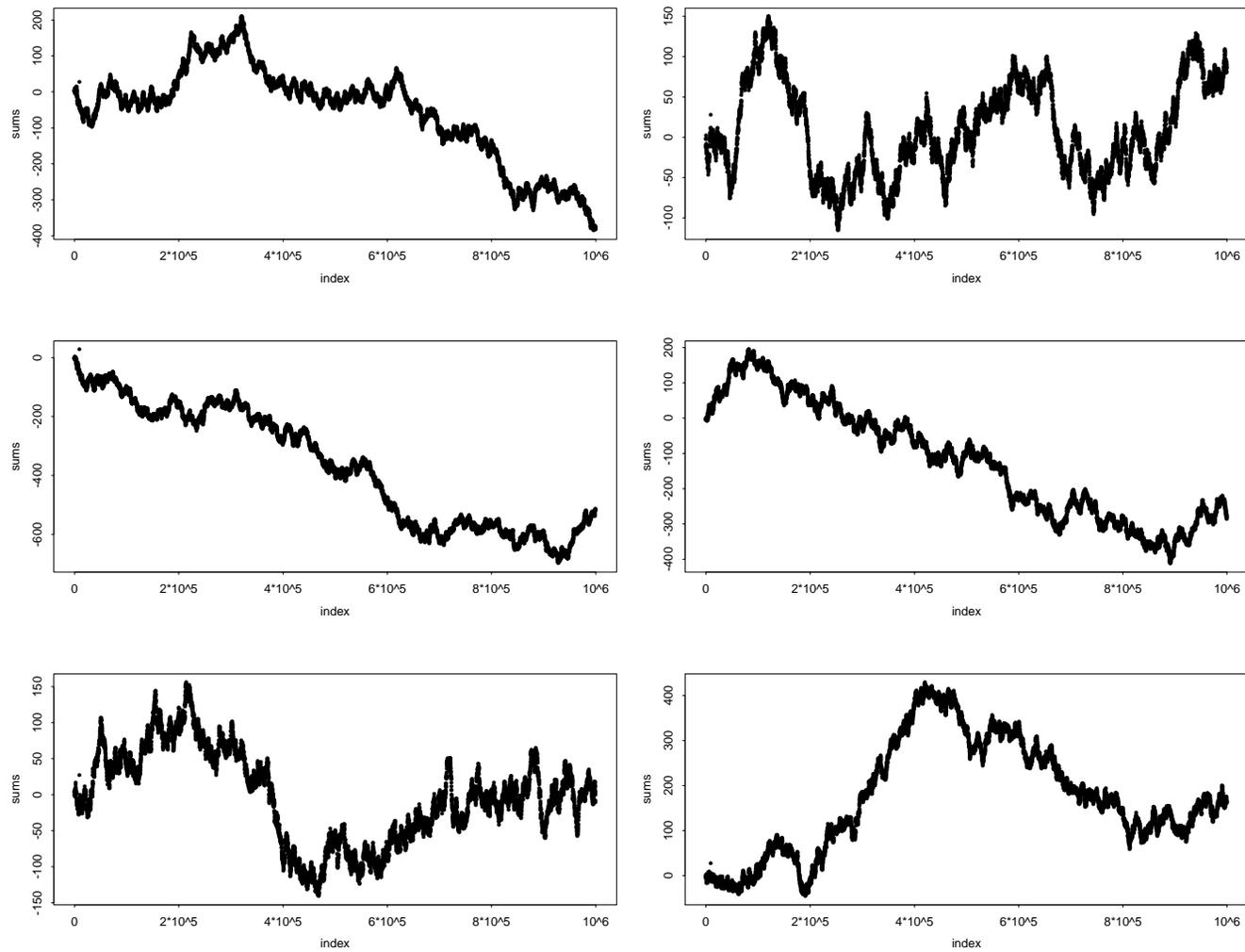
What should we see now?



The centered random walk for  $n = 10^j$  with  
 $j = 1, 2, 3, 4$



six cases for  $n = 10^4$



six cases for  $n = 10^6$

# The Random Plot Limit

---

**Theorem.** If the stochastic-process limit  $\mathcal{S}_n \Rightarrow \mathcal{S}$  holds, where

$$\mathcal{S}_n(t) \equiv (S_{\lfloor nt \rfloor} - m \lfloor nt \rfloor) / c_n, \quad 0 \leq t \leq 1,$$

for some constants  $m$  and  $c_n : n \geq 1$ , and

$$P(\text{range}(\mathcal{S}) = 0) = 0,$$

then

$$\mathbf{plot}(S_k - mk : 0 \leq k \leq n) \Rightarrow \mathbf{plot}(\mathcal{S}).$$

# Invariance Principles

# New Random Steps

---

Let  $Y_i = f(U_i)$ .

$$X_i = Y_i - EY_i.$$

**new centered random walk.**

What should we see now?

# Three Cases

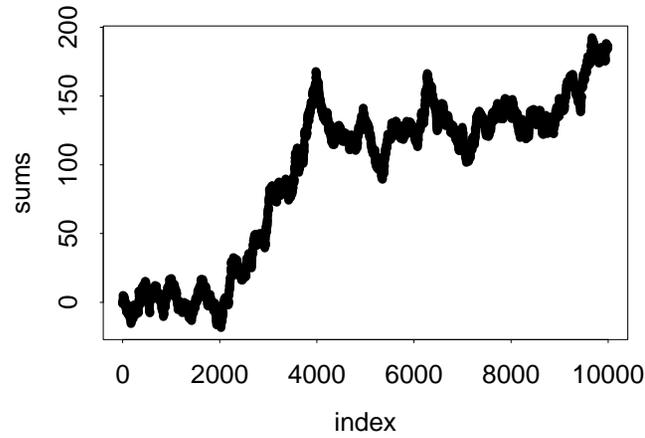
---

$$(i) f(U) = -m \log(U) \quad \text{for } m = 1, 10$$

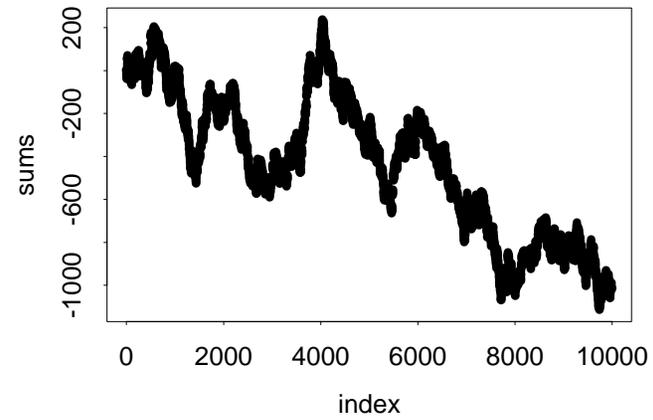
$$(ii) f(U) = U^p \quad \text{for } p = 1/2, 3/2$$

$$(iii) f(U) = U^{-1/p} \quad \text{for } p = 1/2, 3/2$$

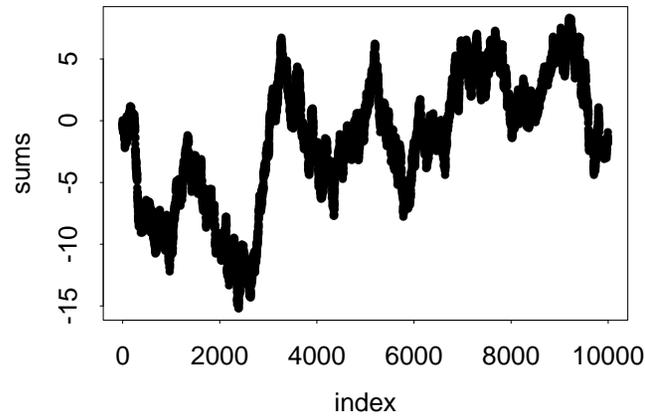
case (i) with  $m = 1$



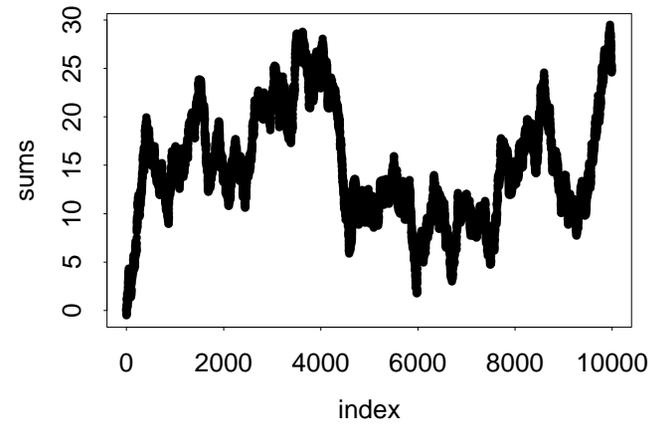
case (i) with  $m = 10$



case (ii) with  $p = 1/2$



case (ii) with  $p = 3/2$



Cases (i) and (ii) for  $n = 10^4$

**The Exception Makes the Rule**

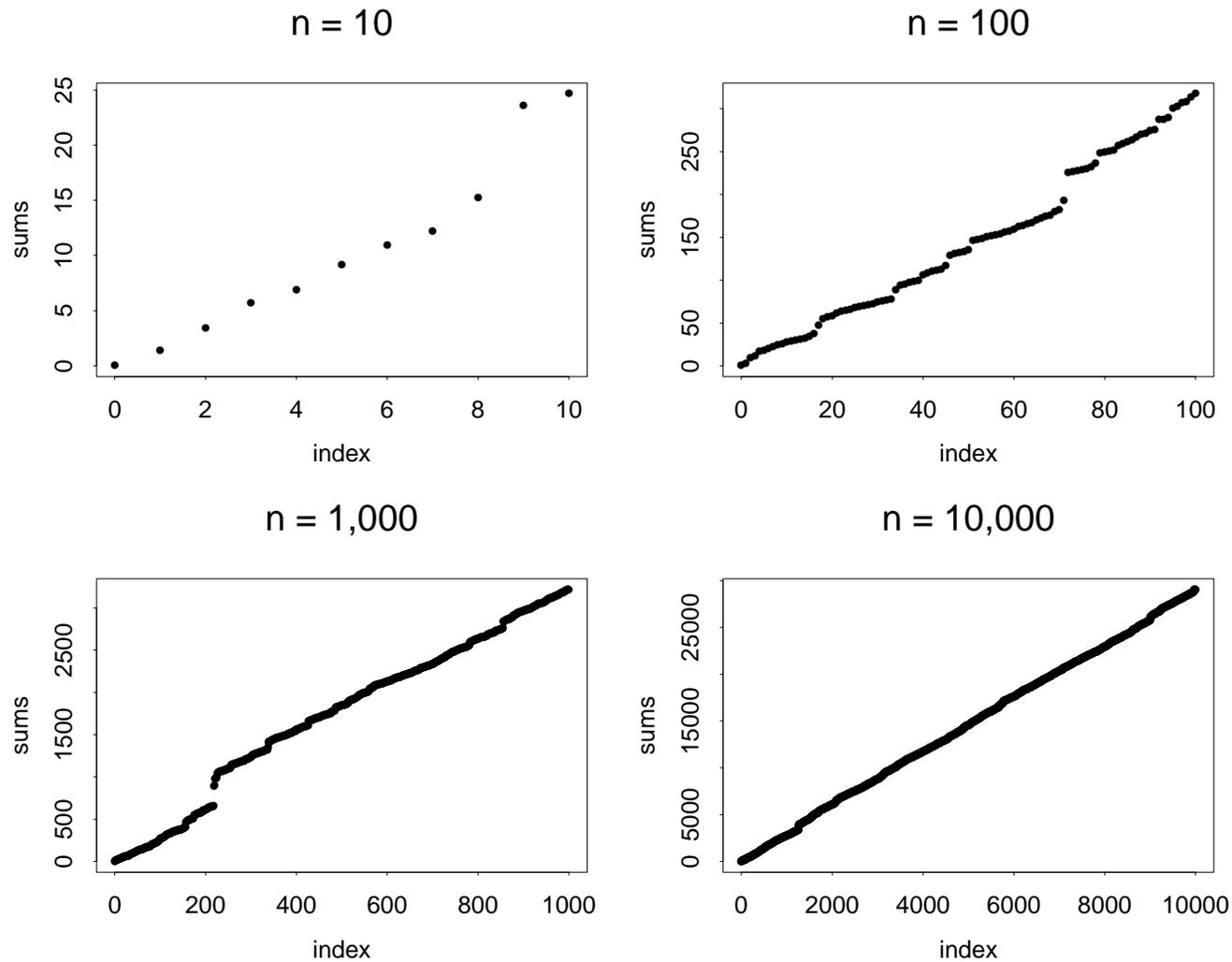
# Case (iii)

---

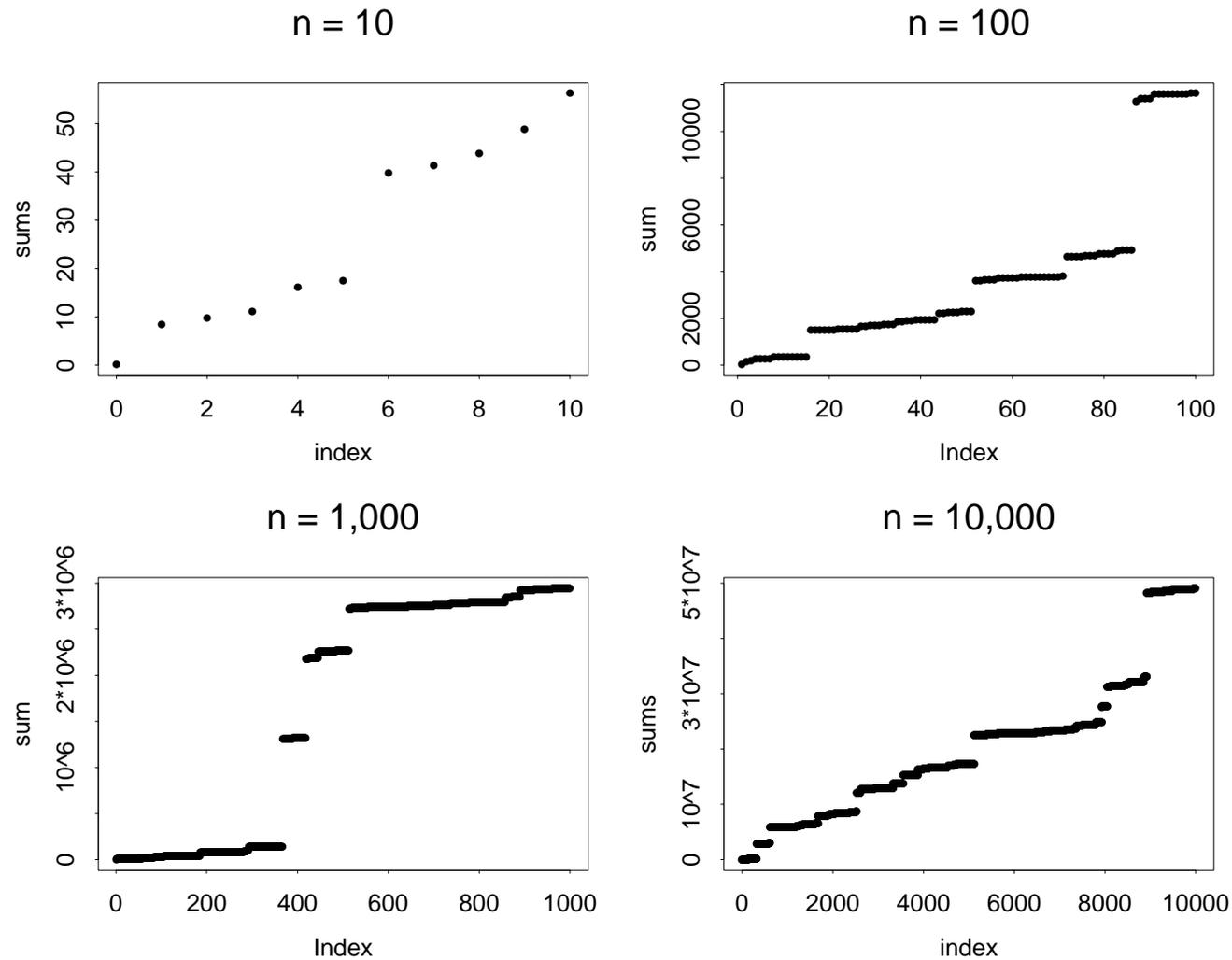
$$(i) f(U) = -m \log(U) \quad \text{for } m = 1, 10$$

$$(ii) f(U) = U^p \quad \text{for } p = 1/2, 3/2$$

$$(iii) f(U) = U^{-1/p} \quad \text{for } p = 1/2, 3/2$$



**Plots of the uncentered random walk**  
**for  $U^{-1/p}$  with  $p = 3/2$**



**Plots of the uncentered random walk**  
**for  $U^{-1/p}$  with  $p = 1/2$**

# Heavy Tails

---

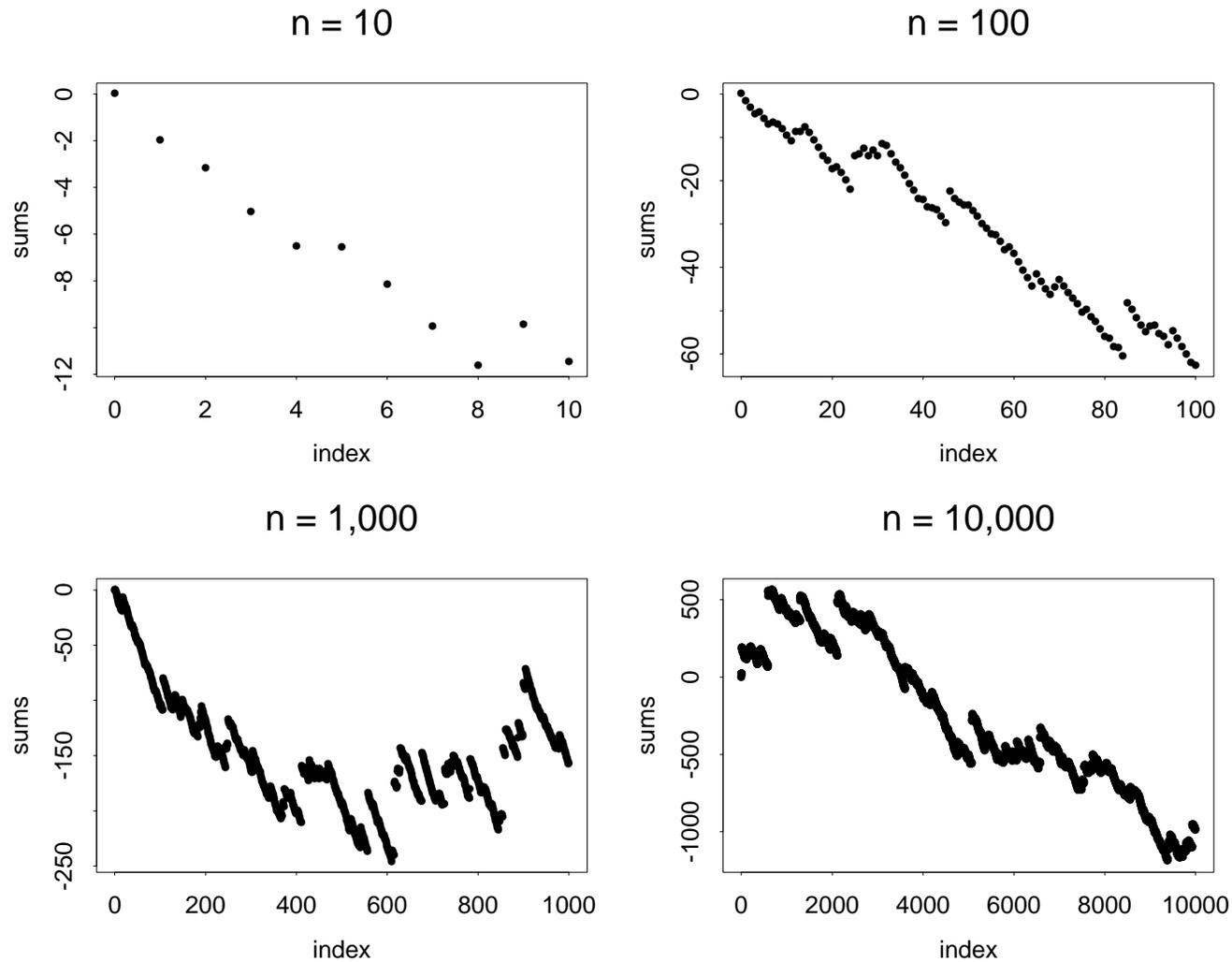
$$\begin{aligned} P(U^{-1/p} > t) &= P(U^{1/p} < t^{-1}). \\ &= P(U < t^{-p}) = t^{-p} \end{aligned}$$

Has **infinite mean** for  $0 < p \leq 1$

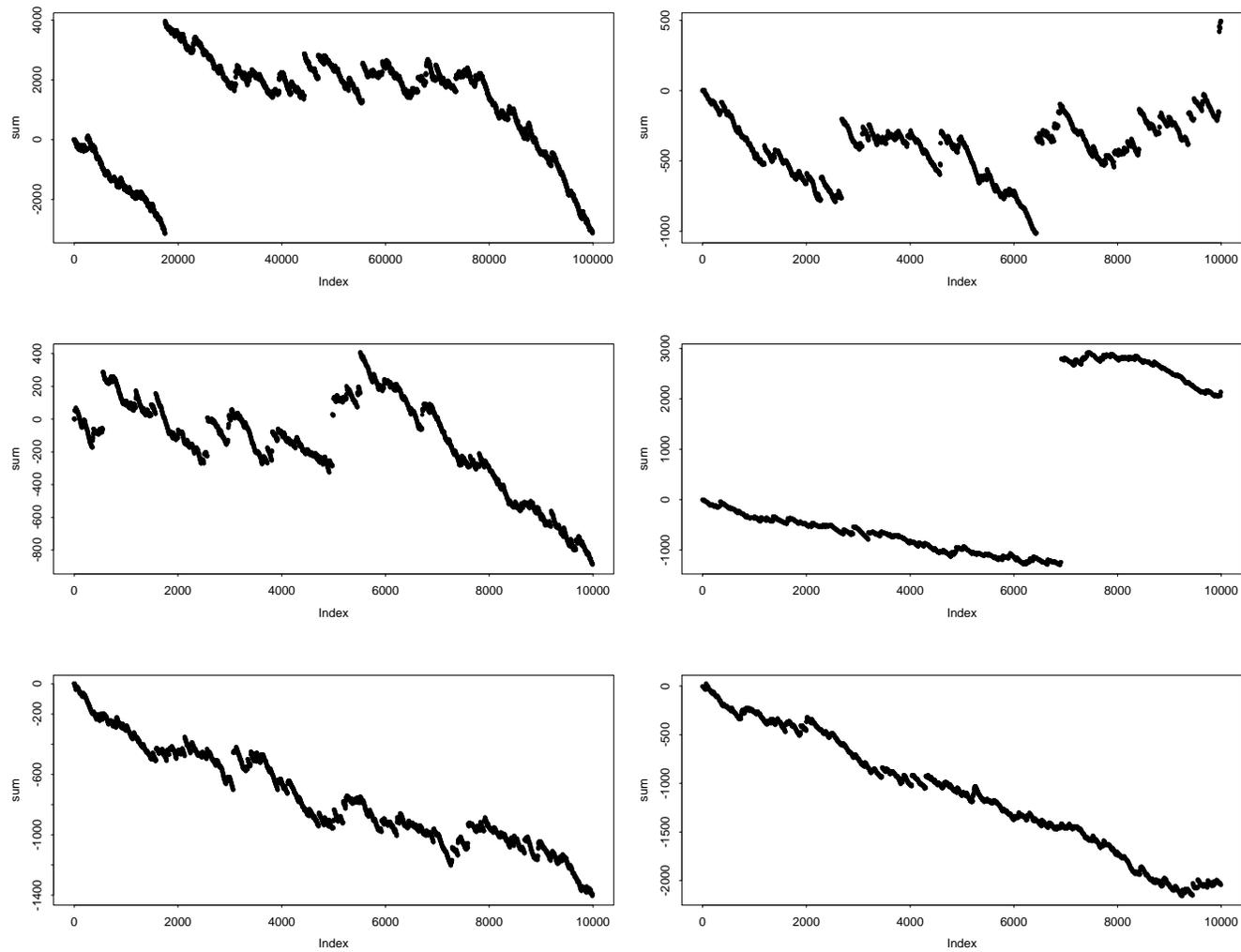
Has **infinite variance** for  $0 < p \leq 2$

# Plots of the Centered Random Walk

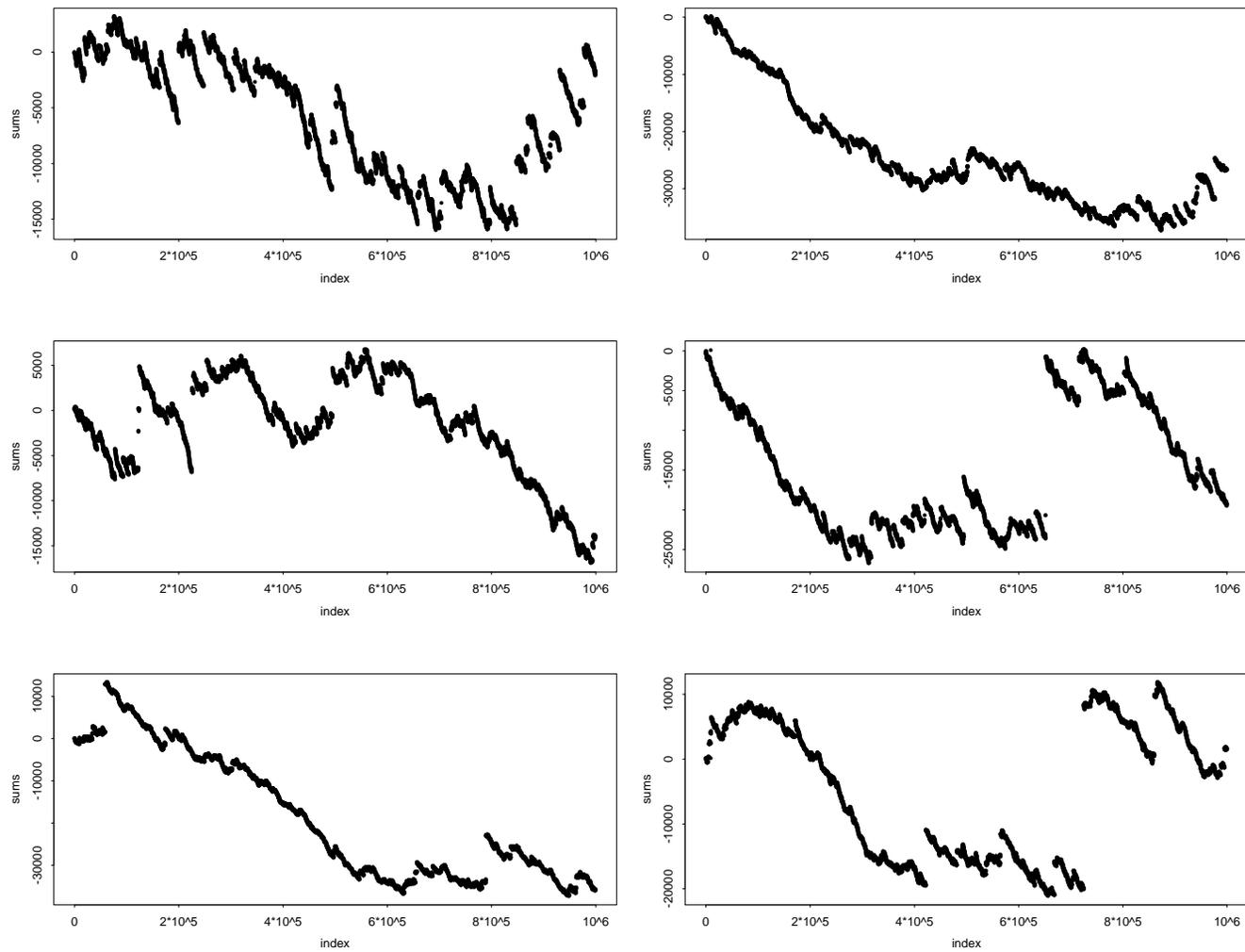
for  $U^{-1/p}$  with  $p = 3/2$



**Plots of the centered random walk**  
**for  $U^{-1/p}$  with  $p = 3/2$**



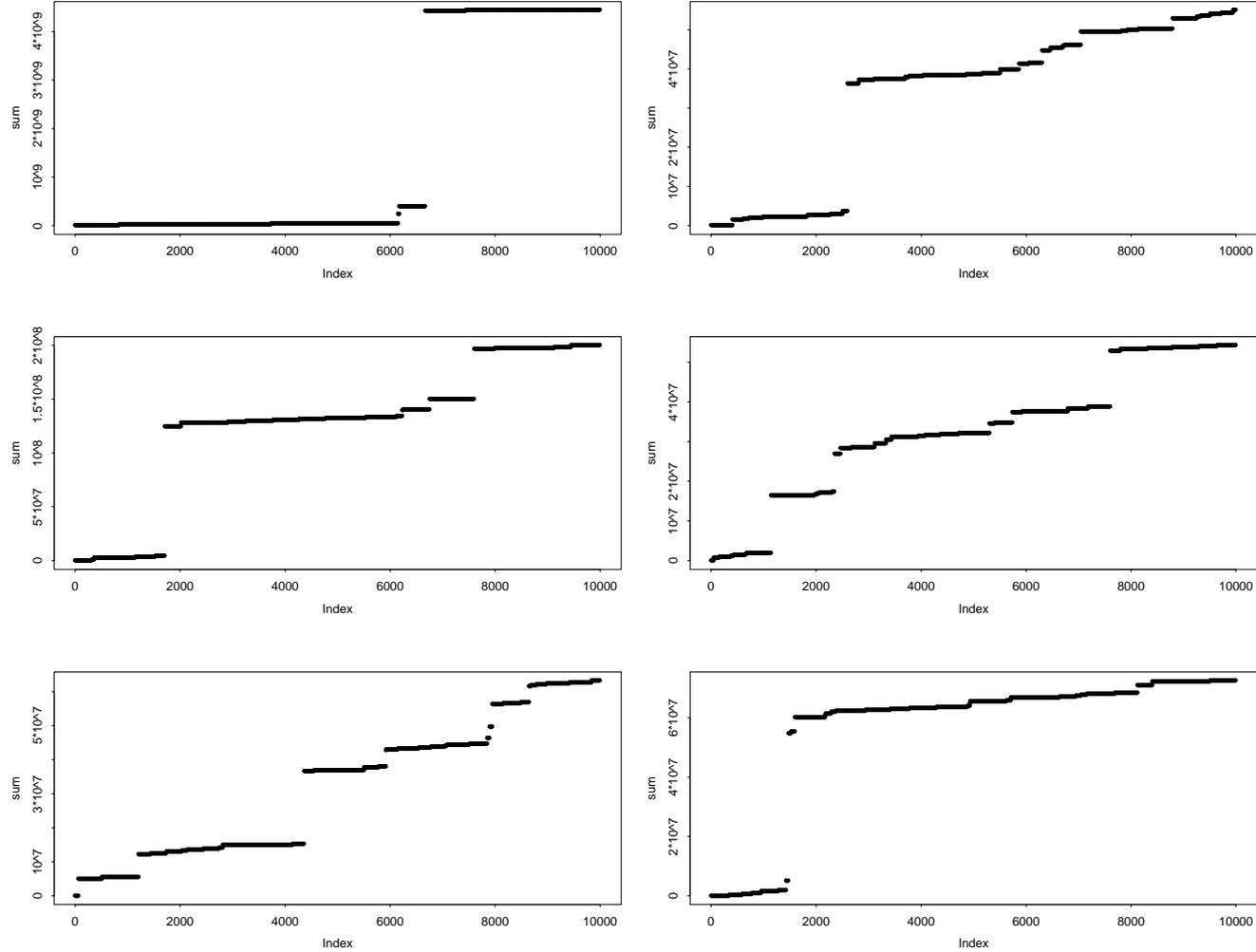
six cases for  $n = 10^4$



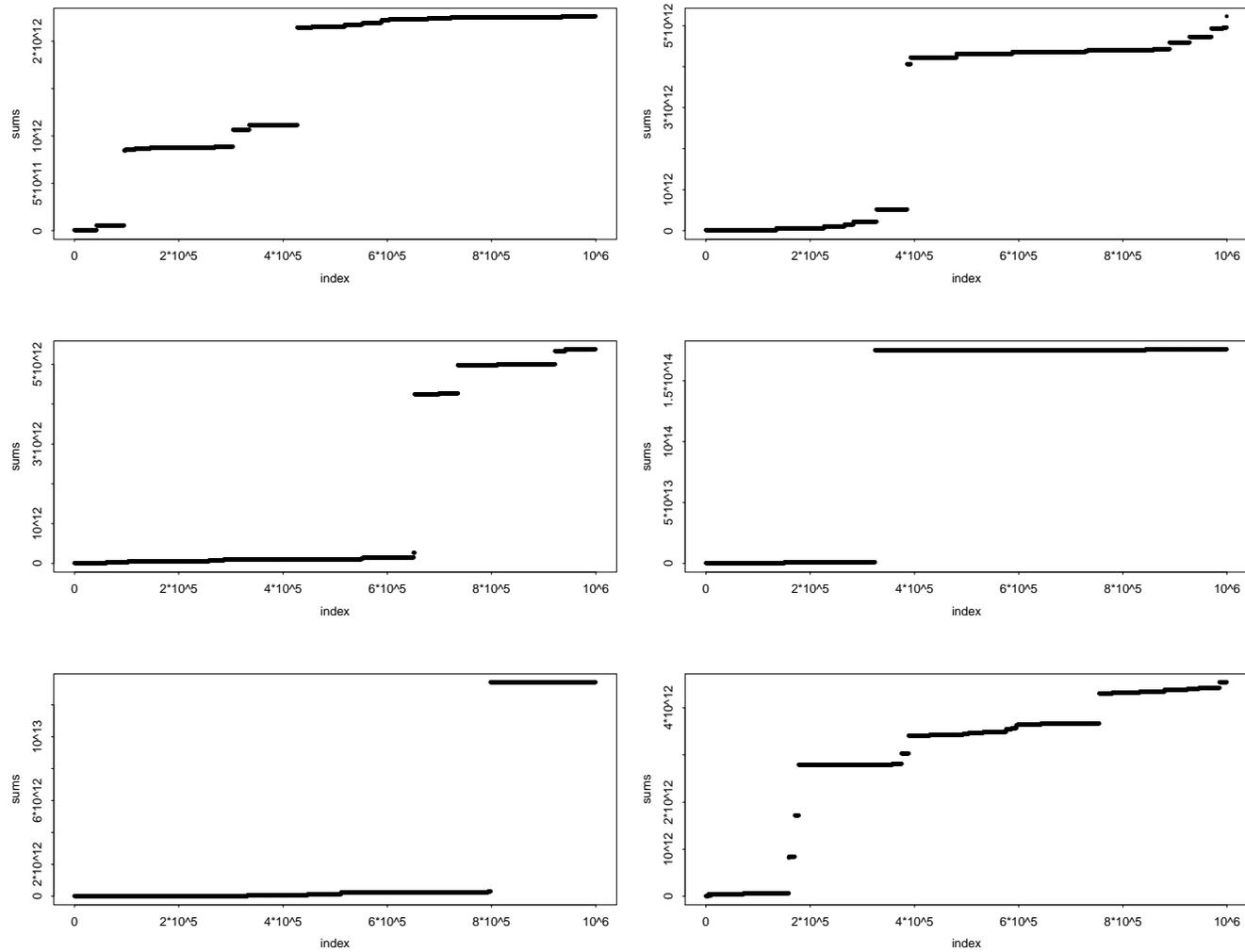
six cases for  $n = 10^6$

# More Plots of the Uncentered Random Walk

for  $U^{-1/p}$  with  $p = 1/2$



six cases for  $n = 10^4$



six cases for  $n = 10^6$

# Conclusions

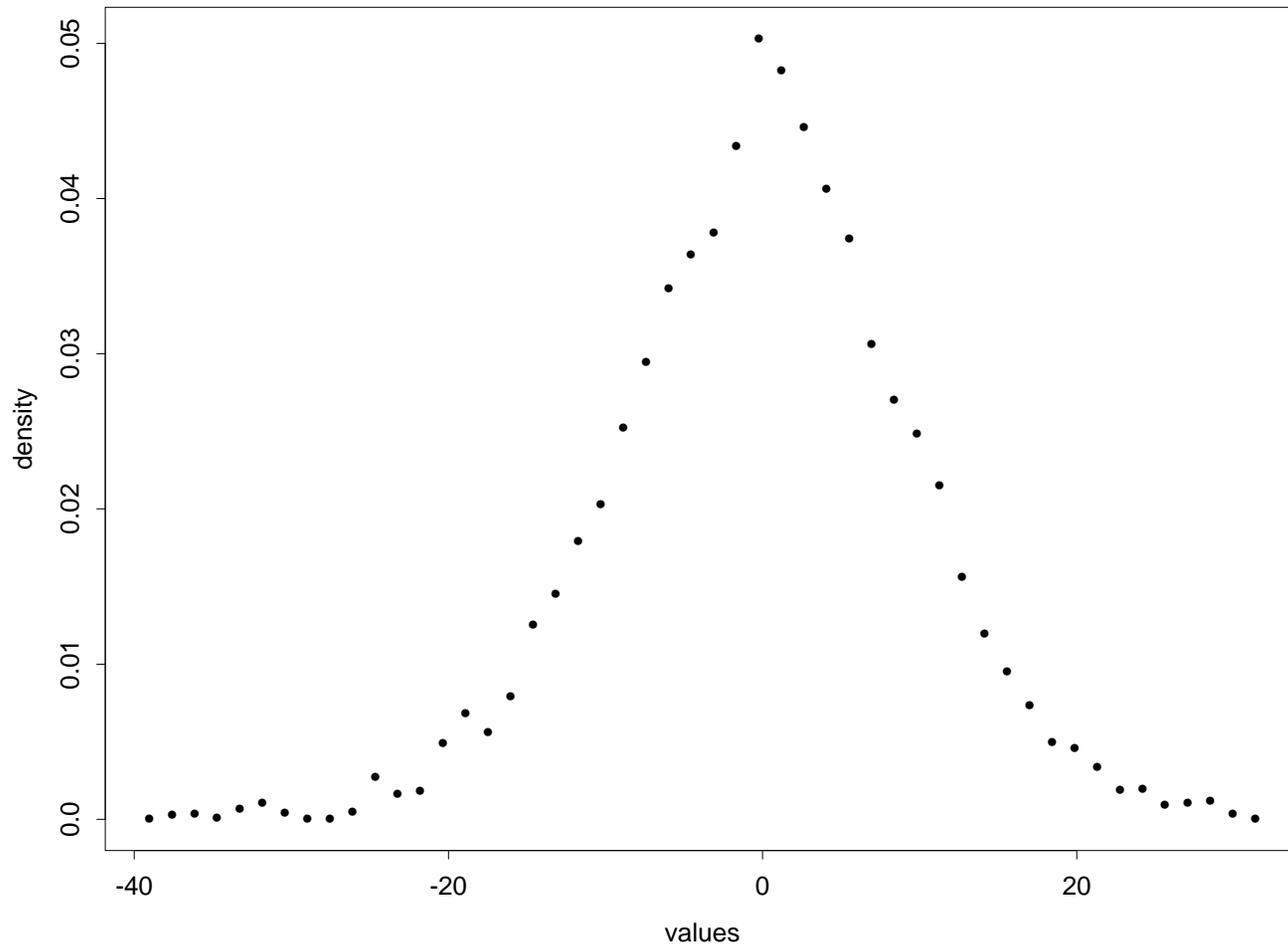
**Plotting reveals statistical regularity.**

**Stochastic-process limits explain the statistical regularity.**

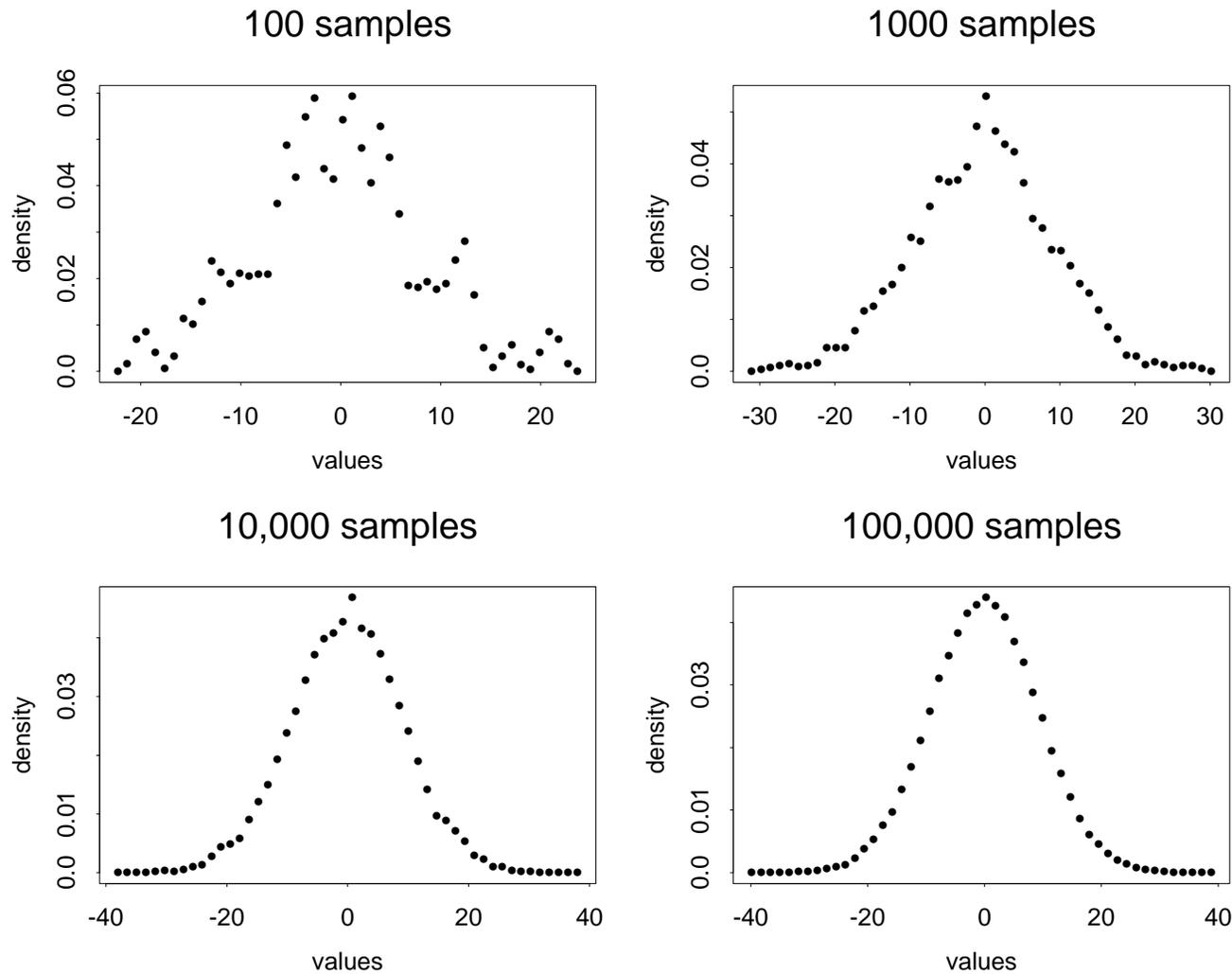
# Density Estimates

for last partial sum  $S_n$

density estimation based on 1000 samples

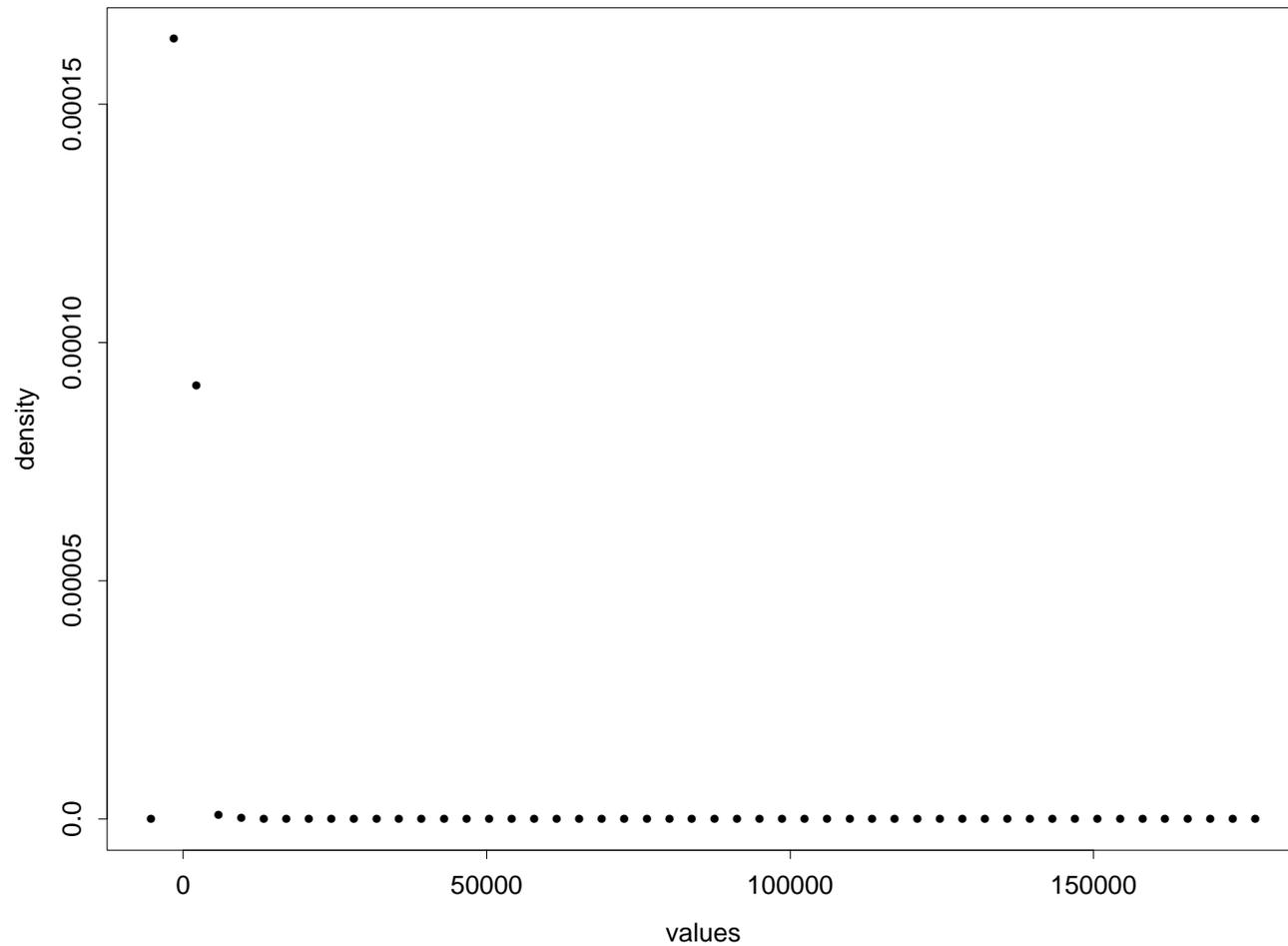


density estimate for  $S_n - mn$  with uniformly distributed summands



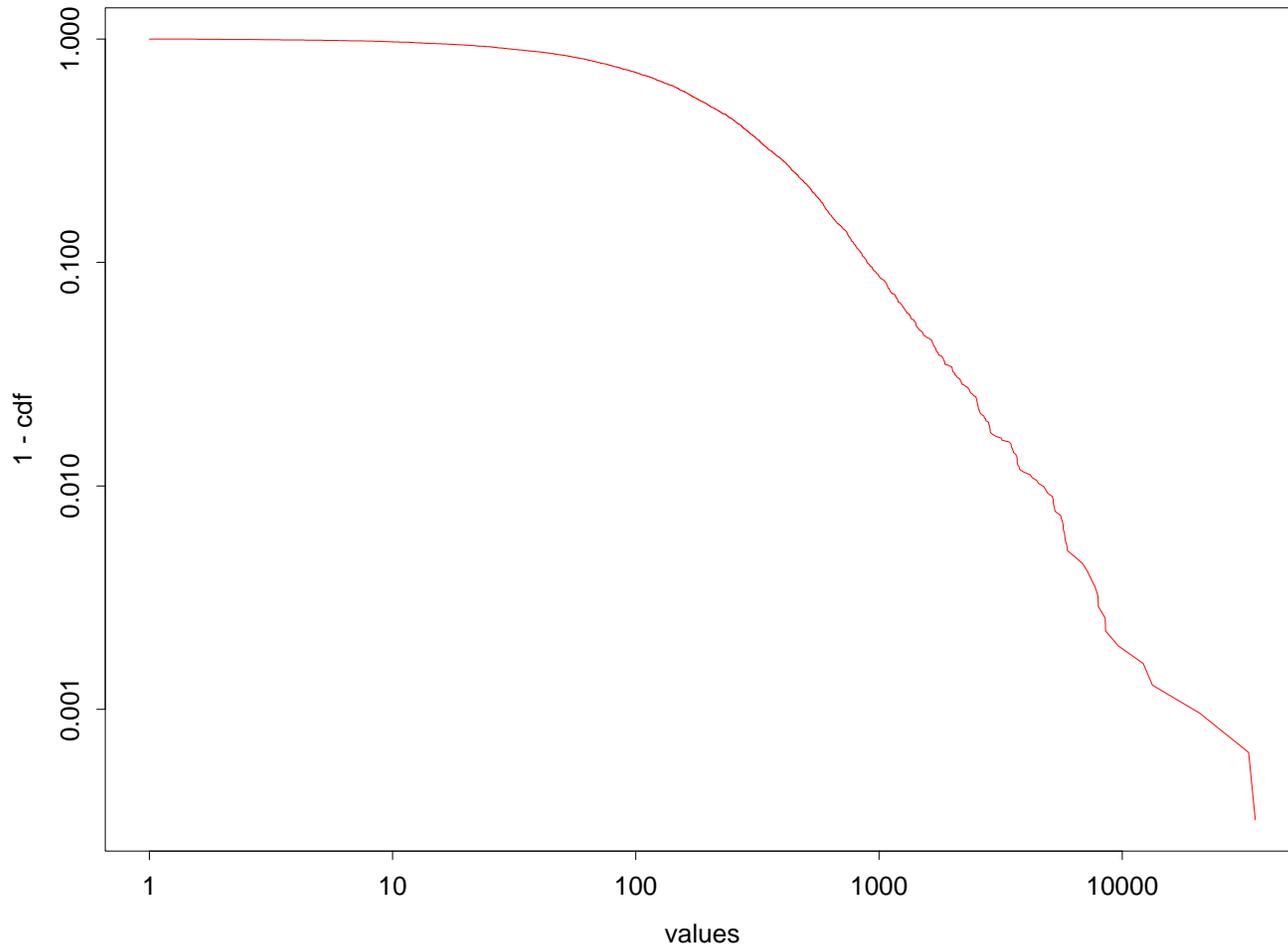
density estimates for  $S_n - mn$  with uniformly distributed summands

density estimate



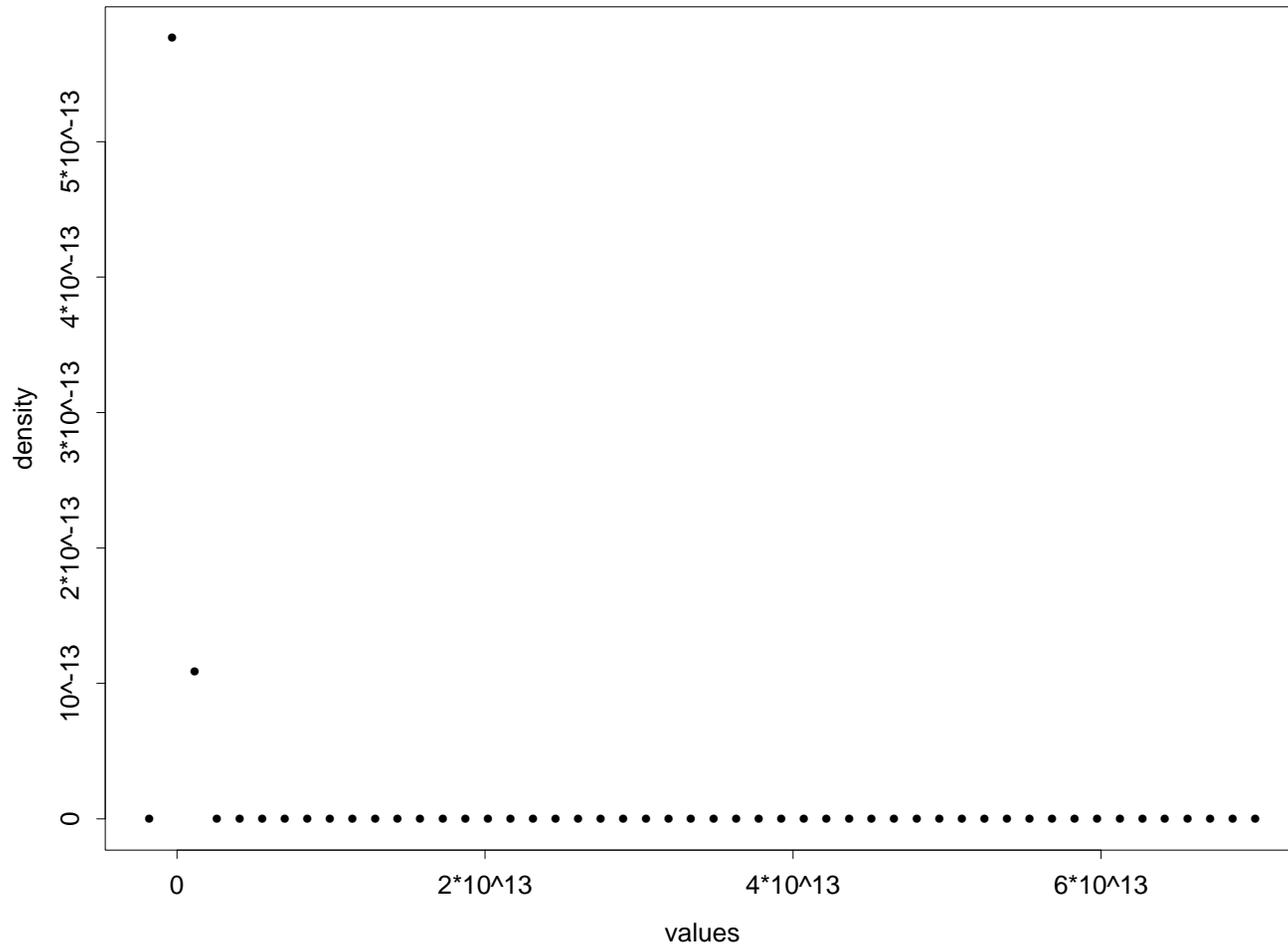
density estimate for  $S_n - mn$  with summands  
 $U^{-1/p}$  for  $p = 3/2$

tail of empirical cdf in log-log scale



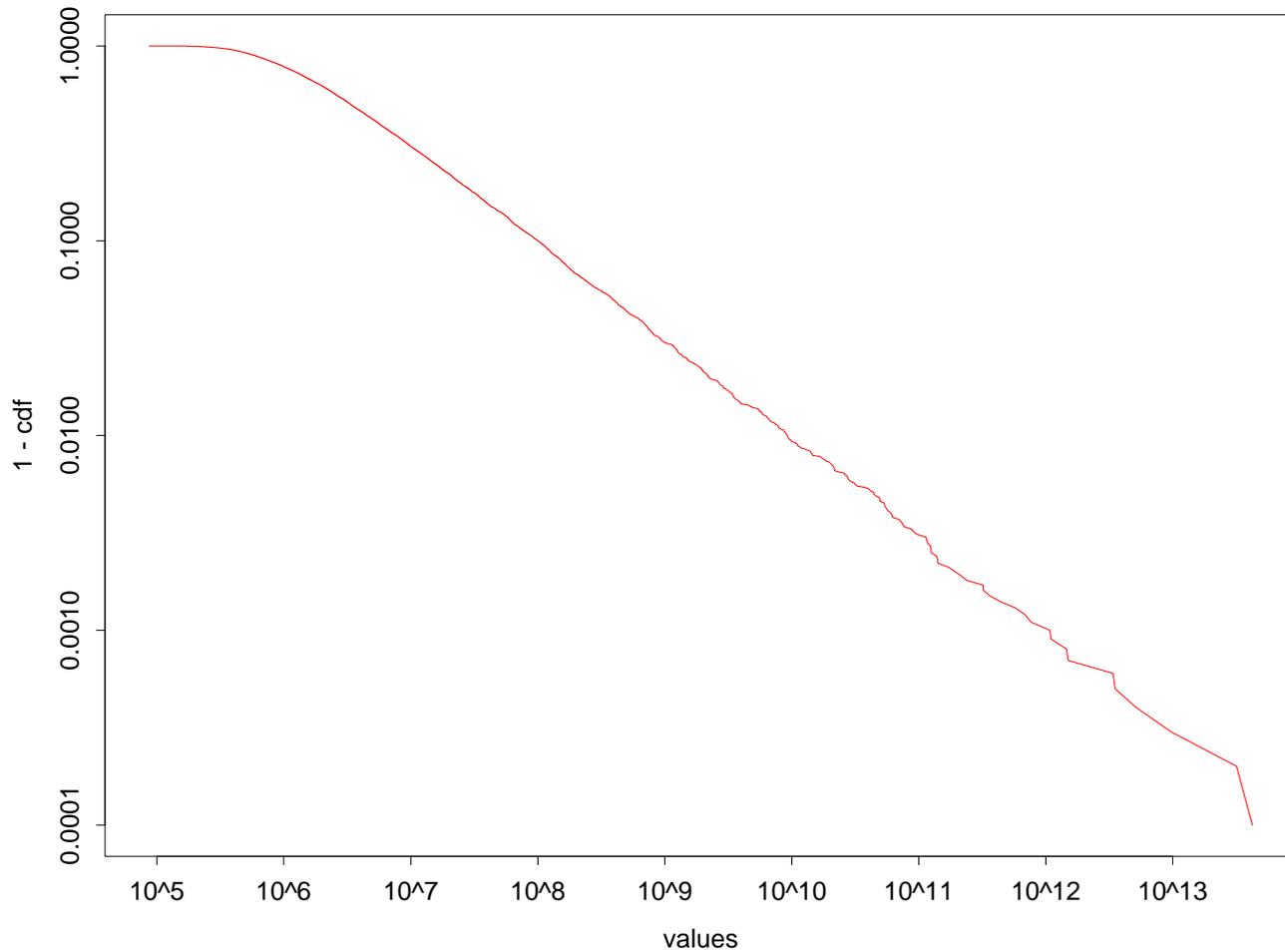
tail of empirical cdf of  $S_n - mn$  with  
summands  $U^{-1/p}$  for  $p = 3/2$

density estimate for  $p = 1/2$



density estimate for  $S_n - mn$  with summands  
 $U^{-1/p}$  for  $p = 1/2$

tail of empirical cdf for  $p = 1/2$  in log-log scale



tail of empirical cdf of  $S_n - mn$  with  
summands  $U^{-1/p}$  for  $p = 1/2$