Longest Increasing Subsequences of a Random Permutation

First moment method, Erdős–Szekeres, and martingale concentration

Contents

1 Introduction **1**
1.1 The problem . . . 1
1.2 Random permutations . . . 2
1.3 Main theorem . . . 2

2 Proof of the Main Theorem **3**
2.1 Tools . . . 3
2.2 Lower bound: the Erdős–Szekeres theorem . . . 3
2.3 Upper bound: the first moment method . . . 4
2.4 Concentration via Doob's martingale . . . 5

3 Simulations **6**
3.1 Computing the longest increasing subsequence . . . 6
3.2 A random instance . . . 6
3.3 Growth of the mean . . . 6
3.4 The first moment is far from the truth . . . 7
3.5 Fluctuations . . . 8
3.6 Code for the figures . . . 8

References **9**

1 Introduction

These notes have been prepared with Claude AI Fable 5.1 on High effort. Siva Athreya takes responsbility of all the content. The length of the longest increasing subsequencs of a random permutaiton is a classical question and was originally posed by Ulam in 1961. It is well studied in the literature.

Definition 1.1. Let $\pi = (\pi(1), \pi(2), \ldots, \pi(n))$ be a permutation of $\{1, 2, \ldots, n\}$. An *increasing subsequence* of π of length k is a set of indices

$$i_1 < i_2 < \cdots < i_k \qquad \text{with} \qquad \pi(i_1) < \pi(i_2) < \cdots < \pi(i_k).$$

We write $L_n(\pi)$ for the length of the longest increasing subsequence of π, and $D_n(\pi)$ for the length of the longest decreasing subsequence (defined analogously, with the inequalities between the $\pi(i_j)$ reversed).

Example 1.2. Let $n = 8$ and $\pi = (3, 1, 4, 7, 5, 8, 2, 6)$. Then $(1, 4, 5, 6)$, $(3, 4, 5, 8)$ and $(1, 4, 7, 8)$ are increasing subsequences of length 4, and one checks that there is no increasing subsequence of length 5. So $L_8(\pi) = 4$. The subsequence $(7, 5, 2)$ is decreasing and $D_8(\pi) = 3$.

More generally, for any sequence $x_1, \ldots, x_n$ of *distinct* real numbers we define $L(x_1, \ldots, x_n)$ and $D(x_1, \ldots, x_n)$ in the same way; only the relative order of the x_i matters.

1.1 Random permutations

Let Π be a uniformly chosen random permutation of $\{1, \dots, n\}$, that is,

$$\mathbb{P}(\Pi = \pi) = \frac{1}{n!} \qquad \text{for every permutation } \pi.$$

It will be convenient to realise Π using i.i.d. Uniform$(0,1)$ random variables. Let $X_1, X_2, \dots$ be i.i.d. Uniform$(0,1)$ random variables and set

$$\Pi(i) := \text{ rank of } X_i \text{ among } X_1, \dots, X_n, \qquad 1 \le i \le n.$$

Exercise 1.3. Show that each of the $n!$ relative orderings of $X_1, \dots, X_n$ is equally likely.

Using Exercise 1.3, Π is a uniform random permutation and

$$L_n := L_n(\Pi) = L(X_1, \dots, X_n).$$

Ulam's problem (1961). How large is L_n for large n?

Simulations (Section 3) suggest that $\mathbb{E}[L_n] \approx 2\sqrt{n}$ and that L_n is sharply concentrated around its mean. The goal of these notes

1.2 Main theorem

The main result discussed in class is the following theorem.

Theorem 1.4. *Let L_n be the length of the longest increasing subsequence of a uniform random permutation of $\{1, \dots, n\}$.*

(a) (Lower bound) *For every $n \ge 1$,*

$$\mathbb{E}[L_n] \ge \sqrt{n}.$$

(b) (Upper bound) *For every $\delta > 0$ and every $k \ge (1+\delta)e\sqrt{n}$,*

$$\mathbb{P}(L_n \ge k) \le (1+\delta)^{-2k},$$

and consequently

$$\limsup_{n\to\infty} \frac{\mathbb{E}[L_n]}{\sqrt{n}} \le e.$$

(c) (Concentration) *For every $n \ge 1$ and $t > 0$,*

$$\mathbb{P}\big(|L_n - \mathbb{E}[L_n]| \ge t\sqrt{n}\big) \le 2e^{-t^2/2}.$$

We remark briefly on what is known in the literature.

Remark 1.5.

1. Hammersley (1972) showed, by a subadditivity argument, that $L_n/\sqrt{n} \to c$ almost surely for some constant $c \in [\pi/2, e]$. Logan–Shepp (1977) and Vershik–Kerov (1977) independently proved that $c = 2$, so $\mathbb{E}[L_n] \sim 2\sqrt{n}$. The first moment constant e in Theorem 1.4(b) is therefore *not* sharp.

2. Part (c) says the fluctuations of L_n are at most of order $\sqrt{n}$. Baik, Deift and Johansson (1999) proved that

$$\frac{L_n - 2\sqrt{n}}{n^{1/6}} \Rightarrow \text{Tracy–Widom distribution},$$

the same limit law as the (rescaled) largest eigenvalue of a large random Hermitian matrix from the Gaussian Unitary Ensemble. The simulations in Section 3 show the $n^{1/6}$ scaling clearly.

2 Proof of the Main Theorem

2.1 Tools

The first three results have been seen in Probability III class and earlier as well.

Proposition 2.1 (Markov's inequality)**.** *If $Y \geq 0$ and $a > 0$ then $\mathbb{P}(Y \geq a) \leq \mathbb{E}[Y]/a$.*

Corollary 2.2 (First moment method)**.** *Let N be a random variable taking values in $\{0, 1, 2, \dots\}$. Then*

$$\mathbb{P}(N \geq 1) \leq \mathbb{E}[N].$$

Proposition 2.3 (Jensen's inequality)**.** *If φ is convex and Y, $\varphi(Y)$ are integrable then $\varphi(\mathbb{E}[Y]) \leq \mathbb{E}[\varphi(Y)]$.*

The next result is a powerful tool in showing concentration inequalities.

Theorem 2.4 (Azuma–Hoeffding inequality)**.** *Let $\{M_k\}_{0 \leq k \leq n}$ be a martingale. Let the observable events till time k be denoted by $\{\mathcal{A}_k\}_{0 \leq k \leq n}$, and suppose $|M_k - M_{k-1}| \leq c_k$ almost surely for $1 \leq k \leq n$. Then for every $t > 0$,*

$$\mathbb{P}\big(|M_n - M_0| \geq t\big) \leq 2 \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right).$$

Proof of Theorem 2.4. Let $D_k = M_k - M_{k-1}$, so that $M_n - M_0 = \sum_{k=1}^{n} D_k$, $\mathbb{E}[D_k \mid \mathcal{A}_{k-1}] = 0$ and $|D_k| \leq c_k$. Fix $\lambda > 0$. Since $\exp(\lambda \sum_{k=1}^{n-1} D_k)$ is observable by time $n-1$, the tower property and Homework 11 Question 2 (a) give

$$\mathbb{E}\big[e^{\lambda(M_n - M_0)}\big] \;=\; \mathbb{E}\Big[e^{\lambda \sum_{k=1}^{n-1} D_k}\, \mathbb{E}\big[e^{\lambda D_n} \mid \mathcal{A}_{n-1}\big]\Big] \;\leq\; e^{\lambda^2 c_n^2/2}\, \mathbb{E}\big[e^{\lambda \sum_{k=1}^{n-1} D_k}\big].$$

Iterating, i.e. peeling off $D_{n-1}, D_{n-2}, \dots, D_1$ in turn,

$$\mathbb{E}\big[e^{\lambda(M_n - M_0)}\big] \leq \exp\left(\frac{\lambda^2}{2} \sum_{k=1}^{n} c_k^2\right) =: e^{\lambda^2 \sigma^2/2}.$$

By Markov's inequality applied to the nonnegative random variable $e^{\lambda(M_n - M_0)}$,

$$\mathbb{P}(M_n - M_0 \geq t) = \mathbb{P}\big(e^{\lambda(M_n - M_0)} \geq e^{\lambda t}\big) \leq e^{-\lambda t}\, \mathbb{E}\big[e^{\lambda(M_n - M_0)}\big] \leq e^{-\lambda t + \lambda^2 \sigma^2/2}.$$

The choice $\lambda = t/\sigma^2$ minimises the exponent and yields $\mathbb{P}(M_n - M_0 \geq t) \leq e^{-t^2/(2\sigma^2)}$. Applying the same bound to the martingale $\{-M_k\}$ controls the lower tail, and the union of the two tails gives the factor 2. □

Recall that the exponentiate-then-Markov inequality step is the *Chernoff bound* technique.

2.2 Lower bound: the Erdős–Szekeres theorem

Theorem 2.5 (Erdős–Szekeres, 1935)**.** *For every sequence $x_1, \dots, x_n$ of distinct real numbers,*

$$L(x_1, \dots, x_n) \cdot D(x_1, \dots, x_n) \geq n.$$

In particular $\max(L, D) \geq \sqrt{n}$ for every such sequence.

Proof. Write $L = L(x_1, \dots, x_n)$ and $D = D(x_1, \dots, x_n)$. For each index i let

- $\ell_i :=$ the length of the longest increasing subsequence *ending at* x_i,

- $d_i :=$ the length of the longest decreasing subsequence *ending at* x_i.

Clearly $1 \le \ell_i \le L$ and $1 \le d_i \le D$. We claim that the map $i \mapsto (\ell_i, d_i)$ is injective. Let $i < j$. If $x_i < x_j$, then appending x_j to a longest increasing subsequence ending at x_i produces an increasing subsequence ending at x_j, so $\ell_j \ge \ell_i + 1$. If $x_i > x_j$, the same argument with decreasing subsequences gives $d_j \ge d_i + 1$. In either case $(\ell_i, d_i) \ne (\ell_j, d_j)$. Thus the n indices are mapped injectively into $\{1, \dots, L\} \times \{1, \dots, D\}$, and $n \le LD$. □

In Example 1.2, $L \cdot D = 4 \cdot 3 = 12 \ge 8$. Exercise 2.6 below shows that the bound is sharp.

Exercise 2.6. Let $n = ab$ with $a, b \ge 1$. Construct a permutation of $\{1, \dots, n\}$ with $L = a$ and $D = b$.

Exercise 2.7. (a) Show carefully that $D_n \overset{d}{=} L_n$. (b) Find the distribution of L_3, compute $\mathbb{E}[L_3]$ and verify that $\mathbb{E}[L_3] \ge \sqrt{3}$.

Proof of Theorem 1.4(a). First, $D_n \overset{d}{=} L_n$ by Exercise 2.7. The map $(X_1, \dots, X_n) \mapsto (1 - X_1, \dots, 1 - X_n)$ preserves the joint law of the X_i and turns increasing subsequences into decreasing ones, so $D(X_1, \dots, X_n) = L(1 - X_1, \dots, 1 - X_n) \overset{d}{=} L(X_1, \dots, X_n)$.

By Theorem 2.5, $L_n \ge n/D_n$. Since $y \mapsto 1/y$ is convex on $(0, \infty)$, Jensen's inequality gives

$$\mathbb{E}[L_n] \ge \mathbb{E}\Big[\frac{n}{D_n}\Big] \ge \frac{n}{\mathbb{E}[D_n]} = \frac{n}{\mathbb{E}[L_n]}.$$

Hence $\mathbb{E}[L_n]^2 \ge n$. □

2.3 Upper bound: the first moment method

Fix $k \ge 1$ and let

$$N_k := \text{ the number of increasing subsequences of } X_1, \dots, X_n \text{ of length } k,$$

that is,

$$N_k = \sum_{I \subset \{1, \dots, n\},\ |I| = k} 1\{X \text{ is increasing on } I\}.$$

For a fixed $I = \{i_1 < \cdots < i_k\}$ the $k!$ relative orderings of $X_{i_1}, \dots, X_{i_k}$ are equally likely, so $\mathbb{P}(X_{i_1} < \cdots < X_{i_k}) = 1/k!$, and by linearity

$$\mathbb{E}[N_k] = \binom{n}{k} \frac{1}{k!}. \tag{1}$$

The key observation is that $\{L_n \ge k\} = \{N_k \ge 1\}$.

Proof of Theorem 1.4(b). By Corollary 2.2 and (1),

$$\mathbb{P}(L_n \ge k) = \mathbb{P}(N_k \ge 1) \le \mathbb{E}[N_k] = \binom{n}{k} \frac{1}{k!}.$$

We use $\binom{n}{k} \le n^k/k!$ and $k! \ge (k/e)^k$, the latter because $e^k = \sum_{j \ge 0} k^j/j! \ge k^k/k!$. Therefore

$$\binom{n}{k} \frac{1}{k!} \le \frac{n^k}{(k!)^2} \le \frac{n^k e^{2k}}{k^{2k}} = \left(\frac{e\sqrt{n}}{k}\right)^{2k}.$$

If $k \ge (1 + \delta) e\sqrt{n}$ then $e\sqrt{n}/k \le (1 + \delta)^{-1}$, which gives $\mathbb{P}(L_n \ge k) \le (1 + \delta)^{-2k}$. Note that this is $\le \exp(-2(1 + \delta) e\sqrt{n} \log(1 + \delta))$, exponentially small in $\sqrt{n}$.

For the expectation, let $k_0 = \lceil (1+\delta)e\sqrt{n} \rceil$. Since L_n takes values in $\{0, 1, 2, \dots\}$,

$$\begin{aligned} \mathbb{E}[L_n] &= \sum_{k \geq 1} \mathbb{P}(L_n \geq k) \;\leq\; (k_0 - 1) + \sum_{k \geq k_0} (1+\delta)^{-2k} \\ &\leq (1+\delta)e\sqrt{n} + \frac{1}{1 - (1+\delta)^{-2}}. \end{aligned}$$

Dividing by $\sqrt{n}$, letting $n \to \infty$ and then $\delta \downarrow 0$ gives $\limsup_n \mathbb{E}[L_n]/\sqrt{n} \leq e$. □

2.4 Concentration via Doob's martingale

The concentration statement rests on one simple combinatorial fact. Recall $L_n = L(X_1, \dots, X_n)$ with $L : (0,1)^n \to \mathbb{N}$.

Lemma 2.8 (Lipschitz property)**.** *If $x, x' \in (0,1)^n$ differ in exactly one coordinate then $|L(x) - L(x')| \leq 1$.*

Proof. Suppose $x_i \neq x'_i$ and $x_j = x'_j$ for $j \neq i$. Take a longest increasing subsequence of x and delete the index i from it if it is present. What remains is an increasing subsequence of x' of length at least $L(x) - 1$, so $L(x') \geq L(x) - 1$. By symmetry $L(x) \geq L(x') - 1$. □

Nothing about the uniform distribution was used: changing one data point changes the length of the longest increasing subsequence by at most one.

Proof of Theorem 1.4(c). Let $\mathcal{A}_k = \sigma(X_1, \dots, X_k)$ for $0 \leq k \leq n$ ($\mathcal{A}_0$ trivial) and define the Doob martingale

$$M_k := \mathbb{E}[L_n \mid \mathcal{A}_k] = g_k(X_1, \dots, X_k), \qquad g_k(x_1, \dots, x_k) := \mathbb{E}\big[L(x_1, \dots, x_k, X_{k+1}, \dots, X_n)\big].$$

The second expression follows from the independence of $(X_{k+1}, \dots, X_n)$ from $\mathcal{A}_k$: one simply integrates out the future coordinates. Then $M_0 = \mathbb{E}[L_n]$, $M_n = L_n$, and by the tower property

$$\mathbb{E}[M_{k+1} \mid \mathcal{A}_k] = \mathbb{E}\big[\mathbb{E}[L_n \mid \mathcal{A}_{k+1}] \mid \mathcal{A}_k\big] = \mathbb{E}[L_n \mid \mathcal{A}_k] = M_k,$$

so $\{M_k\}$ is a martingale.

Bounded differences. Let X'_k be an independent copy of X_k, independent of everything else, and write $X = (X_1, \dots, X_n)$ and $X^{(k)}$ for X with the k-th coordinate replaced by X'_k. Since $(X'_k, X_{k+1}, \dots, X_n)$ is independent of $\mathcal{A}_k$ and has the same law as $(X_k, X_{k+1}, \dots, X_n)$,

$$M_{k-1} = g_{k-1}(X_1, \dots, X_{k-1}) = \mathbb{E}\big[L(X^{(k)}) \mid \mathcal{A}_k\big].$$

Hence, by Lemma 2.8,

$$|M_k - M_{k-1}| = \big|\mathbb{E}\big[L(X) - L(X^{(k)}) \mid \mathcal{A}_k\big]\big| \leq \mathbb{E}\big[|L(X) - L(X^{(k)})| \mid \mathcal{A}_k\big] \leq 1.$$

Theorem 2.4 with $c_k \equiv 1$ gives $\mathbb{P}(|L_n - \mathbb{E}[L_n]| \geq t) \leq 2e^{-t^2/(2n)}$; replace t by $t\sqrt{n}$. □

Remark 2.9. The argument proves a general fact, often called McDiarmid's inequality or the method of bounded differences: if $Y_1, \dots, Y_n$ are independent and f changes by at most c_k when the k-th coordinate is changed, then $f(Y_1, \dots, Y_n)$ is concentrated around its mean with Gaussian tails of variance proxy $\sum_k c_k^2$. See Exercise 2.10.

Exercise 2.10 (Bounded differences)**.** Let $Y = (Y_1, \dots, Y_n)$ have independent coordinates and let f be a function such that $|f(y) - f(y')| \leq c_k$ whenever y, y' differ only in the k-th coordinate. Show that for every $t > 0$,

$$\mathbb{P}\big(|f(Y) - \mathbb{E}f(Y)| \geq t\big) \leq 2\exp\Big(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\Big).$$

Exercise 2.11. Show that

$$\frac{L_n - \mathbb{E}[L_n]}{\sqrt{n}} \to 0 \qquad \text{with probability1.}$$

(Hint: take $t = 2\sqrt{\log n}$ in Theorem 1.4(c).)

3 Simulations

All simulations were done in R. The random seed was `set.seed(2026)`.

3.1 Computing the longest increasing subsequence

The following function computes $L(x_1, \ldots, x_n)$ in $O(n \log n)$ time by *patience sorting*: the values are dealt one at a time onto piles, each value going onto the leftmost pile whose top exceeds it (or starting a new pile on the right). The pile tops are then always increasing from left to right, the number of piles equals L, and following the back-pointers `prev` from the top of the last pile recovers one longest increasing subsequence.

```
lis <- function(x){
  n <- length(x); tops <- numeric(0); topidx <- integer(0)
  prev <- integer(n)
  for(i in 1:n){
    j <- findInterval(x[i], tops) + 1    # leftmost pile with top > x[i]
    tops[j] <- x[i]; topidx[j] <- i
    prev[i] <- if(j > 1) topidx[j-1] else 0
  }
  L <- length(tops); idx <- integer(L); k <- topidx[L]
  for(m in L:1){ idx[m] <- k; k <- prev[k] }
  list(length = L, index = idx)
}
```

On the permutation of Example 1.2:

```
> pi = c(3, 1, 4, 7, 5, 8, 2, 6)
> lis(pi)


$length
[1] 4

$index
[1] 2 3 5 8
```

The indices $2, 3, 5, 8$ correspond to the increasing subsequence $(1, 4, 5, 6)$.

3.2 A random instance

Figure 1 shows $n = 40$ i.i.d. Uniform$(0, 1)$ points (i, X_i) with one longest increasing subsequence highlighted. An increasing subsequence is a chain of points going up and to the right.

3.3 Growth of the mean

For each $n \in \{100, 400, 1600, 6400\}$ we generated 500 independent uniform random permutations (as i.i.d. uniforms) and recorded L_n; the matrix `sims` has one column per n.

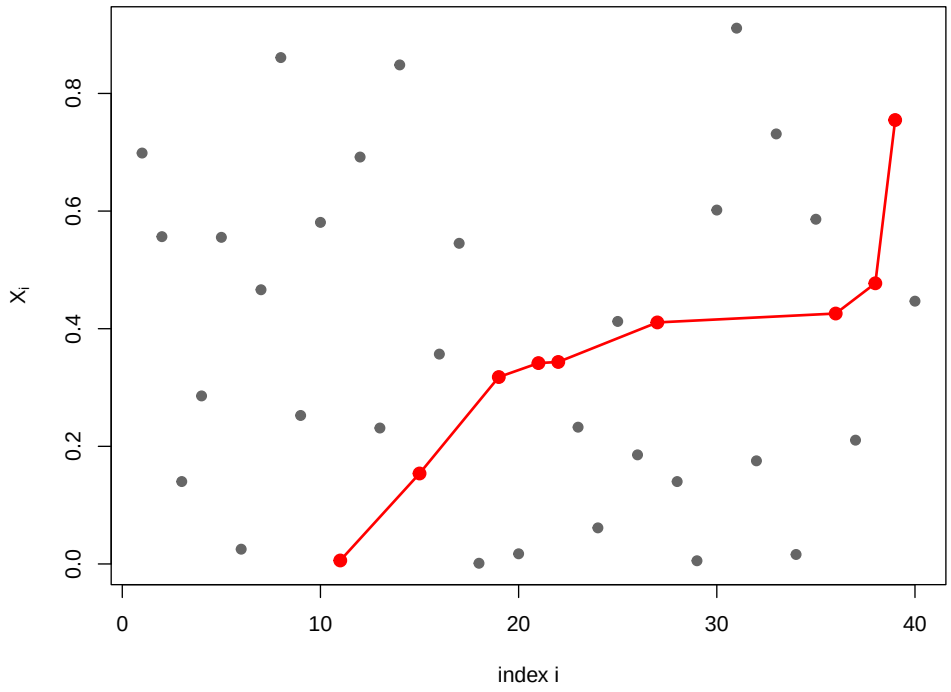


Figure 1: $n = 40$ i.i.d. Uniform$(0, 1)$ points (i, X_i). In red: one longest increasing subsequence, of length 9. Here $2\sqrt{40} \approx 12.6$ and $e\sqrt{40} \approx 17.2$.

```
> ns = c(100, 400, 1600, 6400)
> reps = 500
> sims = sapply(ns, function(n) replicate(reps, lis(runif(n))$length))
> colnames(sims) = ns
> apply(sims, 2, mean)

    100     400    1600    6400
 16.650  35.778  74.270 152.634

> apply(sims, 2, mean)/sqrt(ns)

     100      400     1600     6400
1.665000 1.788900 1.856750 1.907925
```

Each time n grows by a factor 4 the sample mean roughly doubles, consistent with $\mathbb{E}[L_n] \approx c\sqrt{n}$. The ratio $\bar{L}_n/\sqrt{n}$ increases slowly towards 2 from below, in agreement with the Logan–Shepp/Vershik–Kerov theorem and the fact that $\mathbb{E}[L_n] = 2\sqrt{n} - O(n^{1/6})$. Figure 2 compares the simulated ratio with the bounds of Theorem 1.4.

3.4 The first moment is far from the truth

For $n = 10000$ and $k = 240$ (between $2\sqrt{n} = 200$ and $e\sqrt{n} \approx 272$) the expected number of increasing subsequences of length k is astronomically large, although $L_n \geq 240$ essentially never happens.

```
> n = 10000; k = 240
> exp(lchoose(n, k) - lfactorial(k))

[1] 3.354606e+21
```

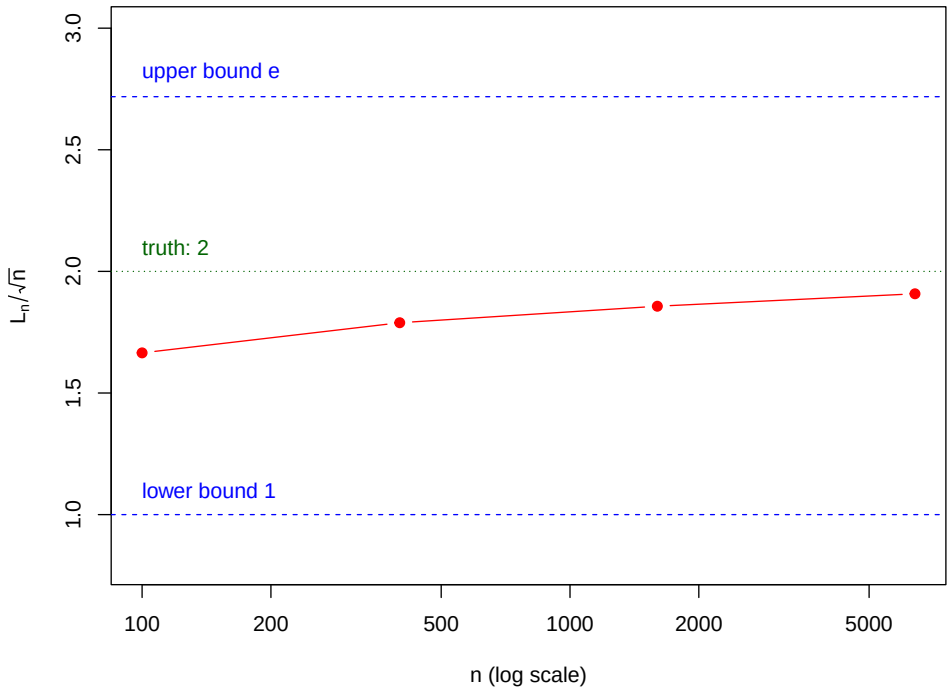


Figure 2: $\bar{L}_n/\sqrt{n}$ from 500 simulations for each n (red), together with the lower bound 1 and upper bound e from Theorem 1.4 (dashed) and the true limit 2 (dotted).

3.5 Fluctuations

Theorem 1.4(c) bounds the fluctuations of L_n by order $\sqrt{n}$. The sample standard deviations are much smaller and scale like $n^{1/6}$, as predicted by the Baik–Deift–Johansson theorem.

```
> s = apply(sims, 2, sd)
> tab = rbind(sd = s, "sd/sqrt(n)" = s/sqrt(ns),
+             "sd/n^(1/6)" = s/ns^(1/6))
> round(tab, 3)
```

```
              100   400  1600  6400
sd          1.613 2.129 2.904 3.680
sd/sqrt(n)  0.161 0.106 0.073 0.046
sd/n^(1/6)  0.749 0.784 0.849 0.854
```

(The standard deviation of the Tracy–Widom distribution is about 0.90, so the last row is heading to the right place.) Figure 3 shows the empirical distribution of L_n for $n = 6400$. Note the skew: the Tracy–Widom distribution is not symmetric, and the sample mean lies to the left of $2\sqrt{n} = 160$.

3.6 Code for the figures

For completeness, the plotting commands:

```
## Figure 1
n = 40; x = runif(n); out = lis(x)
plot(1:n, x, pch=19, col="grey40", xlab="index i", ylab=expression(X[i]))
lines(out$index, x[out$index], col="red", lwd=2)
points(out$index, x[out$index], pch=19, col="red", cex=1.3)
```

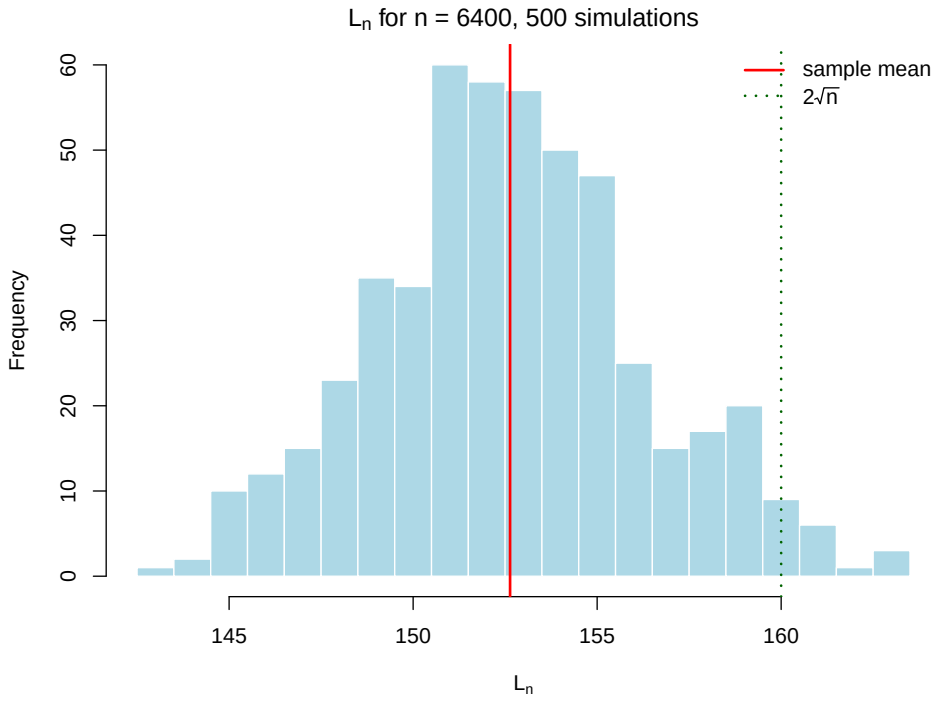


Figure 3: Histogram of L_n for $n = 6400$ over 500 simulations. The solid line is the sample mean and the dotted line is $2\sqrt{n} = 160$.

```
## Figure 2
m = apply(sims, 2, mean)
plot(ns, m/sqrt(ns), log="x", type="b", pch=19, col="red", ylim=c(0.8, 3),
     xlab="n (log scale)", ylab=expression(bar(L)[n]/sqrt(n)))
abline(h=1, lty=2, col="blue"); abline(h=exp(1), lty=2, col="blue")
abline(h=2, lty=3, col="darkgreen")

## Figure 3
n = 6400
hist(sims[, "6400"], breaks=seq(min(sims[,"6400"])-0.5, max(sims[,"6400"])+0.5, by=1),
     col="lightblue", border="white", xlab=expression(L[n]))
abline(v = 2*sqrt(n), col="darkgreen", lwd=2, lty=3)
abline(v = mean(sims[,"6400"]), col="red", lwd=2)
```

References

Textbooks and surveys

[AS] N. Alon and J. H. Spencer, *The Probabilistic Method*, 4th ed., Wiley, 2016. (Chapters 2–4: first and second moment methods; Chapter 7: martingales and bounded differences.)

[Du] R. Durrett, *Probability: Theory and Examples*, 5th ed., Cambridge University Press, 2019. (Chapter 4: martingales.)

[Ro] D. Romik, *The Surprising Mathematics of Longest Increasing Subsequences*, Cambridge University Press, 2015. (A complete, self-contained account of Ulam's problem, including the Baik–Deift–Johansson theorem.)

[St] J. M. Steele, *Probability Theory and Combinatorial Optimization*, CBMS-NSF Regional Conference Series 69, SIAM, 1997. (Chapter 1 treats the longest increasing subsequence exactly as in these notes.)

[AD] D. Aldous and P. Diaconis, Longest increasing subsequences: from patience sorting to the Baik–Deift–Johansson theorem, *Bull. Amer. Math. Soc.* 36 (1999), 413–432.

Original papers

[ES] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* 2 (1935), 463–470.

[Ul] S. M. Ulam, Monte Carlo calculations in problems of mathematical physics, in *Modern Mathematics for the Engineer: Second Series*, McGraw-Hill, 1961, 261–281.

[Ha] J. M. Hammersley, A few seedlings of research, in *Proc. Sixth Berkeley Symp. Math. Statist. Probab.*, Vol. 1, Univ. California Press, 1972, 345–394.

[Ki] J. F. C. Kingman, Subadditive ergodic theory, *Ann. Probab.* 1 (1973), 883–909.

[LS] B. F. Logan and L. A. Shepp, A variational problem for random Young tableaux, *Adv. Math.* 26 (1977), 206–222.

[VK] A. M. Vershik and S. V. Kerov, Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux, *Soviet Math. Dokl.* 18 (1977), 527–531.

[AD2] D. Aldous and P. Diaconis, Hammersley's interacting particle process and longest increasing subsequences, *Probab. Theory Related Fields* 103 (1995), 199–213.

[Fr] A. Frieze, On the length of the longest monotone subsequence in a random permutation, *Ann. Appl. Probab.* 1 (1991), 301–305.

[BB] B. Bollobás and G. Brightwell, The height of a random partial order: concentration of measure, *Ann. Appl. Probab.* 2 (1992), 1009–1018.

[BDJ] J. Baik, P. Deift and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* 12 (1999), 1119–1178.

[Ho] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* 58 (1963), 13–30.

[Az] K. Azuma, Weighted sums of certain dependent random variables, *Tôhoku Math. J.* 19 (1967), 357–367.

[Mc] C. McDiarmid, On the method of bounded differences, in *Surveys in Combinatorics 1989*, London Math. Soc. Lecture Note Ser. 141, Cambridge University Press, 1989, 148–188.