Uniform and Mallows Random Permutations:
Inversions, Levels & Sampling

by

Peter Rabinovitch

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in

Mathematics (Probability and Statistics)

Carleton University
Ottawa, Ontario

© 2012

Peter Rabinovitch

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-89317-3

Our file Notre référence

ISBN: 978-0-494-89317-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Uniformly selected random permutations have been extensively analyzed in the combinatorics and probability literature. Significantly less research has been reported on permutations selected from non-uniform measures. In this thesis, we analyze various characteristics of permutations selected from the Mallows measure: a probability measure on permutations that assigns mass according to the number of inversions of the permutation. In addition, we analyze a new characteristic of the permutation, the maximum element of the inversion table which we call the level. We also develop algorithms for sampling from the Mallows measure, as well as uniformly from all permutations with a fixed number of inversions.

Acknowledgements

It is traditional to acknowledge one's thesis advisor, examination committee, friends and family. And indeed, their guidance, support, patience, questions, and occasional prods were hugely valuable.

I'd also like to take this opportunity to thank the greater mathematical community. Over the many years of my education, I have had many more informal teachers than those at the head of the classroom. I have been extremely fortunate to have spent time, both real and virtual, with mathematicians across the world at conferences, in email exchanges and blog postings. These exchanges, while each small, have had a large cumulative effect. And the community that results is as close to a utopia as I can imagine.

To the community, I say a very deeply felt 'thank you.'

This thesis is dedicated to the memory of Amit Bose: teacher, colleague and most
of all, friend.

Contents

List of Figures	vi
Chapter 1. Introduction	1
1.1. Origins	1
1.2. Context	3
1.3. Outline	3
Chapter 2. Permutations	5
2.1. Inversions, Inversion Tables and Levels	5
2.2. Sequential Construction	17
Chapter 3. The Uniform Measure on Random Permutations	20
3.1. Uniform Measure	20
3.2. Basic Results	20
3.3. Limits	24
3.4. Sampling	31
3.5. Sequential Construction	32
Chapter 4. The Mallows Measure on Random Permutations	38
4.1. Mallows Measure	38
4.2. Basic Results	41
4.3. Limits	51
4.4. Sampling	68
4.5. Sequential Construction	74
Chapter 5. Applications	78

CONTENTS

v

5.1. Sizing of a Reordering Buffer in a Queueing System	78
5.2. Statistics	82
Chapter 6. Conclusion	86
Bibliography	89
Index	91

List of Figures

2.1.1	Staircase diagram of (3,1,4,5,2)	9
2.1.2	Euler's Pentagonal numbers	10
2.1.3	Balls & bars	12
2.1.4	Staircase diagram truncated at level 3	14
2.1.5	$\psi(20, i, l)$	17
3.3.1	Histograms and limiting density of the inversions and the level	30
3.3.2	Joint histogram of the inversions and the level	31
4.2.1	$E_p[v_{50}]$	44
4.2.2	$V_p[v_{20}]$	45
4.2.3	$E_p[\mathbf{I}]$ for $n=50$	46
4.2.4	Distribution of the level, L , for $n=50$	49
4.2.5	$E_p[\mathbf{L}]$ for $n=50$	51
6.0.1	Three Brownian motions?	87
6.0.2	A particle system representation of (35412)	88

CHAPTER 1

Introduction

Permutations are perhaps the most well-known combinatorial object, and properties of uniformly random permutations have been a focus of intense study for some time, culminating in the discovery of the Tracy-Widom distribution for the scaled length of the longest increasing subsequence, as described in [AD]. Less research has been reported on random permutations that are not selected from the uniform distribution, although two are common: the Ewens distribution (see [P]) and the Mallows distribution [D, Mal], which is a focus of this research. Various characteristics of uniform permutations have been studied, such as the number of cycles, the length of the longest increasing subsequence, and the number of inversions. Herein, we focus on the number of inversions of the permutation, and a new characteristic, which we call the level of the permutation: the maximum element of the inversion table of the permutation.

1.1. Origins

Although very interesting in its own right, this work was originally inspired by some very applied work on the size of a reordering buffer required in a high speed telecommunications network router.

The TCP network protocol used by the Internet (and indeed, most data networks) is by design able to handle packets that arrive out of order. Measurement studies indicate that a small amount of misordering can be handled by the end user's system without affecting performance to a noticeable degree. However, larger amounts of misordering do affect performance, and in addition there is the perception that any misordering is a bad thing. In fact in 2000 the popular networking web site "LightReading" [LR] reported results of a core router test, in which one

of the tests was to see if the routers reordered any packets. They reported that Juniper's router misordered a small amount of packets (on the order of 1%) and Juniper got criticized for this, and received bad press. Juniper later updated the software on the router to prevent any reordering of packets.

There are many reasons why packets may be reordered, but the most likely one is that the packets follow different paths through the network, or even within a node. Within a node, one can mark the packet with a sequence number upon ingress, and then keep the packets in a reordering buffer on egress until all the packet's predecessors have left the node.

The ability to reorder packets within a node would provide more flexibility to the router designer, as paths of execution that do not guarantee ordering of outputs could be considered, widening the possible space of designs. This would allow for potentially better performance, for example, by allowing parallel paths of execution through a network processor rather than sequential processing.

It turns out that if we ignore the temporal aspects of this problem, and just focus on the ordering of the packets, then the size of the reordering buffer needed is precisely the largest element of the inversion table of the permutation representing the reordering of the packets (see Theorem 55). If we assume permutations are selected uniformly at random, then the expected value of the reordering buffer is (for large n) approximately $n - \sqrt{\pi n/2}$, where n is the length of the permutation. However, it seems unlikely that the permutations that would be observed on a real network would be from a uniform distribution, as packets that are nearby (in time) should be more likely to be transposed than distant ones. Hence we are interested in a measure that assigns larger probability to permutations with fewer inversions, proportional to p^i where $p > 0$ is a parameter and i is the number of inversions in the permutation. This is the Mallows measure.

1.2. Context

Reordering in queueing systems seems complicated, as exemplified by [BGP, XT] for example, as well as Section 5.1 of this thesis. The combinatorial approach to reordering taken in this thesis was inspired by [Ba] and [O], as well as of course [D].

Recently a limiting empirical measure of a random Mallows permutation was described by [S], and then the length of the longest increasing subsequence of a Mallows permutation was obtained in [SM], extending the results surveyed in [AD]. Unfortunately, applying the wealth of techniques used in these papers to our problem is not fruitful, as there is no obvious relationship between the level of a permutation and any of the row or column statistics of the Young tableaux generated by the RSK correspondence [F, p 40].

Closer to our work is [Mar], in which the asymptotics of the number of inversions of permutations on n symbols with i inversions, for fixed i are derived. These results were then extended in [LP].

It is well known that the inversions of a permutation are what is known as a Mahonian statistic: equidistributed with the major index of a permutation [Bo, p 53]. However, under the Mallows measure, this is not the case, as a simple example illustrates (see Theorem 35), thus preventing direct translation of results about the major index into the Mallows case.

Many of our results are derived by focusing on the inversion table of the permutation, rather than the permutation itself. This is because the elements of the inversion table are independent (although not identically distributed), making much of the analysis feasible.

1.3. Outline

In order to improve the readability of this thesis, it has been organized by topic. However, this is at the cost of new results being interspersed with old. Thus, in

order to aid the reader in determining what is new, all uncited results are new and due to the author.

In Chapter 2, we review standard results about permutations, definitions and notation. After we introduce the level, we determine an explicit formula as well as the generating function for the number of permutations on n symbols with i inversions and level l .

In Chapter 3, we discuss random permutations, chosen from the uniform distribution. We also derive concentration results for both the inversions and level of a uniformly random permutation.

Chapter 4 parallels Chapter 3, but for the case of Mallows random permutations. Many of the results of the previous chapter appear in a more general form in this chapter, as the uniform measure is a special case of the Mallows measure. For both the inversions and level we derive the mean, variance, asymptotics of the means, and concentration of measure results. In addition, we develop new simulation algorithms for sampling permutations from the Mallows measure, as well as uniformly sampling permutations on n symbols with exactly i inversions.

Chapter 5 discusses the applications of the preceeding chapters to reordering buffer sizing and to estimation of the parameter of the Mallows measure, given the inversions or the level.

Chapter 6 concludes with topics for further investigation.

CHAPTER 2

Permutations

A *permutation* of a set is simply a bijection (1 – 1, onto mapping) of the set to itself. In this thesis, we work only with permutations of the set $[n] = \{1, 2, 3, \dots, n\}$ and we frequently refer to a permutation of $[n]$ as an n -permutation. We will occasionally denote the set of all permutations of $[n]$ by $\mathbb{S}_n$, and other than this piece of notation, we will not make use of any group theoretic properties of permutations.

We write all permutations in one-line notation. For example, the permutation σ of $[4]$ given by

$$\begin{aligned}\sigma(1) &= 4 \\ \sigma(2) &= 3 \\ \sigma(3) &= 1 \\ \sigma(4) &= 2\end{aligned}$$

will usually be denoted by $\sigma = (4, 3, 1, 2)$.

Most of the results in this chapter are available in the literature, for example [Bo] although the focus on the level of the permutation is new, as well as some of the proofs.

2.1. Inversions, Inversion Tables and Levels

One measure of the disorder of a permutation is its number of inversions, which we now define.

DEFINITION 1. For any integer i , $1 \leq i \leq n$, the i^{th} element of the *inversion table* of a permutation σ is the number of elements to the left of σ_i that are greater than σ_i . We denote the inversion table of a permutation σ by $v(\sigma)$, and the i^{th}

component of $v(\sigma)$ by $v_i(\sigma)$ or simply v_i if σ is understood from the context. Thus the inversion table of σ has components

$$v_i = \sum_{1 \leq j \leq i} \mathbb{I}(\sigma_j > \sigma_i)$$

Note that $0 \leq v_i(\sigma) \leq i - 1$.

Note that there are other commonly used definitions of the inversion table that are similar. For example, the Mathematica function *ToInversionVector* returns a list of length $n - 1$ for the inversion table of a permutation of length n , since one component is always zero. We will not be concerned with these other notions here.

Sample Mathematica code to calculate the inversion table for a permutation p follows.

```
PermutationToInversionTable[p_List] := Module[{i, x},
  x = Table[Length[Select[Take[p, i - 1], # > p[[i]] &]], {i, 2, Length[p]}];
  x = Prepend[x, 0];
  Return[x]
]
```

EXAMPLE 2. The permutation $\sigma = (5, 2, 3, 1, 4)$ has inversion table $(0, 1, 1, 3, 1)$.

It is well known that there is a bijection between permutations and their inversion tables.

THEOREM 3. [A1, Ex 1.43] *There is a bijection between permutations and inversion tables.*

PROOF. That there is a unique inversion table for each permutation is obvious from the definition. To find the permutation corresponding to an inversion table is only slightly more complicated, and is perhaps clearest explained with the following Mathematica code.

```
InversionTableToPermutation[v_List] := Module[{n, s, p, k, vt},
```

```

n=Length[v];
s=Range[n,1,-1];
p=Table[0,{n}];
For[k=n,k>=1,k--,
  vt=v[[k]]+1;
  p[[k]]=s[[vt]];
  s=DeleteCases[s,s[[vt]]];
Return[p]]

```

In words, denote the inversion table by $v = (v_1, v_2, \dots, v_n)$, and let $s = (1, 2, \dots, n)$. We start from the n^{th} element of v and work backwards. Since v_n denotes how many elements in the permutation are greater than the last element of the permutation, we know that the last element of the permutation is $s_{|s|-v_n}$ where $|S|$ denotes the length of S . We then remove this value from the list s . Next, the $n-1^{st}$ element of the permutation is determined as $s_{|s|-v_{n-1}}$, and again we remove this element from the list s , and continue until the list is empty. $\square$

DEFINITION 4. The *number of inversions* of a permutation σ (or more simply the *inversions*) is the sum of the elements of the inversion table of σ . We denote the inversions of σ by $i(\sigma)$, or simply by i if σ is understood from the context. In symbols

$$\begin{aligned}
 i(\sigma) &= \sum_{j=1}^n v_j(\sigma) \\
 &= \sum_{j=1}^n \sum_{k=1}^j \mathbb{I}(\sigma_k > \sigma_j)
 \end{aligned}$$

We see immediately that

$$0 \leq i(\sigma) \leq \frac{n(n-1)}{2},$$

which will be useful later.

DEFINITION 5. A *transposition* is a permutation that interchanges two elements of the permutation, leaving all other elements fixed. An adjacent transposition is a transposition that interchanges two adjacent elements of the permutation, which is also referred to occasionally as a *swap*.

THEOREM 6. [SF] *The number of inversions of a permutation is the minimum number of adjacent transpositions (swaps) needed to turn the permutation into the identity.*

PROOF. Consider the k^{th} element of the permutation. It has to change places with precisely v_k elements to its left. $\square$

EXAMPLE 7. Let $\sigma = (3, 4, 1, 5, 2)$, and then $v(\sigma) = (0, 0, 2, 0, 3)$, so $i(\sigma) = 5$. A sequence of 5 swaps that transforms σ to the identity is $(3, 4, 1, 5, 2) \rightarrow (3, 4, 1, 2, 5) \rightarrow (3, 1, 4, 2, 5) \rightarrow (3, 1, 2, 4, 5) \rightarrow (1, 3, 2, 4, 5) \rightarrow (1, 2, 3, 4, 5)$. This is known as the *bubblesort* algorithm, and works by moving the largest element to its place, and then moving the second largest element to its place, and so on, all by swaps.

It is also well known that the number of n -permutations with i inversions is the coefficient of x^i in $[n]_x!$ which we denote by $\phi(n, i)$ where the *q-factorial*, is defined by

$$[n]_q! = [n]_q [n-1]_q \cdots [1]_q$$

and the *q-number*, is defined by

$$[n]_q = \frac{1 - q^n}{1 - q}.$$

We now provide a proof of this fact, making use of the inversion table of the permutation.

THEOREM 8. [Bo] *The generating function of the number of permutations on $[n]$ with i inversions is $[n]_x!$.*

FIGURE 2.1.1. Staircase diagram of $(3, 1, 4, 5, 2)$

PROOF. Our proof differs from [Bo] by the focus on the staircase diagram, introduced below. We can view the problem of choosing the elements of the inversion table as putting balls into bins, where the first bin holds at most zero balls, the second bin holds at most one ball, the third bin holds at most two balls, and in general the n^{th} bin holds at most $n - 1$ balls. For example here is such a diagram for a permutation on $[5]$ with 4 inversions, corresponding to the permutation $(3, 1, 4, 5, 2)$ (which has inversion table $(0, 1, 0, 0, 3)$), see Figure 2.1.1. Note that we refer to this as a *staircase diagram* later in this thesis.

There are zero places to put a ball in the first column of the diagram corresponding to a factor of 1 in the generating function for the one and only choice of what goes into column one. There is one place to put a ball in the second column of the diagram, corresponding to a factor of $1 + x$ in the generating function. In general, there are $k - 1$ places in column k of the diagram corresponding to a factor of $1 + x + \cdots + x^{k-1}$ in the generating function. Multiplying all these together we get

$$(1 + x)(1 + x + x^2) \cdots (1 + x + x^2 + \cdots + x^{n-1}) = \prod_{j=1}^n \frac{1 - x^j}{1 - x}$$

□

We denote the number of inversions of n with i inversions by $\phi(n, i)$. Thus $\phi(n, i)$ is the coefficient of x^i in $[n]_x!$

FIGURE 2.1.2. Euler's Pentagonal numbers [W]

There is also an explicit formula (not using generating functions) for the number of permutations of $[n]$ with i inversions due to Euler [Bo]:

$$\binom{n-i-1}{i} + \sum_{j=1}^{\infty} (-1)^j \binom{n+i-u_j-j-1}{i-u_j-j} + \sum_{j=1}^{\infty} (-1)^j \binom{n+i-u_j-1}{i-u_j}$$

where the binomial coefficients are zero when the lower index is negative, and the u_j are the *Euler Pentagonal numbers*

$$u_j = \frac{j(3j-1)}{2}$$

The first few terms of this series are 1, 5, 12, 22, 35, 51, 70 and the n^{th} pentagonal number is the number of distinct dots in a pattern of dots consisting of the outlines of regular pentagons whose sides contain 1 to n dots, overlaid so that they share one vertex [W], see Figure 2.1.2.

Further results on the relationship between the pentagonal numbers and inversions can be found in [Bo, P 51].

There is no standard notation or nomenclature for the maximum element of an inversion table, therefore, we make the following definition.

DEFINITION 9. The *level* of a permutation σ is the largest element of the inversion table of σ . We denote the level of σ by $l(\sigma)$ or simply by l when σ is understood from the context. Thus for a permutation of length n we have $l(\sigma) = \max \{v(\sigma)_j : 1 \leq j \leq n\}$ or equivalently

$$l(\sigma) = \max_{1 \leq j \leq n} \left\{ \sum_{1 \leq k \leq j} \mathbb{I}(\sigma_k > \sigma_j) \right\}$$

EXAMPLE 10. Let $\sigma = (3, 4, 1, 5, 2)$. Then $v(\sigma) = (0, 0, 2, 0, 3)$, $l(\sigma) = 3$ and $i(\sigma) = 5$.

THEOREM 11. i) [SF] *The number of permutations of $[n]$ with level less than k is $k!k^{n-k}$ for $0 \leq k \leq n-1$.*

ii) *The number of permutations of n with level l is $l! \left((l+1)^{n-l} - l^{n-l} \right)$ for $0 \leq l \leq n-1$*

PROOF. i) Since the i^{th} entry of the inversion table is between 0 and $i-1$ for $i \leq k$ there are $k!$ choices for the first k elements. For the remaining $n-k$ elements, each element can be between 0 and $k-1$, for a total of k^{n-k} choices for these elements. Putting these two together, we have the conclusion.

ii) If we let $m(n, k) =$ the number of inversion table of length n with all entries less than or equal to $k-1$, then

$$\begin{aligned} m(n, k) - m(n, k-1) &= k!k^{n-k} - (k-1)!(k-1)^{n-(k-1)} \\ &= k(k-1)!k^{n-k} - (k-1)!(k-1)^{n-k}(k-1) \\ &= (k-1)! \left(k^{n-k+1} - (k-1)^{n-k+1} \right) \end{aligned}$$

and so the number of inversion tables of length n with level l is

$$l! \left((l+1)^{n-l} - l^{n-l} \right)$$

□

FIGURE 2.1.3. Balls & bars

EXAMPLE 12. There are 10 permutations of $[4]$ with level 2, namely $(1, 3, 4, 2), (1, 4, 3, 2), (2, 3, 1, 4), (2, 4, 1, 3), (3, 1, 4, 2), (3, 2, 1, 4), (3, 4, 1, 2), (4, 1, 3, 2), (4, 2, 1, 3)$ and $(4, 3, 1, 2)$.

The next question to be addressed is: what are the possible values for the level of an n -permutation with i inversions, and how many are there for each level?

Clearly, we have the bound

$$l(\sigma) \leq (n-1) \wedge i(\sigma)$$

but determining exactly how many there will take some more work. We present two different approaches, and we begin several combinatorial lemmas.

LEMMA 13. [Ai, Ex 1.59] *The number of solutions to the equation*

$$a_1 + a_2 + \cdots + a_n = k$$

where

$$a_i \geq 0, a_i \in \mathbb{N}$$

is

$$\binom{n+k-1}{k}$$

PROOF. Put k balls in a row. Between the balls insert $n-1$ bars. Figure 2.1.3 is an example with four balls and two bars.

Thus we have a total of $k+n-1$ positions that can be a ball or a bar. $\square$

LEMMA 14. [Ai, p 181] *The number of solutions to the equation*

$$a_1 + a_2 + \cdots + a_n = k$$

where $s > a_i \geq 0$, $a_i \in \mathbb{N}$ for all i is

$$\sum_{j=0}^n (-1)^j \binom{n}{j} \binom{n+k-j s-1}{n-1}$$

PROOF. Consider the number of solutions to the above set of equations, where now $x_1 \geq s, x_2 \geq s, \dots, x_j \geq s$, the other x_k being free. Then let $y_i = x_i - s$ if $i \leq j$ or $y_i = x_i$ if $i > j$. This modification has the same number of solutions as $y_1 + y_2 + \dots + y_n = k - j s$, which equals $\binom{n+k-j s-1}{k-j s} = \binom{n+k-j s-1}{n-1}$ by standard binomial coefficient manipulations:

$$\begin{aligned} \binom{a+b-1}{b} &= \frac{(a+b-1)!}{b!(a+b-1-b)!} \\ &= \frac{(a+b-1)!}{b!(a-1)!} \\ &= \frac{(a+b-1)!}{(a-1)!(a+b-1-(a-1))!} \\ &= \binom{a+b-1}{a-1} \end{aligned}$$

By the inclusion-exclusion formula, we have the conclusion. $\square$

Thus, the number of solutions to the equation

$$a_1 + a_2 + \dots + a_n = k$$

where $s \geq a_i \geq 0$, $a_i \in \mathbb{N}$ for all i is

$$\sum_{j=0}^n (-1)^j \binom{n}{j} \binom{n+k-j(s+1)-1}{n-1}$$

DEFINITION 15. We denote the number of n -permutations with i inversions and level l by $\psi(n, i, l)$

THEOREM 16. *The number of n -permutations with i inversions with level l is*

FIGURE 2.1.4. Staircase diagram truncated at level 3

$$\psi(n, i, l) = \sum_{t=(i-\frac{l(l-1)}{2}) \vee l}^{(n-l)l \wedge i} \phi(l, i-t) a(n-l, t, l)$$

where

$$a(n, i, m) = c(n, i, m) - c(n, i, m-1)$$

and

$$c(n, k, s) = \sum_{j=0}^{n \wedge \frac{k}{s+1} \wedge \frac{n+k-1}{s+1}} (-1)^j \binom{n}{j} \binom{n+k-j(s+1)-1}{n-1}$$

PROOF. We seek the number of solutions $s(n, m, i)$ of the following system of equations, for each $m = 0, 1, \dots, n-1$

$$x_1 + x_2 + \dots + x_{n-1} = i$$

where

$$\forall j : 0 \leq x_j \leq m \wedge j$$

and

$$\exists j : x_j = m$$

This can be accomplished by putting i balls into a *truncated* staircase diagram, as in Theorem 7, but where the maximum height is constrained to be less than or equal to an upper boundary, and where at least one column hits the upper boundary. Figure 2.1.4 displays a staircase diagram truncated at 3.

We put $i - t$ balls into the staircase part, and t into the rectangular part. The number of ways to put t balls into the staircase part is the same as the number of inversions of a permutation $\phi(m, i - t)$, and the number of ways to put t into the rectangular part is $a(n - m, t, m)$, where $a(n, i, m)$ is the number of solutions of

$$x_1 + x_2 + \cdots + x_n = i$$

where

$$\forall j : 0 \leq x_j \leq m$$

and

$$\exists j : x_j = m$$

This is known as a weak composition of i into n parts, with at least one part equaling m , and all parts less than or equal to m . Putting it all together, and being very explicit about indices, we have

$$s(n, m, i) = \sum_{t=\max\{i-\frac{m(m-1)}{2}, m\}}^{\min\{(n-m)m, i\}} \phi(m, i - t) a(n - m, t, m)$$

where

$$a(n, i, m) = c(n, i, m) - c(n, i, m - 1)$$

and

$$c(n, i, m) = \sum_{j=0}^{\min\{n, \frac{i}{m+1}, \frac{n+i-1}{m+1}\}} (-1)^j \binom{n}{j} \binom{n+i-j(m+1)-1}{n-1}$$

□

Unfortunately, this formula does not provide much insight into its asymptotic behavior.

Another approach yields the following.

THEOREM 17. *The generating function of the number of n -permutations with i inversions and level l is*

$$[l]_x! [l+1]_x^{n-l} - [l-1]_x! [l]_x^{n-l+1}$$

PROOF. The generating function of the number of permutations of $[n]$ with i inversions and level at most l is

$$\prod_{j=1}^{n-1} \sum_{k=0}^{j \wedge l} x^k$$

and so the generating function of the number of permutations of $[n]$ with i inversions and level exactly l is

$$\prod_{j=1}^{n-1} \sum_{k=0}^{j \wedge l} x^k - \prod_{j=1}^{n-1} \sum_{k=0}^{j \wedge (l-1)} x^k$$

The conclusion follows after some simplification via the identity

$$\begin{aligned} \prod_{j=1}^{n-1} \sum_{k=0}^{j \wedge l} x^k &= \prod_{j=1}^{n-1} \frac{1 - x^{(j \wedge l) + 1}}{1 - x} \\ &= \prod_{j=1}^l \frac{1 - x^j}{1 - x} \times \left(\frac{1 - x^{l+1}}{1 - x} \right)^{n-l} \\ &= \prod_{j=1}^l \sum_{k=0}^{j-1} x^k \times \left(\sum_{k=0}^l x^k \right)^{n-l} \\ &= \prod_{j=1}^l ([j]_x) \times ([l+1]_x)^{n-l} \\ &= [l]_x! \times [l+1]_x^{n-l} \end{aligned}$$

And so

$$\begin{aligned} \prod_{j=1}^{n-1} \sum_{k=0}^{j \wedge l} x^k - \prod_{j=1}^{n-1} \sum_{k=0}^{j \wedge (l-1)} x^k &= [l]_x! \times [l+1]_x^{n-l} - [l-1]_x! \times [l]_x^{n-l+1} \\ &= [l]_x! [l+1]_x^{n-l} - [l-1]_x! [l]_x^{n-l+1} \end{aligned}$$

□

Figure 2.1.5 is a plot of $\psi(20, i, l)$

FIGURE 2.1.5. $\psi(20, i, l)$

2.2. Sequential Construction

In the previous sections we viewed a permutation as being given “all at once”. In this section we adopt a more sequential view. This new viewpoint leads to interesting questions, such as how does the number of inversions change? How does the level change? We address these issues in this section.

Imagine we already have a permutation on $[n]$, and element $n+1$ arrives which may be inserted anywhere into the permutation. We call this sequential construction of permutations.

EXAMPLE 18. Assume we start with the permutation $(3, 5, 4, 1, 2)$. We then receive 6, which may go into 6 different places, giving the following as possibilities.

$$(6, 3, 5, 4, 1, 2)$$

$$(3, 6, 5, 4, 1, 2)$$

$$(3, 5, 6, 4, 1, 2)$$

$$(3, 5, 4, 6, 1, 2)$$

$$(3, 5, 4, 1, 6, 2)$$

$$(3, 5, 4, 1, 2, 6)$$

We see that by adding the new value, the number of inversions can increase by between 0 and 5, and the level can increase by at most 1. We formalize this next.

THEOREM 19. *Let σ be a fixed permutation on $[n]$ with i inversions and level l . Denote σ after the element $n + 1$ has been added by σ' , its inversions by i' , and its level by l' . Then*

$$i) 0 \leq i' - i \leq n$$

$$ii) 0 \leq l' - l \leq 1$$

PROOF. i) The symbol $n + 1$ can go into $n + 1$ locations, from the leftmost, where it will increase the number of inversions by n , to the rightmost, where it will not increase the number of inversions. In fact, if symbol $n + 1$ is inserted into position k , then $i_{n+1} = i_n + n - k + 1$.

ii) Inserting the new symbol into the permutation at a position leaves all the elements of the inversion table to the left of the position unchanged. All elements to the right of the newly inserted symbol are increased by one. $\square$

Now consider the inversion table of a permutation. If we add a new element to the permutation into position k , then the element of the inversion table at position k is 0, and all elements of the inversion table with indices greater than k get shifted to the right and incremented.

In other words if the new element is inserted into position k , then the inversion table transforms as follows

$$(v_1, v_2, \dots, v_{k-1}, v_k, v_{k+1}, \dots, v_n)$$

to

$$(v_1, v_2, \dots, v_{k-1}, 0, v_k + 1, v_{k+1} + 1, \dots, v_n + 1)$$

Consequently, if position k is selected, the number of inversions goes from

$$v_1 + v_2 + \dots + v_{k-1} + v_k + v_{k+1} + \dots + v_n$$

to

$$v_1 + v_2 + \dots + v_{k-1} + 0 + v_k + 1 + v_{k+1} + 1 + \dots + v_n + 1$$

i.e. the inversions increase by $n + 1 - k$.

The level goes from

$$\max \{v_1, v_2, \dots, v_{k-1}, v_k, v_{k+1}, \dots, v_n\}$$

to

$$\max \{v_1, v_2, \dots, v_{k-1}, 0, v_k + 1, v_{k+1} + 1, \dots, v_n + 1\}$$

or, equivalently, from

$$\max \{v_1, v_2, \dots, v_{k-1}\} \bigvee \max \{v_k, v_{k+1}, \dots, v_n\}$$

to

$$\max \{v_1, v_2, \dots, v_{k-1}\} \bigvee \max \{v_k + 1, v_{k+1} + 1, \dots, v_n + 1\}$$

which equals

$$\max \{v_1, v_2, \dots, v_{k-1}\} \bigvee (1 + \max \{v_k, v_{k+1}, \dots, v_n\})$$

We now look at each case. If the maximal element is in the rightmost part (i.e. positions $k \dots n$), then the level after inserting the new element is the previous level $+1$. If the maximal element is in the leftmost part, then the level after inserting the new element is the previous level. If the maximal element is in both parts, then the level after inserting the new element is the previous level $+1$. Thus we have the following theorem.

THEOREM 20. *Using the notation of Theorem 19, we have $l' = l$ if the position of a maximal element is less than the position of element $n + 1$, and no maximal element is in position greater than that of $n + 1$. Otherwise, $l' = l + 1$.*

CHAPTER 3

The Uniform Measure on Random Permutations

In this chapter we describe the inversions and level of a permutation selected from the uniform probability measure.

3.1. Uniform Measure

The *uniform measure* simply assigns probability $\frac{1}{n!}$ to every permutation of $[n]$. Thus, with $\mathbf{X}$ being a random permutation of $[n]$, we have

$$\mathbb{P}[\mathbf{X} = x] = \frac{1}{n!}.$$

if x is a permutation of $[n]$, and 0 otherwise.

In order to ease the notation, we shall denote by $\mathbf{I}$ the number of inversions of a random permutation, and by $\mathbf{L}$ the level of a random permutation, all under the uniform measure. Unless otherwise stated, all permutations are on $[n]$.

3.2. Basic Results

We begin with some basic results about permutations selected uniformly at random.

THEOREM 21. [Mah] *The components of the inversion table of a uniformly random permutation on $[n]$ are independent random variables, and the k^{th} component of the inversion table has uniform distribution on $\{0, 1, 2, \dots, k-1\}$.*

PROOF. The $\{v_j\}_{j=1, \dots, n}$ are independent: knowing $v_1, v_2, \dots, v_k$ is equivalent to knowing the relative ordering of $1, 2, \dots, k$ in the permutation, but does not provide any information about $k+1$. Thus, v_{k+1} is independent of $v_1, v_2, \dots, v_k$, and so the $\{v_j\}_{j=1, \dots, n}$ are independent. Also, under the uniform measure, in the

random permutation k is equally likely to be in any of the positions $1, 2, \dots, k$ and so

$$\mathbb{P}[v_k = j] = 1/k, \quad j = 0, 1, \dots, k-1$$

□

COROLLARY 22. [Mah] *The mean and variance of the k^{th} component of the inversion table of a uniformly random permutation on $[n]$ are $\frac{k-1}{2}$ and $\frac{k^2-1}{12}$.*

PROOF. From the definitions we have

$$\begin{aligned} \mathbb{E}[v_k] &= \sum_{j=0}^{k-1} \frac{j}{k} \\ &= \frac{k-1}{2} \end{aligned}$$

$$\begin{aligned} \mathbb{E}[v_k^2] &= \sum_{j=0}^{k-1} \frac{j^2}{k} \\ &= \frac{(k-1)(2k-1)}{6} \end{aligned}$$

$$\begin{aligned} \mathbb{E}[v_k^2] - \mathbb{E}[v_k]^2 &= \frac{(k-1)(2k-1)}{6} - \left(\frac{k-1}{2}\right)^2 \\ &= \frac{(k-1)(2k-1)}{6} - \frac{(k-1)^2}{4} \\ &= \frac{2(k-1)(2k-1) - 3(k-1)^2}{12} \\ &= \frac{k^2-1}{12} \end{aligned}$$

□

THEOREM 23. [Bo, 5 and 9] *We have the following:*

$$\begin{aligned}
1) \mathbb{P}[\mathbf{I} = i] &= \frac{\phi(n, i)}{n!} \\
2) \mathbb{P}[\mathbf{L} = l] &= \frac{l! \left((l+1)^{n-l} - l^{n-l} \right)}{n!} \\
3) \mathbb{P}[\mathbf{I} = i | \mathbf{L} = l] &= \frac{\psi(n, i, l)}{l! \left((l+1)^{n-l} - l^{n-l} \right)} \\
4) \mathbb{P}[\mathbf{L} = l | \mathbf{I} = i] &= \frac{\psi(n, i, l)}{\phi(n, i)} \\
5) \mathbb{E}[\mathbf{I}] &= \frac{n(n-1)}{4} \\
6) \mathbb{E}[\mathbf{L}] &= n-1 - \sum_{l=1}^{n-1} \left(\frac{l! l^{n-l}}{n!} \right) \\
7) \mathbb{E}[\mathbf{I} | \mathbf{L}] &= \sum_{i=0}^{\frac{n(n-1)}{2}} \frac{i \psi(n, i, l)}{l! \left((l+1)^{n-l} - l^{n-l} \right)} \\
8) \mathbb{E}[\mathbf{L} | \mathbf{I}] &= \sum_{i=0}^{n-1} \frac{l \psi(n, i, l)}{\phi(n, i)} \\
9) \mathbb{V}[\mathbf{I}] &= \frac{n(n-1)(2n+5)}{72} \\
10) \mathbb{V}[\mathbf{L}] &= n(n-1) - 2 \sum_{l=1}^{n-1} \frac{l! l^{n-l+1}}{n!} - \left(n-1 - \sum_{l=1}^{n-1} \left(\frac{l! l^{n-l}}{n!} \right) \right)^2
\end{aligned}$$

PROOF. 1) Straightforward application of the definitions.

2) Straightforward application of the definitions.

3)

$$\begin{aligned}
\mathbb{P}[\mathbf{I} = i | \mathbf{L} = l] &= \frac{\mathbb{P}[\mathbf{I} = i, \mathbf{L} = l]}{\mathbb{P}[\mathbf{L} = l]} \\
&= \frac{\psi(n, i, l)}{l! \left((l+1)^{n-l} - l^{n-l} \right)}
\end{aligned}$$

4)

$$\begin{aligned}
\mathbb{P}[\mathbf{L} = l | \mathbf{I} = i] &= \frac{\mathbb{P}[\mathbf{I} = i, \mathbf{L} = l]}{\mathbb{P}[\mathbf{I} = i]} \\
&= \frac{\psi(n, i, l)}{\phi(n, i)}
\end{aligned}$$

5)

$$\begin{aligned}
\mathbb{E}[\mathbf{I}] &= \sum_{k=1}^n \mathbb{E}[v_k] \\
&= \sum_{k=1}^n \frac{k-1}{2} \\
&= \frac{n(n-1)}{4}
\end{aligned}$$

6)

$$\begin{aligned}
\mathbb{P}[\mathbf{L}_n] &= \sum_{l=1}^{\infty} \mathbb{P}[\mathbf{L}_n \geq l] \\
&= \sum_{l=1}^{\infty} \left(1 - \frac{l!l^{n-l}}{n!}\right) \\
&= \sum_{l=1}^{\infty} \left(\frac{n! - l!l^{n-l}}{n!}\right) \\
&= \sum_{l=1}^{n-1} \left(\frac{n! - l!l^{n-l}}{n!}\right) \\
&= n-1 - \sum_{l=1}^{n-1} \left(\frac{l!l^{n-l}}{n!}\right)
\end{aligned}$$

7)

$$\begin{aligned}
\mathbb{E}[\mathbf{I} | \mathbf{L}] &= \sum_{i=0}^{\frac{n(n-1)}{2}} i \mathbb{P}[\mathbf{I} = i | \mathbf{L} = l] \\
&= \sum_{i=0}^{\frac{n(n-1)}{2}} \frac{i \psi(n, i, l)}{l! \left((l+1)^{n-l} - l^{n-l}\right)}
\end{aligned}$$

8)

$$\begin{aligned}
\mathbb{E}[\mathbf{L} | \mathbf{I}] &= \sum_{i=0}^{n-1} l \mathbb{P}[\mathbf{L} = l | \mathbf{I} = i] \\
&= \sum_{i=0}^{n-1} \frac{l \psi(n, i, l)}{\phi(n, i)}
\end{aligned}$$

9)

$$\begin{aligned}
\mathbb{V}[\mathbf{I}] &= \sum_{k=1}^n \mathbb{V}[v_k] \\
&= \sum_{k=1}^n \frac{k^2 - 1}{12} \\
&= \frac{2n^3 + 3n^2 - 5n}{72} \\
&= \frac{n(n-1)(2n+5)}{72}
\end{aligned}$$

10) Direct from 6).

□

3.3. Limits

In this section we show three limit theorems as the size of the permutation grows: the Gaussian distribution of the number of inversions, the expected value of the level, and the Rayleigh distribution of the level.

THEOREM 24. [Mah] *In the case of the uniform distribution of permutations on $[n]$*

$$\frac{\mathbf{I}_n - n^2/4}{n^{3/2}} \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{36}\right)$$

PROOF. We check the conditions of Lindeberg's central limit theorem: consider the random variables $\tilde{v}_k = v_k - \mathbb{E}[v_k]$. We have $\sum_{k=1}^n \tilde{v}_k = \mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]$, and let $s_n^2 = \sum_{k=1}^n \mathbb{V}[\tilde{v}_k] = \sum_{k=1}^n \mathbb{V}[v_k] \sim \frac{n^3}{36}$. For fixed $\varepsilon > 0$ we verify that

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \int_{\{|\tilde{v}_k| > \varepsilon s_n\}} \tilde{v}_k^2 d\mathbb{P} = 0.$$

This limit is equal to

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \sum_{|j - \mathbb{E}[v_k]| > \varepsilon s_n} (j - \mathbb{E}[v_k])^2 \mathbb{P}[v_k = j].$$

However, for large n , the set $\{v_k = j : |j - \mathbb{E}[v_k]| > \varepsilon s_n\}$ is empty because $s_n^2 \sim n^3$, whereas $v_k < k \leq n$. Thus for large n , the inner summation is zero. $\square$

DEFINITION 25. $f(n) \sim g(n)$ if for all $\varepsilon > 0$ there exists an N such that for all $n > N$ we have

$$\left| \frac{f(n)}{g(n)} - 1 \right| < \varepsilon.$$

THEOREM 26. [K] *Expected value of $\mathbf{L}_n$ for large n is*

$$\mathbb{E}[\mathbf{L}_n] \sim n - \sqrt{\pi n/2}$$

PROOF. [MSE1] We provide only a heuristic outline of a proof here, following the method of Laplace, as later we provide a more probabilistic proof that has the advantage of also identifying the limiting distributional form of the level. We begin with (replacing l with k to make the reading easier)

$$\mathbb{E}[\mathbf{L}_n] = n - 1 - \sum_{k=1}^{n-1} \left(\frac{k! k^{n-k}}{n!} \right)$$

and so we are led to considering $\sum_{k=1}^{n-1} \frac{k! k^{n-k}}{n!}$ which we approximate by $\int_1^n \frac{k! k^{n-k}}{n!} dk$.

Now consider, by Stirling's approximation

$$\begin{aligned} k! k^{n-k} &\sim \sqrt{2\pi k} \left(\frac{k}{e} \right)^k k^{n-k} \\ &= \sqrt{2\pi k}^{\frac{1}{2}} k^n e^{-k} \\ &= \exp \left(\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(k) + n \log(k) - k \right) \end{aligned}$$

and $\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(k) + n \log(k) - k$ is maximized, for fixed n , by setting $k = n + \frac{1}{2}$. But as we have $k \leq n$, and $\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(k) + n \log(k) - k$ is monotonically increasing in k because

$$1 + \frac{1}{2k} + \frac{n}{k} \geq 0$$

the function is maximized at $k = n$. Setting $k = n$ and expanding about this point we have

$$\begin{aligned} \frac{1}{2} \log(2\pi) + \frac{1}{2} \log(k) + n \log(k) - k &= \frac{1}{2} \log(2\pi) + \frac{1}{2} \log(n) - \frac{(k-n)^2}{2n} - n + n \log n + O(n^3) \\ e^{\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(k) + n \log(k) - k} &\sim \exp\left(\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(n) - \frac{(k-n)^2}{2n} - n + n \log n\right) \\ &= \exp\left(\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(n)\right) \exp\left(-\frac{(k-n)^2}{2n} - n + n \log n\right) \\ &= \exp\left(\log(2\pi)^{\frac{1}{2}}\right) \exp\left(\log\left(n^{\frac{1}{2}}\right)\right) \exp\left(-\frac{(k-n)^2}{2n} - n + n \log n\right) \\ &= \sqrt{2\pi n} e^{-\frac{(k-n)^2}{2n}} e^{-n + n \log n + \frac{1}{2} \log(2\pi n)} \\ &\sim e^{-\frac{(k-n)^2}{2n}} n! \end{aligned}$$

so $k!k^{n-k} \sim e^{-\frac{(k-n)^2}{2n}} n!$ and so $\frac{k!k^{n-k}}{n!} \sim e^{-\frac{(k-n)^2}{2n}}$. Finally we have

$$\begin{aligned} \int_1^n \frac{k!k^{n-k}}{n!} dk &\sim \int_{-\infty}^n e^{-(k-n)^2/2n} dk \\ &= \sqrt{\frac{\pi n}{2}} \end{aligned}$$

□

We now show that the distribution of the (scaled and centered) level converges in distribution to a Rayleigh distribution, and then uniform integrability shows that the moments of the level converge to those of the Rayleigh distribution.

THEOREM 27. [Mah] *The level satisfies*

$$\frac{\mathbf{L}_n - n}{\sqrt{n}} \xrightarrow{\mathcal{D}} -\mathbf{R}$$

where the density of $\mathbf{R}$, a Rayleigh distributed random variable is

$$xe^{-x^2/2} \text{ for } x > 0$$

PROOF. Consider $\mathbf{L}_n^* = (\mathbf{L}_n - \mu_n) / \sigma_n$ for some numbers μ_n and σ_n . Then

$$\begin{aligned} \mathbb{P}[\mathbf{L}_n^* \leq x] &= \mathbb{P}[\mathbf{L}_n \leq \mu_n + x\sigma_n] \\ &= \mathbb{P}[\mathbf{L}_n \leq \lfloor \mu_n + x\sigma_n \rfloor] \end{aligned}$$

We have that

$$\mathbb{P}[\mathbf{L}_n^* \leq x] = \lfloor \mu_n + x\sigma_n \rfloor^{n - \lfloor \mu_n + x\sigma_n \rfloor} \times \frac{\lfloor \mu_n + x\sigma_n \rfloor!}{n!}$$

and $\lfloor \mu_n + x\sigma_n \rfloor = \mu_n + x\sigma_n - \beta_n$ for some $0 \leq \beta_n < 1$. Then

$$\mathbb{P}[\mathbf{L}_n^* \leq x] = (\mu_n + x\sigma_n - \beta_n)^{n - (\mu_n + x\sigma_n - \beta_n)} \times \frac{(\mu_n + x\sigma_n - \beta_n)!}{n!}$$

and by using Stirling's approximation, we have

$$\begin{aligned} \mathbb{P}[\mathbf{L}_n^* \leq x] &\sim (\mu_n + x\sigma_n - \beta_n)^{n - \mu_n - x\sigma_n + \beta_n} \times \frac{\left(\frac{\mu_n + x\sigma_n - \beta_n}{e}\right)^{\mu_n + x\sigma_n - \beta_n} \times \sqrt{2\pi(\mu_n + x\sigma_n - \beta_n)}}{\left(\frac{n}{e}\right)^n \times \sqrt{2\pi n}} \\ &= \frac{(\mu_n + x\sigma_n - \beta_n)^n}{n^n} e^{n - \mu_n - x\sigma_n + \beta_n} \sqrt{\frac{\mu_n + x\sigma_n - \beta_n}{n}} \\ &= \frac{\mu_n^n}{n^n} \left(1 + \frac{x\sigma_n - \beta_n}{\mu_n}\right)^n e^{n - \mu_n - x\sigma_n + \beta_n} \sqrt{\frac{\mu_n + x\sigma_n - \beta_n}{n}} \end{aligned}$$

Using the relation $\left(1 + \frac{f_n}{n}\right)^n \sim e^{f_n}$ when $f_n = O(n^\varepsilon)$ with $\varepsilon < 1$ we take $\mu_n = n$.

As $\mathbf{L}_n < n$, $\mathbb{P}[\mathbf{L}_n^* \leq x] = 1$ for any non-negative x . Any fixed negative x is in the support of this distribution, because $\mathbf{L}_n \geq 1$, i.e. $\mathbf{L}_n^* \geq -(n-1)/\sigma_n$ and so $-(n-1)/\sigma_n \leq x < 0$ if $\sigma_n = o(n)$ and n is large enough. We can write

$$\mathbb{P}[\mathbf{L}_n^* \leq x] \sim \left(1 + \frac{x\sigma_n - \beta_n}{n}\right)^n e^{-x\sigma_n + \beta_n} \sqrt{\frac{n + x\sigma_n}{n}}.$$

Now,

$$\left(1 + \frac{x\sigma_n - \beta_n}{n}\right)^n = \exp \left\{ n \log \left(1 + \frac{x\sigma_n - \beta_n}{n}\right) \right\}$$

Expanding the logarithm as $(x\sigma_n/n - \beta_n/n) - x^2\sigma_n^2/(2n^2) + O(\sigma_n^2/n^2)$ we get

$$\mathbb{P}[\mathbf{L}_n^* \leq x] \sim e^{-x^2\sigma_n^2/(2n) + O(\sigma_n/n)} \sqrt{\frac{n + x\sigma_n}{n}}$$

The right hand side converges if $\sigma_n = c\sqrt{n}$ for any constant c , so let us take $c = 1$ to obtain

$$\mathbb{P}[\mathbf{L}_n^* \leq x] \rightarrow e^{-x^2/2} \text{ for } x < 0.$$

This latter distribution is that of a negative random variable $-\mathbf{R}$, so $\mathbf{R}$ has distribution

$$1 - e^{-x^2/2} \text{ for } x > 0$$

and the density follows by differentiation

$$xe^{-x^2/2} \text{ for } x > 0$$

□

COROLLARY 28. **[Mah]** $\mathbb{E}[\mathbf{L}_n] = n - \sqrt{\frac{\pi n}{2}} + o(\sqrt{n})$ and $\mathbb{V}[\mathbf{L}_n] \sim (2 - \frac{\pi}{2})n$

PROOF. Uniform integrability implies that the moments converge, so we have

$$\begin{aligned} \mathbb{E}\left[\frac{\mathbf{L}_n - n}{\sqrt{n}}\right] &\rightarrow \mathbb{E}[-\mathbf{R}] \\ &= -\int_0^\infty x^2 e^{-x^2/2} dx \\ &= -\sqrt{\frac{\pi}{2}} \end{aligned}$$

implying

$$\mathbb{E}[\mathbf{L}_n] = n - \sqrt{\frac{\pi n}{2}} + o(\sqrt{n})$$

For the second moment we have

$$\mathbb{V}\left[\frac{\mathbf{L}_n - n}{\sqrt{n}}\right] \rightarrow \mathbb{V}[-\mathbf{R}]$$

or

$$\begin{aligned}
 \frac{1}{n} \mathbb{V}[\mathbf{L}_n] &\rightarrow \mathbb{V}[\mathbf{R}] \\
 &= \mathbb{E}[\mathbf{R}^2] - (\mathbb{E}[\mathbf{R}])^2 \\
 &= \int_0^\infty x^3 e^{-x^2/2} dx - \left(-\sqrt{\frac{\pi}{2}}\right)^2 \\
 &= 2 - \frac{\pi}{2}
 \end{aligned}$$

□

Finally we have:

THEOREM 29. **[Mah]** *Convergence in probability*

$$\frac{\mathbf{L}_n}{n} \xrightarrow{P} 1$$

PROOF. Let $\varepsilon > 0$. By Chebyshev's inequality

$$\begin{aligned}
 \mathbb{P}\left[\left|\frac{\mathbf{L}_n}{n} - 1\right| \geq \varepsilon\right] &\leq \frac{\mathbb{V}\left[\frac{\mathbf{L}_n}{n} - 1\right]}{\varepsilon^2} \\
 &= \frac{\mathbb{V}[\mathbf{L}_n]}{n^2 \varepsilon^2} \\
 &= \frac{(2 - \pi/2)n + o(n)}{n^2 \varepsilon^2} \\
 &\rightarrow 0 \text{ as } n \rightarrow \infty
 \end{aligned}$$

□

Note that if for any $\varepsilon > 0$ we have

$$\sum_{n=1}^{\infty} \frac{(2 - \pi/2)n + o(n)}{n^2 \varepsilon^2} < \infty$$

FIGURE 3.3.1. Histograms and limiting density of the inversions and the level

(complete convergence) then we have a.s. convergence. But

$$\begin{aligned}
 \sum_{n=1}^{\infty} \frac{(2 - \pi/2)n}{n^2 \varepsilon^2} &= \sum_{n=1}^{\infty} \frac{2 - \pi/2}{n \varepsilon^2} \\
 &= \frac{(2 - \pi/2)}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{n} \\
 &\rightarrow \infty
 \end{aligned}$$

so we do not have complete convergence here.

For the case of $n=100$, we have histograms (based on 10,000 random permutations) and the limiting density of the inversions and the level shown in Figure 3.3.1.

Shown in figure 3.3.2 is a contour plot of the joint histogram of the inversions and the level of the 10,000 random permutations.

THEOREM 30. *Consider the components of the inversion table of a uniform random permutation of $[n]$. From each component subtract the mean and divide*

FIGURE 3.3.2. Joint histogram of the inversions and the level

by the standard deviation to obtain a sequence of independent, mean zero, unit variance random variables, and interpolate linearly between the values. Then this process converges in distribution to a Brownian motion as $n \rightarrow \infty$.

PROOF. The conclusion follows from a version of Donsker's theorem for independent, mean zero, unit variance random variables [Bi, P 94, 8.4] and the verification of the Lindeberg condition from Theorem 24. $\square$

3.4. Sampling

Selecting a permutation of $[n]$ uniformly at random is simple, and is included here only for completeness. Perhaps the simplest algorithm is the following. Assume we start with the identity permutation $\sigma = (1, 2, \dots, n)$. Now select an element uniformly at random from σ , and place it at the start of a new permutation. Remove the element from σ , and repeat the process.

EXAMPLE 31. Let $n = 5$. $\sigma = (1, 2, 3, 4, 5)$. We select 3 uniformly at random and place it at the start of a new permutation $\pi = (3)$, and remove it from σ , leaving $\sigma = (1, 2, 4, 5)$. Next we select 2, giving $\pi = (3, 2)$ and $\sigma = (1, 4, 5)$. Then we select 1, giving $\pi = (3, 2, 1)$ and $\sigma = (4, 5)$ and then we select 5, giving $\pi = (3, 2, 1, 5)$ and $\sigma = (4)$. Finally, with no choice left, we select 4, giving $\pi = (3, 2, 1, 5, 4)$.

3.5. Sequential Construction

It is clear from the general results on sequential construction of permutations that the inversions has the particularly simple evolution

$$\mathbb{P}[I_{n+1} - I_n = \alpha] = \frac{1}{n+1} I_{\alpha \in \{0, 1, \dots, n\}}$$

and so

$$\begin{aligned} \mathbb{E}[I_{n+1} - I_n] &= \sum_{\alpha=0}^n \frac{\alpha}{n+1} \\ &= \frac{1}{n+1} \sum_{\alpha=0}^n \alpha \\ &= \frac{n}{2} \end{aligned}$$

and

$$\begin{aligned} \mathbb{V}[I_{n+1} - I_n] &= \mathbb{E}[(I_{n+1} - I_n)^2] - (\mathbb{E}[I_{n+1} - I_n])^2 \\ &= \sum_{\alpha=0}^n \frac{\alpha^2}{n+1} - \left(\frac{n}{2}\right)^2 \\ &= \frac{1}{n+1} \left(\frac{n(n+1)(2n+1)}{6} \right) - \left(\frac{n}{2}\right)^2 \\ &= \frac{n(2n+1)}{6} - \frac{n^2}{4} \\ &= \frac{n(2+7n)}{12} \end{aligned}$$

Unfortunately, the level is not so simple. However, two applications of the Azuma Hoeffding inequality, which we next prove, provide concentration results for the inversions and the level.

THEOREM 32. [L, Prop 10.5.1]. *Suppose a sequence of random variables $\mathbf{X}_n$ forms a martingale with respect to a σ -field $\mathcal{F}_n$, and that $\mathbf{X}_0 = 0$. If there exists a sequence of constants c_n such that $\mathbb{P}[|\mathbf{X}_n - \mathbf{X}_{n-1}| \leq c_n] = 1$, then*

$$\mathbb{E}[e^{\beta \mathbf{X}_n}] \leq e^{\frac{\beta^2}{2} \sum_{k=1}^n c_k^2}$$

for all $\beta > 0$. Furthermore

$$\mathbb{P}[\mathbf{X}_n \geq \lambda] \leq e^{-\lambda^2 / (2 \sum_{k=1}^n c_k^2)}$$

and

$$\mathbb{P}[|\mathbf{X}_n| \geq \lambda] \leq 2e^{-\lambda^2 / (2 \sum_{k=1}^n c_k^2)}$$

for all $\lambda > 0$.

PROOF. Because the function $\exp$ is convex, $\exp(\alpha u + (1 - \alpha)v) \leq \alpha \exp(u) + (1 - \alpha) \exp(v)$ for any $\alpha \in [0, 1]$, u and v . Setting $u = -\beta c$, $v = \beta c$ and $\alpha = \frac{c-x}{2c}$ for $x \in [-c, c]$ yields

$$e^{\beta x} \leq \frac{c-x}{2c} e^{-\beta c} + \frac{c+x}{2c} e^{\beta c}$$

If we substitute $\mathbf{X}_n - \mathbf{X}_{n-1}$ for x and c_n for c in this expression and take conditional expectations, then the hypothesis $|\mathbf{X}_n - \mathbf{X}_{n-1}| \leq c_n$ and the fact that $\mathbb{E}[\mathbf{X}_n - \mathbf{X}_{n-1} | \mathcal{F}_{n-1}] = 0$ imply

$$\begin{aligned} \mathbb{E}[e^{\beta \mathbf{X}_n}] &= \mathbb{E}[\mathbb{E}[e^{\beta \mathbf{X}_n} | \mathcal{F}_{n-1}]] \\ &= \mathbb{E}[\mathbb{E}[e^{\beta \mathbf{X}_{n-1} + \beta \mathbf{X}_n - \beta \mathbf{X}_{n-1}} | \mathcal{F}_{n-1}]] \\ &= \mathbb{E}\left[\mathbb{E}\left[e^{\beta \mathbf{X}_{n-1}} e^{\beta(\mathbf{X}_n - \mathbf{X}_{n-1})} | \mathcal{F}_{n-1}\right]\right] \\ &= \mathbb{E}\left[e^{\beta \mathbf{X}_{n-1}} \mathbb{E}\left[e^{\beta(\mathbf{X}_n - \mathbf{X}_{n-1})} | \mathcal{F}_{n-1}\right]\right] \end{aligned}$$

$$\begin{aligned}
&\leq \mathbb{E} \left[e^{\beta \mathbf{X}_{n-1}} \mathbb{E} \left[\frac{c_n - (\mathbf{X}_n - \mathbf{X}_{n-1})}{2c_n} e^{-\beta c_n} + \frac{c_n + (\mathbf{X}_n - \mathbf{X}_{n-1})}{2c_n} e^{\beta c_n} \middle| \mathcal{F}_{n-1} \right] \right] \\
&= \mathbb{E} \left[e^{\beta \mathbf{X}_{n-1}} \left[\frac{c_n - \mathbb{E}[\mathbf{X}_n - \mathbf{X}_{n-1} | \mathcal{F}_{n-1}]}{2c_n} e^{-\beta c_n} + \frac{c_n + \mathbb{E}[\mathbf{X}_n - \mathbf{X}_{n-1} | \mathcal{F}_{n-1}]}{2c_n} e^{\beta c_n} \right] \right] \\
&= \mathbb{E} \left[e^{\beta \mathbf{X}_{n-1}} \left(\frac{c_n - 0}{2c_n} e^{-\beta c_n} + \frac{c_n + 0}{2c_n} e^{\beta c_n} \right) \right] \\
&= \mathbb{E} [e^{\beta \mathbf{X}_{n-1}}] \left(\frac{c_n}{2c_n} e^{-\beta c_n} + \frac{c_n}{2c_n} e^{\beta c_n} \right) \\
&= \mathbb{E} [e^{\beta \mathbf{X}_{n-1}}] \left(\frac{e^{-\beta c_n} + e^{\beta c_n}}{2} \right)
\end{aligned}$$

where the third to last line is because $\mathbf{X}_n$ is a martingale, and $\mathbf{X}_0 = 0$. Induction on n now verifies the first part of the theorem, provided we can show that

$$\frac{e^u + e^{-u}}{2} \leq e^{\frac{u^2}{2}}$$

However, this follows by expanding both sides in a Taylor series

$$\begin{aligned}
\frac{e^u + e^{-u}}{2} &= \frac{1}{2} \sum_{n=0}^{\infty} \frac{u^n + (-u)^n}{n!} \\
e^{\frac{u^2}{2}} &= \sum_{n=0}^{\infty} \left(\frac{u^{2n}}{2^n n!} \right)
\end{aligned}$$

and noting that the corresponding coefficients of u^{2n} satisfy $\frac{1}{(2n)!} \leq \frac{1}{2^n n!}$, as can be verified by induction. Also, the coefficients of the odd powers u^{2n+1} on both sides vanish.

To prove the second part of the theorem, we apply Markov's inequality in the form

$$\begin{aligned}
\mathbb{P}[\mathbf{X}_n \geq \lambda] &\leq \mathbb{E} [e^{\beta \mathbf{X}_n}] e^{-\beta \lambda} \\
&\leq e^{\frac{\beta^2}{2} \sum_{k=1}^n c_k^2 - \beta \lambda}
\end{aligned}$$

for arbitrary $\beta > 0$. The particular β that minimizes the exponent on the right hand side is clearly $\beta = \lambda / \sum_{k=1}^n c_k^2$. Substituting this choice into the above yields

the result. Because $-\mathbf{X}_n$ also satisfies the hypothesis of the theorem, the final part of the theorem follows from the second part. $\square$

$$\text{THEOREM 33. } \mathbb{P} \left[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \lambda n^{\frac{3}{2}} \right] \leq 2 \exp \left(-\frac{2\lambda^2}{3} \right)$$

PROOF. Let us define

$$\mathbf{X}_n = \mathbf{I}_n - \mathbb{E}[\mathbf{I}_n],$$

then the sequence of random variables $\mathbf{X}_n$ forms a martingale with respect to a σ -field $\mathcal{F}_n$, and $\mathbf{X}_0 = 0$. We calculate

$$\begin{aligned} \mathbf{X}_n - \mathbf{X}_{n-1} &= \mathbf{I}_n - \mathbb{E}[\mathbf{I}_n] - \mathbf{I}_{n-1} + \mathbb{E}[\mathbf{I}_{n-1}] \\ &= (\mathbf{I}_n - \mathbf{I}_{n-1}) - (\mathbb{E}[\mathbf{I}_n] - \mathbb{E}[\mathbf{I}_{n-1}]) \end{aligned}$$

and note the simple observation

$$\begin{aligned} \mathbb{E}[\mathbf{I}_n] - \mathbb{E}[\mathbf{I}_{n-1}] &= \frac{n(n-1)}{4} - \frac{(n-2)(n-1)}{4} \\ &= \frac{(n-1)}{2} \end{aligned}$$

so

$$\begin{aligned} |(\mathbf{I}_n - \mathbf{I}_{n-1}) - (\mathbb{E}[\mathbf{I}_n] - \mathbb{E}[\mathbf{I}_{n-1}])| &\leq |\mathbf{I}_n - \mathbf{I}_{n-1}| + |\mathbb{E}[\mathbf{I}_n] - \mathbb{E}[\mathbf{I}_{n-1}]| \\ &\leq n-1 + \frac{n-1}{2} \\ &= \frac{3(n-1)}{2} \end{aligned}$$

and thus

$$\mathbf{X}_n - \mathbf{X}_{n-1} \leq \frac{3}{2}(n-1)$$

By the Azuma Hoeffding inequality, we have

$$\mathbb{P}[|\mathbf{X}_n| \geq \lambda] \leq 2e^{-\lambda^2 / \left(2 \sum_{k=1}^n \left(\frac{3}{2}(k-1) \right)^2 \right)}$$

i.e.

$$\begin{aligned}
 \mathbb{P}[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \lambda] &\leq 2 \exp\left(-\frac{\lambda^2}{2 \times \frac{3n}{8} (1 - 3n + 2n^2)}\right) \\
 &= 2 \exp\left(-\frac{4\lambda^2}{3n(1 - 3n + 2n^2)}\right) \\
 &= 2 \exp\left(-\frac{4\lambda^2}{3n - 9n^2 + 6n^3}\right)
 \end{aligned}$$

Rewriting this in a more convenient form we have

$$\begin{aligned}
 \mathbb{P}[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \beta n^\alpha] &\leq 2 \exp\left(-\frac{2(\beta n^\alpha)^2}{3n^3}\right) \\
 &= 2 \exp\left(-\frac{2\beta^2 n^{2\alpha-3}}{3}\right)
 \end{aligned}$$

so

$$\mathbb{P}\left[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \beta n^{\frac{3}{2}}\right] \leq 2 \exp\left(-\frac{2\beta^2}{3}\right)$$

□

Similarly, we have the following

THEOREM 34. $\mathbb{P}[|\mathbf{L}_n - \mathbb{E}[\mathbf{L}_n]| \geq \lambda] \leq 2e^{-\frac{\lambda^2}{4 \log n}}.$

PROOF. We have

$$\mathbb{E}[\mathbf{L}_n] = n - 1 - \sum_{l=1}^{n-1} \left(\frac{l! l^{n-l}}{n!}\right)$$

so

$$\begin{aligned}
\mathbb{E}[\mathbf{L}_n] - \mathbb{E}[\mathbf{L}_{n-1}] &= \left(n - 1 - \sum_{l=1}^{n-1} \left(\frac{l!l^{n-l}}{n!} \right) \right) - \left((n-1) - 1 - \sum_{l=1}^{n-2} \left(\frac{l!l^{n-l-1}}{(n-1)!} \right) \right) \\
&= n - 1 - \sum_{l=1}^{n-1} \left(\frac{l!l^{n-l}}{n!} \right) - n + 2 + \sum_{l=1}^{n-2} \left(\frac{l!l^{n-l-1}}{(n-1)!} \right) \\
&= 1 - \sum_{l=1}^{n-1} \left(\frac{l!l^{n-l}}{n!} \right) + \sum_{l=1}^{n-2} \left(\frac{l!l^{n-l-1}}{(n-1)!} \right) \\
&= 1 - \sum_{l=1}^{n-1} \left(\frac{l!l^{n-l}}{n!} \right) + \sum_{l=1}^{n-2} \left(\frac{n \times l!l^{n-l-1}}{n!} \right) \\
&= 1 - \sum_{l=1}^{n-2} \left(\frac{l!l^{n-l}}{n!} \right) - \left(\frac{(n-1)! (n-1)^{n-(n-1)}}{n!} \right) + \sum_{l=1}^{n-2} \left(\frac{n \times l!l^{n-l-1}}{n!} \right) \\
&= \frac{1}{n} - \sum_{l=1}^{n-2} \frac{l!l^{n-l} + n \times l!l^{n-l-1}}{n!} \\
&\leq \frac{1}{n}
\end{aligned}$$

and so

$$\mathbf{X}_n - \mathbf{X}_{n-1} \leq 2 \left(1 + \frac{1}{n} \right)$$

By the Azuma Hoeffding inequality we have

$$\begin{aligned}
\mathbb{P}[|\mathbf{X}_n| \geq \lambda] &\leq 2e^{-\lambda^2 / (2 \sum_{k=1}^n 2(1 + \frac{1}{k}))} \\
&= 2e^{-\lambda^2 / (4(1 + H_n))}
\end{aligned}$$

where H_n is the n^{th} harmonic number, defined by

$$H_n = \sum_{k=1}^n \frac{1}{k}.$$

Putting it all together, we have

$$\begin{aligned}
\mathbb{P}[|\mathbf{L}_n - \mathbb{E}[\mathbf{L}_n]| \geq \lambda] &\leq 2 \exp \left(\frac{-\lambda^2}{4(1 + H_n)} \right) \\
&\leq 2 \exp \left(\frac{-\lambda^2}{4 \log n} \right)
\end{aligned}$$

□

CHAPTER 4

The Mallows Measure on Random Permutations

In this chapter we study a nonuniform measure on random permutations that allows weighting of permutations according to their number of inversions.

This measure does not differentiate between a permutation with one element displaced a large amount, and a permutation with several elements displaced small amounts, as long as the total number of inversions are the same. For example, the permutations

$$(2, 3, 4, 5, 1)$$

and

$$(3, 1, 5, 2, 4)$$

both have 4 inversions and hence have the same probability. Of course, here we are interested in the level too, and the first permutation has level 4, whereas the second has level 2.

4.1. Mallows Measure

The *Mallows measure*, a probability measure on $\mathbb{S}_n$, is defined as

$$\mathbb{P}_p(\mathbf{X} = x) = \frac{p^{i(x)}}{[n]_p!}$$

where $\mathbf{X}$ is a random permutation if x is a permutation of $[n]$, and 0 otherwise, and where $0 < p$. Note that for $p = 1$, $[n]_p$ is undefined, and so we take the limit as $p \rightarrow 1$ to get $\lim_{p \rightarrow 1} [n]_p = n$, so we see that, as expected,

$$\mathbb{P}_1(\mathbf{X} = x) = \frac{1}{n!},$$

the uniform measure on permutations. For values of $p < 1$ this probability measure assigns more mass to permutations with fewer inversions, and conversely for values of $p > 1$ more probability is assigned to permutations with more inversions.

It is clear that this is in fact a probability measure, as $\sum_{\sigma \in \mathbb{S}_n} p^{i(\sigma)}$ is just the generating function for the number of inversions discussed in Chapter 2, which equals $[n]_p!$.

EXAMPLE 35. Let $p = \frac{1}{2}$ and $n = 3$, then

$$\begin{aligned}
 \mathbb{P}_{\frac{1}{2}}(\mathbf{X} = x) &= \frac{1}{2^i} \frac{1}{[3]_{\frac{1}{2}}!} \\
 &= \frac{1}{2^i} \frac{1}{[3]_{\frac{1}{2}} [2]_{\frac{1}{2}} [1]_{\frac{1}{2}}} \\
 &= \frac{1}{2^i} \frac{(1 - \frac{1}{2})^3}{(1 - \frac{1}{2^3}) (1 - \frac{1}{2^2}) (1 - \frac{1}{2})} \\
 &= \frac{1}{2^i} \frac{(\frac{1}{2})^3}{(\frac{7}{8}) (\frac{3}{4}) (\frac{1}{2})} \\
 &= \frac{2^{3-i}}{21}; \quad i = 0, 1, 2, 3
 \end{aligned}$$

if x has i inversions. Thus,

$$\begin{aligned}
 \mathbb{P}_{\frac{1}{2}}(\mathbf{X} = (1, 2, 3)) &= \frac{8}{21} \\
 \mathbb{P}_{\frac{1}{2}}(\mathbf{X} = (1, 3, 2)) &= \frac{4}{21} \\
 \mathbb{P}_{\frac{1}{2}}(\mathbf{X} = (2, 1, 3)) &= \frac{4}{21} \\
 \mathbb{P}_{\frac{1}{2}}(\mathbf{X} = (2, 3, 1)) &= \frac{2}{21} \\
 \mathbb{P}_{\frac{1}{2}}(\mathbf{X} = (3, 1, 2)) &= \frac{2}{21} \\
 \mathbb{P}_{\frac{1}{2}}(\mathbf{X} = (3, 2, 1)) &= \frac{1}{21}
 \end{aligned}$$

Similarly, let $p = 2$ and $n = 3$, then

$$\begin{aligned}\mathbb{P}_2(\mathbf{X} = x) &= \frac{2^i}{[3]_2!} \\ &= \frac{2^i}{21}; \quad i = 0, 1, 2, 3\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}_2(\mathbf{X} = (1, 2, 3)) &= \frac{1}{21} \\ \mathbb{P}_2(\mathbf{X} = (1, 3, 2)) &= \frac{2}{21} \\ \mathbb{P}_2(\mathbf{X} = (2, 1, 3)) &= \frac{2}{21} \\ \mathbb{P}_2(\mathbf{X} = (2, 3, 1)) &= \frac{4}{21} \\ \mathbb{P}_2(\mathbf{X} = (3, 1, 2)) &= \frac{4}{21} \\ \mathbb{P}_2(\mathbf{X} = (3, 2, 1)) &= \frac{8}{21}\end{aligned}$$

It is interesting to compare the results of these two examples, and note the reversed order of the probabilities of the second example relative to the first. In fact, we have

$$\begin{aligned}\mathbb{P}_p(I(X) = i) &= \frac{p^i}{[n]_p!} \\ &= \frac{\left(\frac{1}{p}\right)^{\frac{n(n-1)}{2} - i}}{[n]_{\frac{1}{p}}!} \\ &= \frac{p^i p^{-\frac{n(n-1)}{2}}}{[n]_{\frac{1}{p}}!} \\ &= p^i p^{-i} \frac{p^i p^{-\frac{n(n-1)}{2}}}{[n]_{\frac{1}{p}}!} \\ &= \frac{p^{-i}}{[n]_{\frac{1}{p}}!} p^{2i - \frac{n(n-1)}{2}} \\ &= p^{2i - \frac{n(n-1)}{2}} \mathbb{P}_{\frac{1}{p}}(i)\end{aligned}$$

However, there does not appear to be any simple relationship between expected values with parameters p and $1/p$.

Also, note that conditional on a permutation having i inversions, the Mallows measure assigns uniform weights.

Next, we show that in general the inversions are not Mahonian. Recall that a permutation statistic is *Mahonian* if it is equidistributed with the major index of a permutation [Bo, p 53], where the major index is defined by

$$maj(\sigma) = \sum_{\sigma(i) > \sigma(i+1)} i.$$

THEOREM 36. *Under the Mallows measure, the inversions are not (in general) Mahonian.*

PROOF. We display a simple counterexample. Let $n = 3$, then we have the following permutations, inversions, major index, and probabilities under the Mallows measure with parameter p :

Permutation	Inversions	Major Index	Probability
(123)	0	0	$\frac{1}{(1+p)(1+p+p^2)}$
(132)	1	2	$\frac{p}{(1+p)(1+p+p^2)}$
(213)	1	1	$\frac{p}{(1+p)(1+p+p^2)}$
(231)	2	2	$\frac{p^2}{(1+p)(1+p+p^2)}$
(312)	2	1	$\frac{p^2}{(1+p)(1+p+p^2)}$
(321)	3	3	$\frac{p^3}{(1+p)(1+p+p^2)}$

So we see that (for example) the probability of two inversions is $\frac{2p^2}{(1+p)(1+p+p^2)}$ whereas the probability of having major index two is $\frac{p+p^2}{(1+p)(1+p+p^2)}$, and so the two are not equal unless $p = 0$ or $p = 1$. $\square$

4.2. Basic Results

In this section we present a useful representation of the inversion table under the Mallows measure.

DEFINITION 37. For each $n \in \mathbb{N}$ define the *truncated geometric distribution*

$$G_{p,n}(j) = \frac{p^{j-1}}{[n]_p}; j \in [n]$$

DEFINITION 38. The *backward ranks* are defined by

$$\beta_j(\sigma) = \#\{i \leq j : \sigma(i) \leq \sigma(j)\}; j = 1, 2, \dots, n$$

For example, the permutation (1324) has $\beta_1 = 1$, $\beta_2 = 2$, $\beta_3 = 2$, and $\beta_4 = 4$.

Note the identity

$$v = \text{seq}(n) - \beta$$

where v is the inversion table of σ , $\text{seq}(n)$ is the table $(1, 2, 3, \dots, n)$ and $\beta = (\beta_1(\sigma), \beta_2(\sigma), \dots, \beta_n(\sigma))$.

The correspondence $\sigma \longleftrightarrow \beta$ is a bijection between $\mathbb{S}_n$ and the Cartesian product $[1] \times [2] \times \dots \times [n]$, and so we have the following important theorem:

THEOREM 39. [GO, Mal] *Under the Mallows measure the backward ranks are independent and each variable $j - \beta_j + 1$ is distributed according to $G_{p,j}$.*

PROOF. Decompose the number of inversions as

$$i(\sigma) = \sum_{j=1}^n (j - \beta_j)$$

and multiply probabilities of the truncated geometric distribution to see that the Mallows measure coincides with the product measure. $\square$

REMARK 40. We have that

$$\begin{aligned} j - \beta_j + 1 &= j + (v_j - j) + 1 \\ &= v_j + 1 \end{aligned}$$

so

$$v_j + 1 \sim G_{p,j}$$

or equivalently

$$\mathbb{P}[v_j + 1 = k] = \frac{p^{k-1}}{[j]_p} \mathbb{I}_{k \in \{1, 2, \dots, j\}}$$

so, by setting $k = m + 1$

$$\begin{aligned} \mathbb{P}[v_j = m] &= \frac{p^m}{[j]_p} \mathbb{I}_{m+1 \in \{1, 2, \dots, j\}} \\ &= \frac{p^m}{[j]_p} \mathbb{I}_{m \in \{0, 1, 2, \dots, j-1\}} \end{aligned}$$

DEFINITION 41. The *shifted truncated geometric distribution* has mass function

$$\tilde{G}_{p,j}(m) = \frac{p^m}{[j]_p} \mathbb{I}_{m \in \{0, 1, 2, \dots, j-1\}}$$

Let us now determine the mean and variance of the components of the inversion table of a random permutation under the Mallows measure.

THEOREM 42. $\mathbb{E}_p[v_j] = p \frac{1 - jp^{j-1} - (1-j)p^j}{(1-p)(1-p^j)}$

PROOF.

$$\begin{aligned} \mathbb{E}_p[v_j] &= \sum_{m=0}^{j-1} \frac{mp^m}{[j]_p} \\ &= \sum_{m=0}^{j-1} \frac{(1-p)mp^m}{1-p^j} \\ &= \frac{p - jp^j + (-1+j)p^{1+j}}{(-1+p)(-1+p^j)} \\ &= p \frac{1 - jp^{j-1} - (1-j)p^j}{(1-p)(1-p^j)} \end{aligned}$$

□

Figure 4.2.1 is a plot of $\mathbb{E}_p[v_{50}]$

FIGURE 4.2.1. $\mathbb{E}_p [v_{50}]$

$$\text{THEOREM 43. } \mathbb{V}_p [v_j] = \frac{p - j^2 p^j + 2(-1 + j^2) p^{j+1} - j^2 p^{2+j} + p^{2j+1}}{(-1+p)^2 (-1+p^j)^2}$$

PROOF.

$$\begin{aligned} \mathbb{E}_p [v_j^2] &= \sum_{m=0}^{j-1} \frac{m^2 p^m}{[j]_p} \\ &= \sum_{m=0}^{j-1} \frac{(1-p) m^2 p^m}{1-p^j} \\ &= \frac{-p - p^2 + j^2 p^j + (1 + 2j - 2j^2) p^{j+1} + (-1 + j)^2 p^{2+j}}{(-1+p)^2 (-1+p^j)} \\ &= \frac{p - j^2 p^j + 2(-1 + j^2) p^{j+1} - j^2 p^{2+j} + p^{2j+1}}{(-1+p)^2 (-1+p^j)^2} \end{aligned}$$

□

It is interesting that although the mean of the inversion table components increases as p increases, the situation is different for the variance, as seen in Figure 4.2.2.

THEOREM 44.

$$\mathbb{E}_p [\mathbf{I}] = \frac{p(n-1)}{1-p} - \sum_{j=1}^{n-1} \frac{j p^j}{1-p^j}$$

FIGURE 4.2.2. $\mathbb{V}_p[v_{20}]$

PROOF.

$$\begin{aligned}
 \mathbb{E}_p[\mathbf{I}] &= \sum_{j=1}^{n-1} \mathbb{E}_p[v_j] \\
 &= \sum_{j=1}^{n-1} p \frac{1 - jp^{j-1} - (1-j)p^j}{(1-p)(1-p^j)} \\
 &= \frac{p}{1-p} \sum_{j=1}^{n-1} \frac{1 - jp^{j-1} - (1-j)p^j}{1-p^j} \\
 &= \frac{p}{1-p} \left(\sum_{j=1}^{n-1} \frac{1}{1-p^j} - \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} - \sum_{j=1}^{n-1} \frac{(1-j)p^j}{1-p^j} \right) \\
 &= \frac{p}{1-p} \left(\sum_{j=1}^{n-1} \frac{1}{1-p^j} - \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} - \sum_{j=1}^{n-1} \frac{p^j}{1-p^j} + \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \right)
 \end{aligned}$$

FIGURE 4.2.3. $\mathbb{E}_p[\mathbf{I}]$ for $n = 50$

$$\begin{aligned}
&= \frac{p}{1-p} \left(\sum_{j=1}^{n-1} \frac{1}{1-p^j} - \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} - \sum_{j=1}^{n-1} \frac{p^j}{1-p^j} + p \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} \right) \\
&= \frac{p}{1-p} \left(\left(\sum_{j=1}^{n-1} \frac{1}{1-p^j} - \sum_{j=1}^{n-1} \frac{p^j}{1-p^j} \right) + \left(- \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} + p \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} \right) \right) \\
&= \frac{p}{1-p} \left(\left(\sum_{j=1}^{n-1} 1 \right) - (1-p) \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} \right) \\
&= \frac{p}{1-p} \left(n-1 - (1-p) \sum_{j=1}^{n-1} \frac{jp^{j-1}}{1-p^j} \right) \\
&= \frac{p(n-1)}{1-p} - \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j}
\end{aligned}$$

□

Figure 4.2.3 displays $\mathbb{E}_p[\mathbf{I}]$.

To determine the level of the permutation, we need to determine the maximum of a collection of several $\tilde{G}$ distributed random variables.

THEOREM 45. *The cumulative distribution of the level $\mathbf{L}$ of a permutation selected from the Mallows distribution with parameter p is*

$$\begin{aligned}\mathbb{P}_p[\mathbf{L} \leq l] &= \frac{[l]_p!}{[n]_p!} [1+l]_p^{n-l} \\ &= \frac{(1-p^{1+l})^{n-l}}{\prod_{k=l+1}^n (1-p^k)}\end{aligned}$$

PROOF. Since each component of the inversion table has independent distribution $\tilde{G}_{p,j}$, we see that the distribution of the level is the distribution of $\max_{1 \leq i \leq n} \mathbf{Y}_i$, where each $\mathbf{Y}_i$ has distribution $\tilde{G}_{p,i}$.

$$\begin{aligned}\mathbb{P}_p[\mathbf{Y}_i \leq x] &= \sum_{k=0}^{\lfloor x \rfloor} \mathbb{P}_p[\mathbf{Y}_i = k] \\ &= \sum_{k=0}^{\lfloor x \rfloor} \frac{p^k}{[i]_p} \mathbb{I}_{k \in \{0,1,2,\dots,i-1\}} \\ &= \frac{1}{[i]_p} \sum_{k=0}^{\lfloor x \rfloor \wedge (i-1)} p^k \\ &= \frac{1}{[i]_p} \frac{1 - p^{1+(\lfloor x \rfloor \wedge (i-1))}}{1 - p} \\ &= \frac{1}{[i]_p} \frac{1 - p^{(1+\lfloor x \rfloor) \wedge i}}{1 - p} \\ &= \frac{[(1 + \lfloor x \rfloor) \wedge i]_p}{[i]_p}\end{aligned}$$

We also have

$$\begin{aligned}\mathbb{P}_p[\mathbf{Y}_i \leq 0] &= \mathbb{P}_p[\mathbf{Y}_i = 0] \\ &= \frac{1}{[i]_p}\end{aligned}$$

$$\begin{aligned}
\mathbb{P}_p[\mathbf{L} \leq x] &= \prod_{i=1}^n \mathbb{P}_p[\mathbf{Y}_i \leq x] \\
&= \prod_{i=1}^n \frac{[(1+x) \wedge i]_p}{[i]_p} \\
&= \frac{\prod_{i=1}^n [(1+x) \wedge i]_p}{\prod_{i=1}^n [i]_p} \\
&= \frac{\prod_{i=1}^n [(1+x) \wedge i]_p}{[n]_p!}
\end{aligned}$$

and

$$\begin{aligned}
\prod_{i=1}^n [(1+x) \wedge i]_p &= [1]_p [2]_p [3]_p \cdots [x]_p [1+x]_p [1+x]_p \cdots [1+x]_p \\
&= [x]_p! [1+x]_p^{n-x}
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathbb{P}_p[\mathbf{L} \leq l] &= \frac{\prod_{i=1}^n [(1+[l]) \wedge i]_p}{[n]_p!} \\
&= \frac{[[l]]_p! [1+[l]]_p^{n-[l]}}{[n]_p!}
\end{aligned}$$

The other form follows by simple algebra. □

COROLLARY 46. *The probability mass function of the level $\mathbf{L}$ of a permutation selected from the Mallows distribution with parameter p is*

$$\mathbb{P}_p[\mathbf{L} = l] = \frac{[l]_p!}{[n]_p!} \left([1+l]_p^{n-l} - [l]_p^{n-l} \right) \mathbb{I}_{l \in \{n-1\}} + \frac{1}{[n]_p!} \mathbb{I}_{l=0}$$

FIGURE 4.2.4. Distribution of the level, $\mathbf{L}$, for $n = 50$

PROOF.

$$\begin{aligned}
 \mathbb{P}_p[L = l] &= \mathbb{P}_p[L \leq l] - \mathbb{P}_p[L \leq l-1] \\
 &= \frac{[l]_p!}{[n]_p!} \left([1 + l]_p \right)^{n-l} - \frac{[l-1]_p!}{[n]_p!} \left([l]_p \right)^{n-l+1} \\
 &= \frac{[l]_p!}{[n]_p!} \left([1 + l]_p^{n-l} - [l]_p^{n-l+1} \right)
 \end{aligned}$$

The above is only valid where $l \geq 1$, and for $l = 0$ we have $\mathbb{P}_p[\mathbf{L} = 0] = \frac{1}{[n]_p!}$. $\square$

Figure 4.2.4 displays the distribution of the level for varying values of p .

THEOREM 47. *The expected value of the level $\mathbf{L}$ of a permutation selected from the Mallows distribution with parameter p is*

$$\begin{aligned}
 \mathbb{E}_p[\mathbf{L}] &= \sum_{l=1}^{n-1} l \frac{[l]_p!}{[n]_p!} \left([1 + l]_p^{n-l} - [l]_p^{n-l+1} \right) \\
 &= n - 1 - \sum_{l=1}^{n-1} \frac{(1 - p^l)^{n-l+1}}{\prod_{k=l}^n (1 - p^k)}
 \end{aligned}$$

PROOF. From the definition of the expected value we have

$$\begin{aligned}
\mathbb{E}_p[\mathbf{L}] &= \sum_{l=0}^{n-1} l \mathbb{P}_p[\mathbf{L} = l] \\
&= \sum_{l=1}^{n-1} l \mathbb{P}_p[\mathbf{L} = l] \\
&= \sum_{l=1}^{n-1} l \frac{[l]_p!}{[n]_p!} \left([1+l]_p^{n-l} - [l]_p^{n-l} \right).
\end{aligned}$$

Also

$$\mathbb{E}_p[\mathbf{L}] = \sum_{l=1}^{n-1} \mathbb{P}_p[\mathbf{L} \geq l]$$

and

$$\begin{aligned}
\mathbb{P}_p[\mathbf{L} \geq l] &= 1 - \mathbb{P}_p[\mathbf{L} < l] \\
&= 1 - \mathbb{P}_p[\mathbf{L} \leq l-1] \\
&= 1 - \frac{(1-p)^n}{(1-p)^{l-1}} \frac{\prod_{k=1}^{l-1} (1-p^k)}{\prod_{k=1}^n (1-p^k)} \left(\frac{1-p^l}{1-p} \right)^{n-l+1} \\
&= 1 - (1-p^l)^{n-l+1} \frac{\prod_{k=1}^{l-1} (1-p^k)}{\prod_{k=1}^n (1-p^k)} \\
&= 1 - (1-p^l)^{n-l+1} \frac{\prod_{k=1}^{l-1} (1-p^k)}{\prod_{k=1}^{l-1} (1-p^k) \prod_{k=l}^n (1-p^k)} \\
&= 1 - \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)}
\end{aligned}$$

So

$$\mathbb{E}_p[\mathbf{L}] = n - 1 - \sum_{l=1}^{n-1} \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)}$$

□

Figure 4.2.5 displays $\mathbb{E}_p[\mathbf{L}]$ for varying values of p .

Summarizing, we have the following:

FIGURE 4.2.5. $\mathbb{E}_p[\mathbf{L}]$ for $n = 50$

THEOREM 48.

$$\begin{aligned}
1) \mathbb{P}_p[\mathbf{I} = i] &= \frac{p^i}{[n]_p!} \\
2) \mathbb{P}_p[\mathbf{L} = l] &= \frac{[l]_p!}{[n]_p!} \left([1+l]_p^{n-l} - [l]_p^{n-l} \right) \mathbb{I}_{l \in [n-1]} + \frac{1}{[n]_p!} \mathbb{I}_{l=0} \\
3) \mathbb{E}_p[\mathbf{I}] &= \frac{p(n-1)}{1-p} - \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \\
4) \mathbb{E}_p[\mathbf{L}] &= n-1 - \sum_{l=1}^{n-1} \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)} \\
5) \mathbb{V}_p[\mathbf{I}] &= \sum_{j=1}^n \frac{p - j^2 p^j + 2(-1+j^2)p^{j+1} - j^2 p^{2+j} + p^{2j+1}}{(-1+p)^2 (-1+p^j)^2}
\end{aligned}$$

PROOF. 1) is the definition; 2), 3), and 4) were proved above; and 5) is due to additivity of variances of independent random variables. $\square$

4.3. Limits

It appears, from the plots in the previous section, that for small p (i.e. $p < 1$) and large n both the inversions and the level are small, and that for large p (i.e. $p > 1$) both the inversions and level are large (large and small being relative to their maximums). Of course, we already know the behavior when $p = 1$ from Chapter 3. In this section we make these observations precise.

THEOREM 49. For fixed $p < 1$ we have that $\mathbb{E}_p[\mathbf{I}] \sim \frac{np}{1-p}$

PROOF. First,

$$\frac{jp^j}{1} \leq \frac{jp^j}{1-p^j} \leq \frac{jp^j}{1-p}$$

so

$$\begin{aligned} \sum_{j=1}^{n-1} jp^j &\leq \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \leq \sum_{j=1}^{n-1} \frac{jp^j}{1-p} \\ \frac{p - np^n + (n-1)p^{n+1}}{(1-p)^2} &\leq \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \leq \frac{p - np^n + (n-1)p^{n+1}}{(1-p)^3} \end{aligned}$$

Now, we see that

$$\frac{p - np^n + (n-1)p^{n+1}}{(1-p)^2} \sim \frac{p}{(1-p)^2}$$

because

$$\begin{aligned} \frac{\frac{p - np^n + (n-1)p^{n+1}}{(1-p)^2}}{\frac{p}{(1-p)^2}} &= \frac{p - np^n + (n-1)p^{n+1}}{(1-p)^2} \cdot \frac{(1-p)^2}{p} \\ &= \frac{p - np^n + (n-1)p^{n+1}}{p} \\ &= \frac{p}{p} - \frac{np^n}{p} + \frac{np^{n+1}}{p} - \frac{p^{n+1}}{p} \\ &= 1 - np^{n-1} + np^n - p^n \end{aligned}$$

and as $n \rightarrow \infty$ we have

$$np^{n-1} \rightarrow 0$$

$$np^n \rightarrow 0$$

$$p^n \rightarrow 0$$

Similarly, we have that

$$\frac{p - np^n + (n-1)p^{n+1}}{(1-p)^3} \sim \frac{p}{(1-p)^3}$$

and

$$\frac{p(n-1)}{1-p} \sim \frac{np}{1-p}$$

so combining the above, we have

$$\begin{aligned} \mathbb{E}_p[\mathbf{I}] &= \frac{p(n-1)}{1-p} - \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \\ &\sim \frac{np}{1-p} \end{aligned}$$

□

THEOREM 50. *For fixed $p > 1$ we have that $\mathbb{E}_p[\mathbf{I}] \sim \frac{n^2}{2}$*

PROOF. We have that

$$\frac{jp^j}{1-p^j} = -j \frac{p^j}{p^j - 1}$$

and

$$\frac{p^j}{p^j - 1} \geq 1$$

So, we seek a J such that for all $j \leq J$ we have

$$\begin{aligned} \frac{p^j}{p^j - 1} &\leq 1 + \varepsilon \\ p^j &\leq (1 + \varepsilon)(p^j - 1) \\ p^j &\leq p^j - 1 + \varepsilon p^j - \varepsilon \end{aligned}$$

so

$$\begin{aligned}
0 &\leq \varepsilon p^j - (1 + \varepsilon) \\
1 + \varepsilon &\leq \varepsilon p^j \\
\frac{1 + \varepsilon}{\varepsilon} &\leq p^j \\
\log\left(\frac{1 + \varepsilon}{\varepsilon}\right) &\leq j \log p \\
\frac{\log\left(\frac{1 + \varepsilon}{\varepsilon}\right)}{\log p} &\leq j
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{j=1}^{n-1} j \frac{p^j}{p^j - 1} &= \sum_{j=1}^{\frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p}} j \frac{p^j}{p^j - 1} + \sum_{j=\frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p}}^{n-1} j \frac{p^j}{p^j - 1} \\
&\leq \sum_{j=1}^{\frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p}} j \frac{p^j}{p^j - 1} + (1 + \varepsilon) \sum_{j=\frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p}}^{n-1} j \\
&= \sum_{j=1}^{\frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p}} j \frac{p^j}{p^j - 1} \\
&\quad + (1 + \varepsilon) \frac{(\log\left(1 + \frac{1}{\varepsilon}\right) + (n-1) \log p) (-\log\left(1 + \frac{1}{\varepsilon}\right) + n \log p)}{2 \log^2 p} \\
&\leq \frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p} \sum_{j=1}^{\frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p}} \frac{p^j}{p^j - 1} \\
&\quad + (1 + \varepsilon) \frac{(\log\left(1 + \frac{1}{\varepsilon}\right) + (n-1) \log p) (-\log\left(1 + \frac{1}{\varepsilon}\right) + n \log p)}{2 \log^2 p}
\end{aligned}$$

we see by making the following definitions to ease notation

$$\begin{aligned}
 K &:= \frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p} \sum_{j=1}^{\frac{\log\left(\frac{1+\varepsilon}{\varepsilon}\right)}{\log p}} \frac{p^j}{p^j - 1} \\
 A &:= \frac{(1+\varepsilon)}{2 \log^2 p} \\
 B &:= \log\left(1 + \frac{1}{\varepsilon}\right) \\
 C &:= \log p \\
 D &:= K + ABC - AB^2 \\
 E &:= AC^2
 \end{aligned}$$

that

$$\begin{aligned}
 \sum_{j=1}^{n-1} j \frac{p^j}{p^j - 1} &\leq K + (1+\varepsilon) \frac{(\log(1 + \frac{1}{\varepsilon}) + (n-1) \log p) (-\log(1 + \frac{1}{\varepsilon}) + n \log p)}{2 \log^2 p} \\
 &\leq K + A(B + (n-1)C)(-B + nC) \\
 &\leq K - AB^2 + nABC + -nABC + ABC + n^2 AC^2 - nAC^2 \\
 &\leq K - AB^2 + ABC + (n-1)nE \\
 &\leq D - nE + n^2 E
 \end{aligned}$$

and we know

$$\begin{aligned}
 E &= AC^2 \\
 &\geq \frac{1}{2}
 \end{aligned}$$

We also have

$$\frac{p^j}{p^j - 1} \geq 1$$

so

$$\begin{aligned} \sum_{j=1}^{n-1} \frac{jp^j}{p^j-1} &\geq \sum_{j=1}^{n-1} j \\ &= \frac{n(n-1)}{2} \end{aligned}$$

Thus

$$\begin{aligned} \frac{n(n-1)}{2} &\leq \sum_{j=1}^{n-1} \frac{jp^j}{p^j-1} \leq D - nE + n^2E \\ -\frac{n(n-1)}{2} &\geq \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \geq -D + nE - n^2E \\ \frac{n(n-1)}{2} &\leq -\sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \leq D - nE + n^2E \\ \frac{p(n-1)}{1-p} + \frac{n(n-1)}{2} &\leq \frac{p(n-1)}{1-p} - \sum_{j=1}^{n-1} \frac{jp^j}{1-p^j} \leq \frac{p(n-1)}{1-p} + D - nE + n^2E \\ \frac{p(n-1)}{1-p} + \frac{n(n-1)}{2} &\leq \mathbb{E}_p[\mathbf{I}] \leq \frac{p(n-1)}{1-p} + D - nE + n^2E \end{aligned}$$

and thus $\mathbb{E}_p[\mathbf{I}] \sim \frac{n^2}{2}$

□

We next turn to $\mathbb{E}_p[\mathbf{L}_n]$ for $p < 1$. After several lemmata, we will show that

$$\mathbb{E}_p[\mathbf{L}_n] = o(n)$$

LEMMA 51. If $0 < \varepsilon < 1$ and $l \geq \frac{\log\left(1-(1-\varepsilon)^{\frac{1}{n+1}}\right)}{\log p}$ then $\prod_{k=l}^n (1-p^k) \geq 1-\varepsilon$

PROOF.

$$\begin{aligned}
l &\geq \frac{\log \left(1 - (1 - \varepsilon)^{\frac{1}{n+1}} \right)}{\log p} \\
l \log p &\leq \log \left(1 - (1 - \varepsilon)^{\frac{1}{n+1}} \right) \\
p^l &\leq 1 - (1 - \varepsilon)^{\frac{1}{n+1}} \\
-p^l &\geq (1 - \varepsilon)^{\frac{1}{n+1}} - 1 \\
1 - p^l &\geq (1 - \varepsilon)^{\frac{1}{n+1}} \\
1 - p^l &\geq \exp \left(\frac{\log (1 - \varepsilon)}{n+1} \right) \\
\log (1 - p^l) &\geq \frac{\log (1 - \varepsilon)}{n+1} \\
(n+1) \log (1 - p^l) &\geq \log (1 - \varepsilon)
\end{aligned}$$

Now

$$(n - l + 1) \log (1 - p^l) \geq (n + 1) \log (1 - p^l)$$

and

$$\sum_{k=l}^n \log (1 - p^k) \geq (n - l + 1) \log (1 - p^l)$$

so

$$\prod_{k=l}^n (1 - p^k) \geq 1 - \varepsilon$$

□

Thus, by letting $l^* = \frac{\log \left(1 - (1 - \varepsilon)^{\frac{1}{n+1}} \right)}{\log p}$ we have that

$$\sum_{l=l^*+1}^{n-1} (1 - p^l)^{n-l+1} \leq \sum_{l=l^*+1}^{n-1} \frac{(1 - p^l)^{n-l+1}}{\prod_{k=l}^n (1 - p^k)} \leq \frac{1}{1 - \varepsilon} \sum_{l=l^*+1}^{n-1} (1 - p^l)^{n-l+1}$$

so we now bound $\sum_{l=l^*+1}^{n-1} (1 - p^l)^{n-l+1}$.

$$\text{LEMMA 52. } 0 \leq \sum_{l=1}^{l^*} \frac{(1 - p^l)^{n-l+1}}{\prod_{k=l}^n (1 - p^k)} \leq l^*$$

PROOF. $(1 - p^x)$ is a monotone increasing function of x , so

$$\begin{aligned} \prod_{k=l}^n (1 - p^n) &> \prod_{k=l}^n (1 - p^k) > \prod_{k=l}^n (1 - p^l) \\ (1 - p^n)^{n+1-l} &> \prod_{k=l}^n (1 - p^k) > (1 - p^l)^{n+1-l} \\ \frac{1}{(1 - p^n)^{n+1-l}} &< \frac{1}{\prod_{k=l}^n (1 - p^k)} < \frac{1}{(1 - p^l)^{n+1-l}} \end{aligned}$$

so

$$\begin{aligned} \sum_{l=1}^{l^*} \frac{(1 - p^l)^{n-l+1}}{(1 - p^n)^{n+1-l}} &< \sum_{l=1}^{l^*} \frac{(1 - p^l)^{n-l+1}}{\prod_{k=l}^n (1 - p^k)} < \sum_{l=1}^{l^*} \frac{(1 - p^l)^{n-l+1}}{(1 - p^l)^{n+1-l}} \\ \sum_{l=1}^{l^*} \left(\frac{1 - p^l}{1 - p^n} \right)^{n-l+1} &< \sum_{l=1}^{l^*} \frac{(1 - p^l)^{n-l+1}}{\prod_{k=l}^n (1 - p^k)} < \sum_{l=1}^{l^*} 1 \end{aligned}$$

Of course we have that $\sum_{l=1}^{l^*} 1 = l^*$, and since

$$0 < \frac{1 - p^l}{1 - p^n} < 1$$

we have

$$0 < \sum_{l=1}^{l^*} \left(\frac{1 - p^l}{1 - p^n} \right)^{n-l+1} < l^*$$

□

LEMMA 53. If $0 < \varepsilon < 1$ and $l < \frac{\log(1-\sqrt{\varepsilon})}{\log p}$ then $(1 - p^l)^{n-l+1} < \varepsilon$

PROOF. Note that for $1 \leq l \leq n - 1$ we have

$$2 \leq (n - l + 1) \leq n$$

and so

$$n \log(1 - p^l) < (n - l + 1) \log(1 - p^l) < 2 \log(1 - p^l).$$

Now, if

$$\begin{aligned}
 l &< \frac{\log(1 - \sqrt{\varepsilon})}{\log p} \\
 l &< \frac{\log\left(1 - \exp\left(\frac{\log \varepsilon}{2}\right)\right)}{\log p} \\
 l \log p &> \log\left(1 - \exp\left(\frac{\log \varepsilon}{2}\right)\right) \\
 p^l &> 1 - \exp\left(\frac{\log \varepsilon}{2}\right) \\
 -p^l &< \exp\left(\frac{\log \varepsilon}{2}\right) - 1 \\
 1 - p^l &< \exp\left(\frac{\log \varepsilon}{2}\right) \\
 \log(1 - p^l) &< \frac{\log \varepsilon}{2} \\
 2 \log(1 - p^l) &< \log \varepsilon
 \end{aligned}$$

and so

$$\begin{aligned}
 (n - l + 1) \log(1 - p^l) &< \log \varepsilon \\
 \log(1 - p^l)^{n-l+1} &< \log \varepsilon \\
 (1 - p^l)^{n-l+1} &< \varepsilon
 \end{aligned}$$

□

Now we want to know for which l we have that $(1 - p^l)^{n-l+1} > 1 - \varepsilon$.

LEMMA 54. *If $0 < \varepsilon < 1$ and $l > \frac{\log(1 - (1 - \varepsilon)^{1/n})}{\log p}$ then $(1 - p^l)^{n-l+1} > 1 - \varepsilon$*

PROOF.

$$\begin{aligned}
l &> \frac{\log \left(1 - (1 - \varepsilon)^{1/n} \right)}{\log p} \\
l &> \frac{\log \left(1 - \exp \left(\frac{\log(1-\varepsilon)}{n} \right) \right)}{\log p} \\
l \log p &< \log \left(1 - \exp \left(\frac{\log(1-\varepsilon)}{n} \right) \right) \\
p^l &< 1 - \exp \left(\frac{\log(1-\varepsilon)}{n} \right) \\
-p^l &> \exp \left(\frac{\log(1-\varepsilon)}{n} \right) - 1 \\
1 - p^l &> \exp \left(\frac{\log(1-\varepsilon)}{n} \right) \\
\log(1 - p^l) &> \frac{\log(1-\varepsilon)}{n} \\
n \log(1 - p^l) &> \log(1-\varepsilon) \\
(n - l + 1) \log(1 - p^l) &> \log(1-\varepsilon) \\
(1 - p^l)^{n-l+1} &> 1 - \varepsilon
\end{aligned}$$

□

Now, how many terms are between ε and $1 - \varepsilon$? These are the terms greater than $\frac{\log(1-\sqrt{\varepsilon})}{\log p}$ and less than $\frac{\log(1-(1-\varepsilon)^{1/n})}{\log p}$, i.e. there are

$$\frac{\log \left(1 - (1 - \varepsilon)^{1/n} \right)}{\log p} - \frac{\log(1 - \sqrt{\varepsilon})}{\log p} = \frac{\log \left(1 - (1 - \varepsilon)^{1/n} \right) - \log(1 - \sqrt{\varepsilon})}{\log p}$$

of them. So, now putting all the pieces together, we have the following.

We split the sum $S := \sum_{l=1}^{n-1} \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)}$ into two sums, call them S_1 and S_2 where

$$\begin{aligned}
S_1 &= \sum_{l=1}^{l^*} \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)} \\
S_2 &= \sum_{l=l^*+1}^{n-1} \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)}
\end{aligned}$$

and found the bound

$$0 \leq S_1 \leq \frac{\log \left(1 - (1 - \varepsilon)^{\frac{1}{n+1}} \right)}{\log p}$$

and

$$\sum_{l=l^*+1}^{n-1} (1 - p^l)^{n-l+1} \leq S_2 \leq \frac{1}{1 - \varepsilon} \sum_{l=l^*+1}^{n-1} (1 - p^l)^{n-l+1}$$

Then we split $\sum_{l=l^*+1}^{n-1} (1 - p^l)^{n-l+1}$ into three sums, expand the range of summation to start at 1, and call them T_1 , T_2 and T_3 defined by

$$\begin{aligned} T_1 &= \sum_{l=1}^{\frac{\log(1-\sqrt{\varepsilon})}{\log p}} (1 - p^l)^{n-l+1} \\ T_2 &= \sum_{l=\frac{\log(1-\sqrt{\varepsilon})}{\log p}}^{\frac{\log(1-(1-\varepsilon)^{1/n})}{\log p}} (1 - p^l)^{n-l+1} \\ T_3 &= \sum_{l=\frac{\log(1-(1-\varepsilon)^{1/n})}{\log p}}^{n-1} (1 - p^l)^{n-l+1} \end{aligned}$$

which have bounds

$$\begin{aligned} 0 &\leq T_1 \leq \frac{\varepsilon \log(1 - \sqrt{\varepsilon})}{\log p} \\ \varepsilon \frac{\log(1 - (1 - \varepsilon)^{1/n}) - \log(1 - \sqrt{\varepsilon})}{\log p} &\leq T_2 \leq (1 - \varepsilon) \frac{\log(1 - (1 - \varepsilon)^{1/n}) - \log(1 - \sqrt{\varepsilon})}{\log p} \\ \left(n - 1 - \frac{\log(1 - (1 - \varepsilon)^{1/n})}{\log p} \right) (1 - \varepsilon) &\leq T_3 \leq n - 1 - \frac{\log(1 - (1 - \varepsilon)^{1/n})}{\log p} \end{aligned}$$

Thus we have that

$$\begin{aligned} \varepsilon \frac{\log(1 - (1 - \varepsilon)^{1/n}) - \log(1 - \sqrt{\varepsilon})}{\log p} + \left(n - 1 - \frac{\log(1 - (1 - \varepsilon)^{1/n})}{\log p} \right) (1 - \varepsilon) &\leq S_2 \\ (n - 1)(1 - \varepsilon) + \frac{(2\varepsilon - 1) \log(1 - (1 - \varepsilon)^{1/n}) - \varepsilon \log(1 - \sqrt{\varepsilon})}{\log p} &\leq S_2 \end{aligned}$$

and

$$\begin{aligned}
S_2 &\leq \frac{\varepsilon \log(1 - \sqrt{\varepsilon})}{\log p} + (1 - \varepsilon) \frac{\log(1 - (1 - \varepsilon)^{1/n}) - \log(1 - \sqrt{\varepsilon})}{\log p} \\
&\quad + n - 1 - \frac{\log(1 - (1 - \varepsilon)^{1/n})}{\log p} \\
S_2 &\leq n - 1 + \frac{(2\varepsilon - 1) \log(1 - \sqrt{\varepsilon}) - \varepsilon \log(1 - (1 - \varepsilon)^{1/n})}{\log p}
\end{aligned}$$

and so we have that

$$(n - 1)(1 - \varepsilon) + \frac{(2\varepsilon - 1) \log(1 - (1 - \varepsilon)^{1/n}) - \varepsilon \log(1 - \sqrt{\varepsilon})}{\log p} \leq S$$

and

$$\begin{aligned}
S &\leq \frac{\log(1 - (1 - \varepsilon)^{\frac{1}{n+1}})}{\log p} + n - 1 + \frac{(2\varepsilon - 1) \log(1 - \sqrt{\varepsilon}) - \varepsilon \log(1 - (1 - \varepsilon)^{1/n})}{\log p} \\
S &\leq n - 1 + \frac{(1 - \varepsilon) \log(1 - (1 - \varepsilon)^{\frac{1}{n}}) + (2\varepsilon - 1) \log(1 - \sqrt{\varepsilon})}{\log p}
\end{aligned}$$

Recall that

$$\begin{aligned}
\mathbb{E}_p[\mathbf{L}_n] &= n - 1 - S \\
n - 1 - \mathbb{E}_p[\mathbf{L}_n] &= S
\end{aligned}$$

so

$$\begin{aligned}
(n - 1)(1 - \varepsilon) + \frac{(2\varepsilon - 1) \log(1 - (1 - \varepsilon)^{1/n}) - \varepsilon \log(1 - \sqrt{\varepsilon})}{\log p} &\leq n - 1 - \mathbb{E}_p[\mathbf{L}_n] \\
&\text{and} \\
n - 1 + \frac{(1 - \varepsilon) \log(1 - (1 - \varepsilon)^{\frac{1}{n}}) + (2\varepsilon - 1) \log(1 - \sqrt{\varepsilon})}{\log p} &\geq n - 1 - \mathbb{E}_p[\mathbf{L}_n]
\end{aligned}$$

$$-\varepsilon(n-1) + \frac{(2\varepsilon-1)\log\left(1-(1-\varepsilon)^{1/n}\right) - \varepsilon\log(1-\sqrt{\varepsilon})}{\log p} \leq -\mathbb{E}_p[\mathbf{L}_n]$$

and

$$\frac{(1-\varepsilon)\log\left(1-(1-\varepsilon)^{\frac{1}{n}}\right) + (2\varepsilon-1)\log(1-\sqrt{\varepsilon})}{\log p} \geq -\mathbb{E}_p[\mathbf{L}_n]$$

$$\frac{-(1-\varepsilon)\log\left(1-(1-\varepsilon)^{\frac{1}{n}}\right) - (2\varepsilon-1)\log(1-\sqrt{\varepsilon})}{\log p} \leq \mathbb{E}_p[\mathbf{L}_n]$$

and

$$\varepsilon(n-1) - \frac{(2\varepsilon-1)\log\left(1-(1-\varepsilon)^{1/n}\right) - \varepsilon\log(1-\sqrt{\varepsilon})}{\log p} \geq \mathbb{E}_p[\mathbf{L}_n]$$

Let $K = \frac{\log(1-\sqrt{\varepsilon})}{\log p}$ and $C = \frac{(2\varepsilon-1)}{\log p}$ then we have

$$\mathbb{E}_p[\mathbf{L}_n] \leq \varepsilon(n-1) - C\log\left(1-(1-\varepsilon)^{1/n}\right) - \varepsilon K$$

Since $\log(1-x) \sim -x$, we have that

$$\log\left(1-(1-\varepsilon)^{1/n}\right) \sim -(1-\varepsilon)^{1/n}$$

and

$$\mathbb{E}_p[\mathbf{L}_n] \leq \varepsilon(n-1) + C(1-\varepsilon)^{1/n} - \varepsilon K$$

As $\varepsilon > 0$ is arbitrary, we see that

$$\mathbb{E}_p[\mathbf{L}_n] = o(n)$$

THEOREM 55. *For fixed $p > 1$ we have that $\mathbb{E}_p[\mathbf{L}_n] \sim n - \sqrt{\frac{\pi}{2\log p}}$*

PROOF. First, note that

$$\begin{aligned} \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)} &= \frac{(-1)^{n-l+1} (p^l-1)^{n-l+1}}{(-1)^{n-l+1} \prod_{k=l}^n (p^k-1)} \\ &= \frac{(p^l-1)^{n-l+1}}{\prod_{k=l}^n (p^k-1)} \end{aligned}$$

and because $p > 1$,

$$\begin{aligned} \prod_{k=l}^n (p^k-1) &\sim \prod_{k=l}^n p^k \\ &= p^{\frac{1}{2}(l-l^2+n+n^2)} \end{aligned}$$

and

$$\begin{aligned} (p^l-1)^{n-l+1} &\sim (p^l)^{n-l+1} \\ &= p^{ln-l^2+l} \end{aligned}$$

We have

$$\begin{aligned} \frac{p^{ln-l^2+l}}{p^{\frac{1}{2}(l-l^2+n+n^2)}} &= p^{ln-l^2+l-\frac{1}{2}(l-l^2+n+n^2)} \\ &= p^{\frac{l}{2}-\frac{l^2}{2}} p^{-\frac{n}{2}-\frac{n^2}{2}} p^{ln} \\ &= p^{-\frac{1}{2}(n-l)(n-l-1)} \\ &\sim p^{-\frac{1}{2}(n-l)^2} \end{aligned}$$

and then

$$\begin{aligned}
n-1-\sum_{l=1}^{n-1} \frac{(1-p^l)^{n-l+1}}{\prod_{k=l}^n (1-p^k)} &\sim n-\sum_{l=1}^n p^{-\frac{(n-l)^2}{2}} \\
&\sim n-\frac{\sqrt{\frac{\pi}{2}}}{\sqrt{\log p}} \operatorname{Erf}\left(\frac{(n-1)\sqrt{\log p}}{\sqrt{2}}\right) \\
&\sim n-\sqrt{\frac{\pi}{2\log p}}
\end{aligned}$$

by approximating the sum by an integral.

Specifically, to evaluate $\sum_{l=1}^n p^{-\frac{(n-l)^2}{2}}$ we use the Euler Maclaurin summation formula [SF, p 184]:

$$\int_1^N f(x) dx = \sum_{1 \leq k \leq N} f(k) - \frac{1}{2} (f(1) + f(N)) - \int_1^N \left(\{x\} - \frac{1}{2}\right) f'(x) dx$$

in our case, $f(x) = p^{-(n-x)^2/2}$, so $f'(x) = (n-x)p^{-(n-x)^2/2} \log p$, thus

$$\begin{aligned}
\int_1^n p^{-(n-x)^2/2} dx &= \sum_{1 \leq k \leq n} p^{-(n-k)^2/2} \\
&\quad - \frac{1}{2} \left(p^{-(n-1)^2/2} + p^{-(n-n)^2/2} \right) \\
&\quad - \int_1^n \left(\{x\} - \frac{1}{2} \right) (n-x) p^{-(n-x)^2/2} \log p dx
\end{aligned}$$

Recalling the error function

$$\operatorname{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

we have

$$\begin{aligned}
\sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) &= \sum_{1 \leq k \leq n} p^{-(n-k)^2/2} - \frac{1}{2} \left(p^{-(n-1)^2/2} + 1 \right) \\
&\quad - \log p \int_1^n \left(\{x\} - \frac{1}{2} \right) (n-x) p^{-(n-x)^2/2} dx \\
\sum_{1 \leq k \leq n} p^{-(n-k)^2/2} &= \sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) \\
&\quad + \frac{1}{2} \left(p^{-(n-1)^2/2} + 1 \right) \\
&\quad + \log p \int_1^n \left(\{x\} - \frac{1}{2} \right) (n-x) p^{-(n-x)^2/2} dx
\end{aligned}$$

$$\frac{\sum_{1 \leq k \leq n} p^{-(n-k)^2/2} - \sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) - \frac{1}{2} \left(p^{-(n-1)^2/2} + 1 \right)}{\log p} = \int_1^n \left(\{x\} - \frac{1}{2} \right) (n-x) p^{-(n-x)^2/2} dx$$

$$-\frac{1}{2} \int_1^n (n-x) p^{-(n-x)^2/2} dx \leq \int_1^n \left(\{x\} - \frac{1}{2} \right) (n-x) p^{-(n-x)^2/2} dx$$

and

$$\int_1^n \left(\{x\} - \frac{1}{2} \right) (n-x) p^{-(n-x)^2/2} dx \leq \frac{1}{2} \int_1^n (n-x) p^{-(n-x)^2/2} dx$$

so

$$-\frac{1 - p^{-\frac{1}{2}(n-1)^2}}{\log p} \leq \int_1^n \left(\{x\} - \frac{1}{2} \right) (n-x) p^{-(n-x)^2/2} dx \leq \frac{1 - p^{-\frac{1}{2}(n-1)^2}}{2 \log p}$$

Thus

$$\sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) + \frac{1}{2} \left(p^{-(n-1)^2/2} + 1 \right) - \frac{1 - p^{-\frac{1}{2}(n-1)^2}}{2} \leq \sum_{1 \leq k \leq n} p^{-(n-k)^2/2}$$

and

$$\sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) + \frac{1}{2} \left(p^{-(n-1)^2/2} + 1 \right) + \frac{1 - p^{-\frac{1}{2}(n-1)^2}}{2} \geq \sum_{1 \leq k \leq n} p^{-(n-k)^2/2}$$

$$\begin{aligned} \sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) + \frac{1}{2} \left(p^{-(n-1)^2/2} + 1 \right) - \frac{1 - p^{-\frac{1}{2}(n-1)^2}}{2} = \\ \sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) + p^{-\frac{1}{2}(n-1)^2} \end{aligned}$$

As $n \rightarrow \infty$ $\operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) \rightarrow 1$, so we have

$$\begin{aligned} \sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) + \frac{1}{2} \left(p^{-(n-1)^2/2} + 1 \right) + \frac{1 - p^{-\frac{1}{2}(n-1)^2}}{2} = \\ 1 + \sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) \\ \leq 1 + \sqrt{\frac{\pi}{2 \log p}} \end{aligned}$$

Thus

$$\sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) + p^{-\frac{1}{2}(n-1)^2} \leq \sum_{1 \leq k \leq n} p^{-(n-k)^2/2} \leq 1 + \sqrt{\frac{\pi}{2 \log p}}$$

and so

$$n - 1 - \sqrt{\frac{\pi}{2 \log p}} \operatorname{Erf} \left(\frac{(n-1) \sqrt{\log p}}{\sqrt{2}} \right) - p^{-\frac{1}{2}(n-1)^2} \geq \mathbb{E}_p [\mathbf{L}_n]$$

and

$$nn - 2 - \sqrt{\frac{\pi}{2 \log p}} \leq \mathbb{E}_p [\mathbf{L}_n]$$

□

Summarizing this section, we see that for large n the inversions and level are both small as compared to their maximums for small p ($p < 1$) and large as compared to their maximums for large p ($p > 1$), thus explaining figures 4.2.3 and 4.2.5.

4.4. Sampling

4.4.1. Sampling from the Mallows measure.

4.4.1.1. *Propp Wilson Algorithm.* This algorithm, due to [PW], generates a sample from the Mallows distribution with parameter p .

- (1) Let $X_0 = \iota$, the identity permutation
- (2) Select any adjacent pair of elements uniformly at random
- (3) Toss a coin with probability of landing heads $\frac{1}{1+p}$
- (4) If the coin comes up heads, order the two selected elements from smallest to largest. If the coin comes up tails, order the elements from largest to smallest.
- (5) Repeat steps 2-4 many times.

This algorithm can also be adapted to yield a sample uniformly from the set of all permutations with i inversions by running the algorithm, and discarding the samples until one permutation has i inversions. However, this is quite wasteful. For example, assume we require a permutation with length 100 and 2500 inversions. The value of p that yields the largest probability of obtaining such a permutation is about 1.00088. But the probability that a permutation of $[n]$ generated by this algorithm has i inversions is $\frac{\phi(n,i)p^i}{[n]_p}$, only about 2.3×10^{-3} in this case.

4.4.1.2. *Gnedin Algorithm.* A simpler algorithm, based on the representation theorem (Theorem 28), is to simply choose backwards ranks according to the shifted truncated geometric distribution. This is simple and efficient, but still does not provide an easy way to get a permutation with exactly i inversions.

4.4.2. Sampling uniformly from permutations with i inversions. In the course of this research, it became desirable to be able to sample uniformly from

permutations with exactly i inversions. Arndt's thesis [Ar] lists this as an open problem. Here we present two new algorithms to solve this problem.

4.4.2.1. *Algorithm 1.* In this section we present a new algorithm for sampling uniformly from the set of permutations with exactly i inversions.

This is an MCMC algorithm, and although no analysis of convergence time has yet been completed, the running time of the algorithm is better than that of the naive algorithm mentioned above.

The algorithm works by starting at a known configuration with the correct number (i) of inversions, and then moving to neighboring configurations that maintain the number of inversions. After sufficiently many steps, we return the current configuration as the random permutation with i inversions.

First, we need a starting configuration, and this can be obtained by simply choosing any permutation with i inversions. To make the choice concrete, we choose an inversion table such that all i inversions are put into the leftmost possible places. For example, if $n = 5$ and $i = 3$, then the corresponding inversion table is $(0, 1, 2, 0, 0)$, or if $i = 4$, then the inversion table is $(0, 1, 2, 1, 0)$.

Next, we need to define the neighbors of a permutation. Consider an inversion table $v = (v_1, v_2, \dots, v_n)$. An inversion table is its neighbor if 1) it is an inversion table, and 2) it can be obtained from v by decrementing one of its components and incrementing another. For example, $(0, 1, 2, 1, 0)$ and $(0, 0, 2, 1, 1)$ are neighbors, but $(0, 1, 2, 1, 0)$ and $(1, 0, 2, 1, 0)$ are not (because the latter is not an inversion table), nor are $(0, 0, 2, 1, 0)$ and $(0, 1, 2, 1, 2)$, because they do not differ by a decrement/increment operation.

Note that it is not feasible to compute all the neighbors of all permutations (of size n) because there are, for large n , too many. So, we need a way to construct the neighbors of a permutation (inversion table) in an online manner. We do this by choosing, as the target of the decrement, any component that is greater than zero, and as the target of the increment any component k that is less than $k - 1$.

For example, starting with the inversion table $(0, 1, 2, 1, 0)$ we choose the next step from the set $\{(0, 0, 2, 2, 0), (0, 0, 2, 1, 1), (0, 1, 1, 2, 0), (0, 1, 1, 1, 1), (0, 1, 2, 0, 1)\}$.

It is obvious that it is possible to move from any permutation with i inversions to any other inversion with i inversions by some sequence of neighbors. So, the next question is to determine the probabilities with which the neighbors are chosen. As we know that the steady state distribution of a random walk on a graph is uniform if the next edge is chosen such that the probability of going from node j to node k is $\deg(j)/\deg(k)$ [LPW], the problem becomes one of determining the degrees of the various nodes (configurations), and this is simply the cardinality of the set of neighbors.

Let us look at the example of our favorite permutation $(3, 1, 4, 5, 2)$ which has inversion table $(0, 1, 0, 0, 3)$, and so has 4 inversions. So, we seek a permutation that is uniformly distributed on all permutations with 4 inversions. There are 20 permutations on $[5]$ with 4 inversions. Their inversion tables are as follows:

(00004)	(00013)	(00022)	(00031)
(00103)	(00112)	(00121)	(00130)
(00202)	(00211)	(00220)	(01003)
(01012)	(01021)	(01030)	(01102)
(01111)	(01120)	(01201)	(01210)

We can consider two graphs on this set of vertices. In the first graph, each vertex is connected to every other vertex that results from a decrement in a component of one vertex and an increment in a component of the other vertex. For example, vertices (00112) and (0111) are connected because the 2 is decremented and the second 0 is incremented. We call this Algorithm 1-A In the second graph, the same rules apply, but now the two components must be adjacent. Thus vertices (00112) and (0111) are not connected because the two components to be changed are the second and fifth, but (00112) and (00121) are. We call this Algorithm 1-B

Let us take a look at the transition matrix of a simpler example, permutations on $[4]$ with 3 inversions. Explicitly, they are

$$(3214), (2341), (2413), (4123), (1432), (3142).$$

Their inversion tables are, respectively,

$$(0120), (0003), (0021), (0111), (0012), (0102).$$

If we look at the all neighbors version (Algorithm 1-A) we have the following transition matrix:

	(0003)	(0012)	(0021)	(0102)	(0111)	(0120)
(0003)	$\frac{5}{12}$	$\frac{1}{4}$	0	$\frac{1}{3}$	0	0
(0012)	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	0
(0021)	0	$\frac{1}{4}$	$\frac{1}{6}$	0	$\frac{1}{4}$	$\frac{1}{3}$
(0102)	$\frac{1}{3}$	$\frac{1}{4}$	0	$\frac{1}{6}$	$\frac{1}{4}$	0
(0111)	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{4}$
(0120)	0	0	$\frac{1}{3}$	0	$\frac{1}{4}$	$\frac{5}{12}$

If we look at the nearest neighbors version (Algorithm 1-B), we have the following transition matrix:

	(0003)	(0012)	(0021)	(0102)	(0111)	(0120)
(0003)	$\frac{2}{3}$	$\frac{1}{3}$	0	0	0	0
(0012)	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	0	0
(0021)	0	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	0
(0102)	0	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	0
(0111)	0	0	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$
(0120)	0	0	0	0	$\frac{1}{3}$	$\frac{2}{3}$

Of course, both have stationary distribution $(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$, as both are doubly stochastic.

4.4.2.2. *Algorithm 2.* Again, let us start with the permutation (31452) which has inversion table (01003), 4 inversions and level 3. If we transpose two adjacent elements, then the number of inversions will increase or decrease by 1. This suggests the next algorithm: choose any adjacent pair of elements uniformly at random and transpose them, resulting in a new permutation. Without loss of generality, assume that this increases the number of inversions. Now, from this new permutation, choose any pair of elements uniformly at random and transpose them. If this last transposition decreases the number of inversions (i.e. returns the number of inversions to the starting point) accept it, otherwise try this last step again until it does. Note that there may be only one such transposition that works, i.e. that returns the permutation to its starting point. Now repeat this process a large number of times. This process yields a uniformly distributed sample from the set of all permutations with the same number of inversions as the starting permutation.

Let us take a look at a simple example, permutations on $[4]$ with 3 inversions. Explicitly, they are

$$(3214), (2341), (2413), (4123), (1432), (3142).$$

Starting from, for example (3214), we can go to (2314), (3124), or (3241) by swaps, each with probability $\frac{1}{3}$, the first two decreasing the number of inversions, and the last increasing the inversions. From (2314) we can go to (3214) or (2341), each with probability $\frac{1}{2}$. From (3124) we can go to (3214) or (3142), each with probability $\frac{1}{2}$. Finally, from (3241) we can go to (3214) or (2341), each with probability $\frac{1}{2}$. Thus, from (3214) we can end up at (3214) with probability $\frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2}$, and (2341) with probability $\frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{3}$, (3142) with probability $\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$.

Similarly, we can construct the following (symmetric) transition matrix.

	(1432)	(2341)	(2413)	(3142)	(3214)	(4123)
(1432)	$\frac{1}{2}$	0	0	$\frac{1}{6}$	0	$\frac{1}{3}$
(2341)	0	$\frac{1}{2}$	$\frac{1}{6}$	0	$\frac{1}{3}$	0
(2413)	0	$\frac{1}{6}$	$\frac{2}{3}$	0	0	$\frac{1}{6}$
(3142)	$\frac{1}{6}$	0	0	$\frac{2}{3}$	$\frac{1}{6}$	0
(3214)	0	$\frac{1}{3}$	0	$\frac{1}{6}$	$\frac{1}{2}$	0
(4123)	$\frac{1}{3}$	0	$\frac{1}{6}$	0	0	$\frac{1}{2}$

For a permutation on $[n]$ with i inversions, this matrix will be of dimension $\phi(n, i) \times \phi(n, i)$, and if $i = 0, \frac{n(n-1)}{2}$ then the rows will have only one entry, corresponding to the fact that there is only one permutation of size n with 0 inversions and one permutation of size n with $\frac{n(n-1)}{2}$ inversions. If $i \neq 0, \frac{n(n-1)}{2}$, then the situation is more complicated.

Since a Markov chain with a symmetric transition matrix has a uniform stationary distribution, we only need to show that this matrix is symmetric in general.

LEMMA 56. *The transition matrix of this Markov chain is symmetric*

PROOF. We need to show that for any two permutations x, y of $[n]$ with i inversions, $p(x, y) = p(y, x)$. Each path from x to y starts by going to a permutation z of either higher or lower inversions, where the probability of going from x to z is $\frac{1}{n-1}$. Once at z , the path goes to x . Note that there is only one path from z to x , and that the degree of node z is $n-1$. However, it is possible that several edges emanating from node z do not go to nodes with i inversions. Assume, without loss of generality, that the path from x to z increases the inversions. Going from z to y must then decrease the inversions. Let d be the number of nodes connected to z that decrease the inversions, and note that one of these nodes is y . Thus, the probability of going from z to y is $\frac{1}{d}$. Now consider the reverse path, from y to x . Once again, at node z the probability of going from z to x is $\frac{1}{d}$, and since the

probability of going from x to z is the same as the probability of going from y to z , we have the desired conclusion. $\square$

In general, this graph is not regular. For example, consider the case of $n = 5$, $i = 3$. The permutation (15234) has degree 4: $\{(12543), (13524), (15234), (21534)\}$, but (31425) has degree 5: $\{(13452), (14325), (32145), (31254), (31425)\}$.

Discussion. Next, we look at how long is necessary to run the algorithm before achieving approximate uniformity. Although a rigorous analysis of mixing times has not been completed, some preliminary remarks are in order.

In the above examples, the SLEM (second largest eigenvalue in modulus) of these algorithms, which govern their convergence rate [LPW, Sec. 12.2], are approximately 0.694 for Algorithm 1A, 0.805 for Algorithm 1B, and 0.833 for Algorithm 2. This agrees with intuition that Algorithm 1A should converge the fastest, as in each step it explores a potentially larger region of the state space than the other algorithms. The other two algorithms were intended to be simpler to analyze, but that has not been the case.

One can also compare the number of steps needed by these algorithms to the Gnedin algorithm (Section 4.4.1.2) to see if these more sophisticated algorithms are even needed, depending on the parameters of the problem. For example, the number of trials until success in the Gnedin algorithm will be geometrically distributed with parameter $\frac{\phi(n,i)p^i}{[n]_p}$, and this may be reasonable for some combinations of n, p and i .

4.5. Sequential Construction

In this section, we present concentration of measure results for the inversions and the level under the Mallows measure. We do not use any properties of the Mallows measure, and so these bounds are true for general distributions on permutations.

THEOREM 57. $\mathbb{P} \left[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \lambda n^{\frac{3}{2}} \right] \leq 2 \exp \left(-\frac{3\lambda^2}{4} \right).$

PROOF. Let us define

$$\mathbf{X}_n = \mathbf{I}_n - \mathbb{E}[\mathbf{I}_n],$$

then the sequence of random variables $\mathbf{X}_n$ forms a martingale with respect to a σ -field $\mathcal{F}_n$, and $\mathbf{X}_0 = 0$. We calculate

$$\begin{aligned}\mathbf{X}_n - \mathbf{X}_{n-1} &= \mathbf{I}_n - \mathbb{E}[\mathbf{I}_n] - \mathbf{I}_{n-1} + \mathbb{E}[\mathbf{I}_{n-1}] \\ &= (\mathbf{I}_n - \mathbf{I}_{n-1}) - (\mathbb{E}[\mathbf{I}_n] - \mathbb{E}[\mathbf{I}_{n-1}])\end{aligned}$$

and we already know that

$$\mathbf{I}_n - \mathbf{I}_{n-1} \leq n - 1$$

so

$$\mathbb{E}[\mathbf{I}_n] - \mathbb{E}[\mathbf{I}_{n-1}] \leq n - 1$$

Thus we have

$$\mathbf{X}_n - \mathbf{X}_{n-1} \leq 2(n - 1)$$

and so, by the Azuma Hoeffding inequality, we have

$$\mathbb{P}[|\mathbf{X}_n| \geq \lambda] \leq 2e^{-\lambda^2 / (2 \sum_{k=1}^n (2(k-1))^2)}$$

i.e.

$$\begin{aligned}
\mathbb{P}[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \lambda] &\leq 2 \exp\left(-\frac{\lambda^2}{2 \times \frac{2n}{3} (1 - 3n + 2n^2)}\right) \\
&= 2 \exp\left(-\frac{3\lambda^2}{4n(1 - 3n + 2n^2)}\right) \\
&\sim 2 \exp\left(-\frac{3\lambda^2}{4n^3}\right) \\
\mathbb{P}[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \beta n^\alpha] &\leq 2 \exp\left(-\frac{3(\beta n^\alpha)^2}{4n^3}\right) \\
&= 2 \exp\left(-\frac{3\beta^2 n^{2\alpha}}{4n^3}\right) \\
&= 2 \exp\left(-\frac{3\beta^2 n^{2\alpha-3}}{4}\right) \\
\mathbb{P}\left[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \beta n^{\frac{3}{2}}\right] &\leq 2 \exp\left(-\frac{3\beta^2}{4}\right) \\
\mathbb{P}\left[|\mathbf{I}_n - \mathbb{E}[\mathbf{I}_n]| \geq \lambda n^{\frac{3}{2}}\right] &\leq 2 \exp\left(-\frac{3\lambda^2}{4}\right)
\end{aligned}$$

□

Similarly, we have the following:

THEOREM 58. $\mathbb{P}[|\mathbf{L}_n - \mathbb{E}[\mathbf{L}_n]| \geq \lambda\sqrt{n}] \leq 2e^{-\frac{\lambda^2}{8}}.$

PROOF. Let

$$\mathbf{X}_n = \mathbf{L}_n - \mathbb{E}[\mathbf{L}_n],$$

then the sequence of random variables $\mathbf{X}_n$ forms a martingale with respect to a σ -field $\mathcal{F}_n$, and $\mathbf{X}_0 = 0$. We calculate

$$\begin{aligned}
\mathbf{X}_n - \mathbf{X}_{n-1} &= \mathbf{L}_n - \mathbb{E}[\mathbf{L}_n] - \mathbf{L}_{n-1} + \mathbb{E}[\mathbf{L}_{n-1}] \\
&= (\mathbf{L}_n - \mathbf{L}_{n-1}) - (\mathbb{E}[\mathbf{L}_n] - \mathbb{E}[\mathbf{L}_{n-1}])
\end{aligned}$$

and we already know that

$$\mathbf{L}_n - \mathbf{L}_{n-1} \leq 1$$

so

$$\mathbb{E}[\mathbf{L}_n] - \mathbb{E}[\mathbf{L}_{n-1}] \leq 1$$

Thus we have

$$\mathbf{X}_n - \mathbf{X}_{n-1} \leq 2$$

and by the Azuma Hoeffding inequality we have

$$\mathbb{P}[|\mathbf{X}_n| \geq \lambda] \leq 2e^{-\lambda^2/(2\sum_{k=1}^n 4)}$$

i.e.

$$\begin{aligned} \mathbb{P}[|\mathbf{L}_n - \mathbb{E}[\mathbf{L}_n]| \geq \lambda] &\leq 2e^{-\lambda^2/(8n)} \\ \mathbb{P}[|\mathbf{L}_n - \mathbb{E}[\mathbf{L}_n]| \geq \beta\sqrt{n}] &\leq 2\exp\left(-\frac{(\beta\sqrt{n})^2}{8n}\right) \\ &= 2\exp\left(-\frac{\beta^2}{8}\right) \\ \mathbb{P}[|\mathbf{L}_n - \mathbb{E}[\mathbf{L}_n]| \geq \lambda\sqrt{n}] &\leq 2e^{-\lambda^2/8} \end{aligned}$$

□

CHAPTER 5

Applications

5.1. Sizing of a Reordering Buffer in a Queueing System

In this section, we display some of the complexity of developing a queueing theoretic model of reordering based on an $M/M/\infty$ queue, for comparison to the random permutation approach of previous chapters. We are concerned only with the buffer that must hold the jobs for reordering, not the actual mechanism that scrambles the jobs. Thus, we choose the simplest possible models, where the jobs arrive in order, separated by some random amount of time, and reordering occurs due to variations in the service time of each job, which are assumed to be independent random variables.

The key observation that sparked the random permutation approach was the following.

THEOREM 59. *The size of the reordering buffer needed for a given permutation is the maximum element of the inversion table of the permutation.*

PROOF. $iv(\pi)_k$ is the number of elements to the left of position k that are greater than k . Thus, a reordering buffer would have to hold those elements until element k arrived in order to restore the correct order. The total buffer size required would be the maximum over all elements of the inversion table. $\square$

5.1.1. Continuous Time Queueing Model. Consider jobs arriving in continuous time to an $M/M/\infty$ queueing system, where the arrivals are Poisson with rate λ and the server has rate μ . We begin by determining the probability that job 1 completes service after job k , for $1 < k$.

Let $\mathbf{S}_i$ be i.i.d. exponential random variables with parameter μ , representing the service times, and let $\mathbf{A}_i$ be i.i.d. exponential random variables with parameter λ , representing the inter-arrival times. We say there is a k -missequencing if

$$\mathbf{S}_1 > \mathbf{S}_{k+1} + \sum_{i=1}^k \mathbf{A}_i$$

i.e. if the $k + 1^{st}$ job leaves before the first job. By stationarity, this is the same for any two jobs separated by k arrivals.

Let us consider first the simple case of $k = 1$. The probability density function f (PDF) and cumulative distribution function F (CDF) of $\mathbf{S}$ and $\mathbf{A}$ are respectively

$$\begin{aligned} f_{\mathbf{S}}(x) &= \mu e^{-\mu x} \mathbb{I}_{x>0} \\ F_{\mathbf{S}}(x) &= (1 - e^{-\mu x}) \mathbb{I}_{x>0} \end{aligned}$$

and

$$\begin{aligned} f_{\mathbf{A}}(x) &= \lambda e^{-\lambda x} \mathbb{I}_{x>0} \\ F_{\mathbf{A}}(x) &= (1 - e^{-\lambda x}) \mathbb{I}_{x>0} \end{aligned}$$

We repeatedly use the following well known facts about independent random variables.

FACT 60. [R, Ch 2, Ex. 43] *Let $\mathbf{X}$ and $\mathbf{Y}$ be independent random variables with densities f_x and f_y respectively. Then the density of $\mathbf{X} + \mathbf{Y}$ is given by the convolution formula*

$$f_{\mathbf{X}+\mathbf{Y}}(z) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(z-y) f_{\mathbf{Y}}(y) dy$$

and let $\mathbf{X}$ have distribution function F_x . Then

$$\mathbb{P}[\mathbf{X} \leq \mathbf{Y}] = \int_{-\infty}^{\infty} F_{\mathbf{X}}(y) f_{\mathbf{Y}}(y) dy$$

Thus

$$\mathbb{P}[\mathbf{X} > \mathbf{Y}] = 1 - \int_{-\infty}^{\infty} F_{\mathbf{X}}(y) f_{\mathbf{Y}}(y) dy$$

Let us now compute the density of $\mathbf{S} + \mathbf{A}$, i.e. the density of the departure time of the second job.

$$\begin{aligned} f_{\mathbf{S}+\mathbf{A}}(z) &= \int_{-\infty}^{\infty} f_{\mathbf{S}}(x) f_{\mathbf{A}}(z-x) dx \\ &= \int_{-\infty}^{\infty} \mu e^{-\mu x} \mathbb{I}_{x>0} \lambda e^{-\lambda(z-x)} \mathbb{I}_{z-x>0} dx \\ &= \int_0^z \mu e^{-\mu x} \lambda e^{-\lambda(z-x)} dx \\ &= \frac{\lambda \mu}{\mu - \lambda} (e^{-\lambda z} - e^{-\mu z}) \mathbb{I}_{z>0} \end{aligned}$$

Now, what is the probability that job 1 completes service after job 2? Denote by $\tilde{\mathbf{S}}$ a random variable with the same distribution as, but independent of, $\mathbf{S}$.

$$\begin{aligned} P[\tilde{\mathbf{S}} \leq \mathbf{S} + \mathbf{A}] &= \int_{-\infty}^{\infty} F_{\tilde{\mathbf{S}}}(y) f_{\mathbf{S}+\mathbf{A}}(y) dy \\ &= \int_{-\infty}^{\infty} (1 - e^{-\mu y}) \frac{\lambda \mu}{\mu - \lambda} (e^{-\lambda y} - e^{-\mu y}) \mathbb{I}_{y>0} dy \\ &= \frac{\lambda \mu}{\mu - \lambda} \int_0^{\infty} (1 - e^{-\mu y}) (e^{-\lambda y} - e^{-\mu y}) dy \\ &= 1 - \frac{\lambda}{2(\lambda + \mu)} \end{aligned}$$

so

$$\mathbb{P}[\tilde{\mathbf{S}} > \mathbf{S} + \mathbf{A}] = \frac{\lambda}{2(\lambda + \mu)}.$$

Thus we have shown the following.

THEOREM 61. *The probability of a 1-missequencing is $\frac{\lambda}{2(\lambda+\mu)}$.*

We see that the probability of a 1-missequencing is always less than or equal to $\frac{1}{2}$, which happens when the service rate μ is vanishingly small compared to the arrival rate λ .

With that simple calculation complete, let us now consider the case of $k > 1$. Note that the departure time of the k^{th} packet is distributed as $\mathbf{S} + \sum_{i=1}^k \mathbf{A}_i$, i.e. one service time plus k independent exponentially distributed inter-arrival times. So, $\mathbf{Z}_k = \sum_{i=1}^k \mathbf{A}_i$ is distributed as a gamma random variable with parameters k and $\frac{1}{\lambda}$:

$$f_{\mathbf{Z}_k}(z) = \frac{1}{\Gamma(k)} e^{-z\lambda} z^{k-1} \lambda^k$$

Thus the PDF of the departure time of the k^{th} job is

$$\begin{aligned} f_{\mathbf{S}+\mathbf{Z}_k}(z) &= \int_{-\infty}^{\infty} f_{\mathbf{S}}(x) f_{\mathbf{Z}_k}(z-x) dx \\ &= \int_{-\infty}^{\infty} \mu e^{-\mu x} \mathbb{I}_{x>0} \frac{1}{\Gamma(k)} e^{-(z-x)\lambda} (z-x)^{k-1} \lambda^k \mathbb{I}_{z-x>0} dx \\ &= \int_0^z \mu e^{-\mu x} \frac{1}{\Gamma(k)} e^{-(z-x)\lambda} (z-x)^{k-1} \lambda^k dx \\ &= \frac{e^{-\mu z} \mu (\Gamma(k) - \Gamma(k, z(\lambda - \mu)))}{\Gamma(k)} \frac{\lambda^k}{(\lambda - \mu)^k} \end{aligned}$$

where, as usual,

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

is the Euler gamma function, and

$$\Gamma(a, z) = \int_z^{\infty} t^{a-1} e^{-t} dt$$

is the incomplete gamma function.

Then

$$\begin{aligned}
\mathbb{P}[\tilde{\mathbf{S}} \leq \mathbf{Z}_k + \mathbf{S}] &= \int_{-\infty}^{\infty} F_{\tilde{\mathbf{S}}}(y) f_{\mathbf{Z}_k + \mathbf{S}}(y) dy \\
&= \int_{-\infty}^{\infty} (1 - e^{-\mu y}) \frac{\mu e^{-y\mu}}{\Gamma(k)} \left(\frac{\lambda}{\lambda - \mu} \right)^k (\Gamma(k) - \Gamma(k, y(\lambda - \mu))) \mathbb{I}_{y>0} dy \\
&= \frac{\mu}{\Gamma(k)} \left(\frac{\lambda}{\lambda - \mu} \right)^k \int_0^{\infty} (e^{-y\mu} - e^{-2y\mu}) (\Gamma(k) - \Gamma(k, y(\lambda - \mu))) dy \\
&= \frac{\mu}{\Gamma(k)} \left(\frac{\lambda}{\lambda - \mu} \right)^k \frac{\Gamma(k)}{2\mu} \left(2 \left(\frac{\lambda}{\lambda - \mu} \right)^{-k} - \left(\frac{\lambda + \mu}{\lambda - \mu} \right)^{-k} \right) \\
&= \frac{1}{2} \left(\frac{\lambda}{\lambda - \mu} \right)^k \left(2 \left(\frac{\lambda - \mu}{\lambda} \right)^k - \left(\frac{\lambda - \mu}{\lambda + \mu} \right)^k \right) \\
&\quad 1 - \frac{1}{2} \left(\frac{\lambda}{\lambda + \mu} \right)^k
\end{aligned}$$

so

$$\mathbb{P}[\tilde{\mathbf{S}} > \mathbf{Z}_k + \mathbf{S}] = \frac{1}{2} \left(\frac{\lambda}{\lambda + \mu} \right)^k$$

and thus we have shown the following:

THEOREM 62. *The probability of a k -missequencing is $\frac{1}{2} \left(\frac{\lambda}{\lambda + \mu} \right)^k$.*

Note that these results may also be derived by exploiting the exponential distributions lack of memory property.

5.2. Statistics

It would be useful to estimate the value of the parameter p of the Mallows distribution from observations of the inversions and/or level. We take steps towards that goal in this section.

5.2.0.1. Estimation of p , given i and n , single permutation. Note that under the Mallows model, i is a sufficient statistic for p . Consider first values of $p < 1$. Then we have the likelihood

$$\begin{aligned}
\mathcal{L}(p|i) &= \frac{p^i}{[n]_p!} \\
&= \frac{p^i (1-p)^n}{\prod_{k=1}^n (1-p^k)}
\end{aligned}$$

and so the log-likelihood is

$$\log \mathcal{L}(p|i) = i \log p + n \log (1-p) - \sum_{k=1}^n \log (1-p^k)$$

and the derivative of the log-likelihood is

$$\frac{d}{dp} \log \mathcal{L}(p|i) = \frac{i}{p} - \frac{n}{1-p} + \sum_{k=1}^n \frac{kp^{k-1}}{1-p^k}$$

Straightforward maximization in closed form is not feasible, due to the summation term. However, if we neglect this term, we find a starting point for numerical maximization of the log-likelihood of $p_0 = \frac{i}{i+n}$. For example with $n = 100$ and $i = 100$, we begin the maximization with $p_0 = 0.0291262$, and the optimal value is numerically found to be quite close, at 0.029429.

For large values of p (i.e. $p > 1$), $p_0 = \frac{i}{i+n}$ still works well as a starting point for numerical maximization. However, in contrast to the $p < 1$ case, it no longer approximates the MLE well. This is because the series in the log-likelihood is negative, and very large in absolute value, and cannot be neglected.

5.2.0.2. Estimation of p , given i and n , and m permutations. Here we assume m permutations, each with length n , and observed values $i_1, i_2, \dots, i_m$. Then, denoting $i = i_1 + i_2 + \dots + i_m$, we have the likelihood

$$\begin{aligned}
\mathcal{L}(p|i) &= \frac{p^i (1-p)^{mn}}{\prod_{k=1}^n (1-p^k)^m} \\
\log \mathcal{L}(p|i) &= i \log p + mn \log (1-p) - m \sum_{k=1}^n \log (1-p^k) \\
\frac{d}{dp} \log \mathcal{L}(p|i) &= \frac{i}{p} - \frac{mn}{1-p} + m \sum_{k=1}^n \frac{kp^{k-1}}{1-p^k}
\end{aligned}$$

and so similarly, we determine the MLE numerically, starting from $p_0 = \frac{i}{i+mn}$. Our recommendation in general remains the same as in the previous case.

5.2.0.3. *Estimation of p , given l and n , for a single permutation.* Here we have one permutation of length n with observed level l . Then, we have likelihood

$$\mathcal{L}(p|l) = \frac{[l]_p!}{[n]_p!} \left([1 + l]_p^{n-l} - [l]_p^{n-l} \right) \mathbb{I}_{l \in [n-1]} + \frac{1}{[n]_p!} \mathbb{I}_{l=0}$$

First, assume $l = 0$. Then

$$\begin{aligned} \mathcal{L}(p|0) &= \frac{1}{[n]_p!} \\ &= \frac{(1-p)^n}{\prod_{k=1}^n (1-p^k)} \\ \log \mathcal{L}(p|0) &= n \log(1-p) - \sum_{k=1}^n \log(1-p^k) \\ \frac{d}{dp} \log \mathcal{L}(p|0) &= \frac{-n}{1-p} + \sum_{k=1}^n \frac{kp^{k-1}}{1-p^k} \end{aligned}$$

and numerically maximizing the likelihood yields an MLE of 0.

Next, assume $l \neq 0$, then the likelihood function is

$$\begin{aligned} \mathcal{L}(p|l) &= \frac{[l]_p!}{[n]_p!} \left([1 + l]_p^{n-l} - [l]_p^{n-l} \right) \\ &= \frac{\prod_{k=1}^l (1-p^k)}{(1-p)^l} \frac{(1-p)^n}{\prod_{k=1}^n (1-p^k)} \left(\left(\frac{1-p^{l+1}}{1-p} \right)^{n-l} - \left(\frac{1-p^l}{1-p} \right)^{n-l} \right) \\ &= \frac{(1-p^{l+1})^{n-l} - (1-p^l)^{n-l}}{\prod_{k=l+1}^n (1-p^k)} \end{aligned}$$

and the log-likelihood is

$$\log \mathcal{L}(p|l) = \log \left((1-p^{l+1})^{n-l} - (1-p^l)^{n-l} \right) - \sum_{k=l+1}^n \log(1-p^k)$$

Here, even neglecting the series term in order to determine a starting point for numerical maximization does not yield enough simplification.

In addition, using the log-likelihood here involves logarithms, which for $p > 1$ have negative arguments. Thus using the likelihood in this case is preferred to the log-likelihood. Hence, the suggestion is to start the numerical maximization at, say, $1/2$ if the level is below $n - \sqrt{n\pi/2}$, the mean value of the level in the case of $p = 1$, and to start the maximization at, say, 2 otherwise.

CHAPTER 6

Conclusion

This research can be seen as the first steps in extending the wealth of knowledge about uniform random permutations to the Mallows measure, and promoting the use of the level as an informative characteristic of permutations in general.

Future research could include overcoming the challenges discussed below.

Analyzing the mixing times of the algorithms in Section 4.4 in order to be able to make operational suggestions as to which algorithm is optimal under varying scenarios would be of great practical use. The main difficulty in doing so is that the graphs obtained from the various algorithms are not regular, and are in general difficult to describe precisely enough to be able to calculate, for example, eigenvalues of the transition matrices on these graphs.

Theorem 29 suggests an interesting approach to understanding the limiting behavior of the inversion tables, and therefore the level. If a similar result applies to the Mallows case, then perhaps results about the maximum of a Brownian motion can be translated into results about the level. The difficulty at this point is in verifying the Lindeberg condition in the Mallows case, although numerical experiments suggest that it may hold. In Figure 6.0.1, centered and scaled inversion tables for the cases of $p = 1/4$, 1, and 4 are shown, and it is difficult to tell which is which.

Finally, using the results of Section 2.2, we can view the sequential construction of permutations in two additional ways.

In the first case, which we call the inversion table process, at each time t the configuration is the inversion table of the permutation after symbol t has been added. Thus, at time 0 we have the configuration $\emptyset$, and at time t a configuration is a list of length t of integers such that the k^{th} element is less than k . At time

FIGURE 6.0.1. Three Brownian motions?

t , a new particle is created and placed at any integer location k between 1 and t , inclusive. All the elements to the right of k are shifted and increased via the rule

$$(v_1, v_2, \dots, v_{k-1}, v_k, v_{k+1}, \dots, v_n) \mapsto (v_1, v_2, \dots, v_{k-1}, 0, v_k + 1, v_{k+1} + 1, \dots, v_n + 1)$$

Of course, the level can be read off by simply choosing the maximal value in the configuration.

The second representation is similar, but more useful. In it, we track the position of each particle. At time 0 we have the configuration $\emptyset$. At time 1, we have the particle labeled 1 at position 1. At time 2, we have either particle 1 at position 1 and particle 2 at position 2, or particle 1 at position 2 and particle 2 at position 1. At each time step t , we add a new particle with label t in any of the t positions. We then connect the world lines of each particle. Thus we end up with Figure 6.0.2 for the permutation $(3, 5, 4, 1, 2)$.

In this representation, we can view the process as that of non-intersecting paths in a wedge of a grid, where each point in the grid has a particle. Perhaps some of the methods described in [KRB] may be applicable.

FIGURE 6.0.2. A particle system representation of (35412)

Bibliography

- [Ai] Martin Aigner (2007) “A Course in Enumeration,” Springer-Verlag, Berlin
- [AD] David Aldous and Persi Diaconis (1999) “Longest Increasing Subsequences: From Patience Sorting to the Baik-Deift-Johansson Theorem,” *Bull. Amer. Math. Soc.* 36, 413-432
- [Ar] Jorg Arndt (2010) “Generating Random Permutations,” PhD Thesis available at <http://maths.anu.edu.au/~brent/pd/Arndt-thesis.pdf>
- [BGP] François Baccelli, Erol Gelenbe and Brigitte Plateau (1984) “An End-to-End Approach to the Resequencing Problem,” *Journal of the ACM* Volume 31 Issue 3, 474-485
- [Ba] Yuliy M. Baryshnikov (2001) “GUEs and Queues,” *Probab. Theory Related Fields*, 119(2):256-274
- [Bi] Patrick Billingsley (1999) “Convergence of Probability Measures,” 2nd Edition, Wiley Series in Probability and Statistics
- [Bo] Miklos Bona (2004). “Combinatorics of Permutations,” Chapman and Hall/CRC
- [D] Persi Diaconis (1988) “Group Representations in Probability and Statistics,” IMS Lecture Notes - Monograph Series, 11 Institute of Mathematical Statistics, Hayward Ca.
- [F] William Fulton (1996) “Young Tableaux: With Applications to Representation Theory and Geometry.” Cambridge University Press
- [GO] Alexander Gnedin and Grigori Olshanski (2010) “Q-Exchangeability via Quasi-Invariance,” *Ann. Probab.* Volume 38, Number 6, 2103-2135
- [K] Donald E. Knuth (1998) “Art of Computer Programming, Volume 3: Sorting and Searching,” 2nd Edition, Addison-Wesley Professional
- [KRB] Pavel L. Krapivsky, Sidney Redner and Eli Ben-Naim (2010) “A Kinetic View of Statistical Physics,” Cambridge University Press
- [L] Kenneth Lange (2010) “Applied Probability,” 2nd Edition, Springer
- [LP] Guy Louchard and Helmut Prodinger (2003) “The Number of Inversions in Permutations: A Saddle Point Approach,” *Journal of Integer Sequences*, Vol. 6, Article 03.2.8
- [LPW] David A. Levin, Yuval Peres and Elizabeth L. Wilmer (2008) “Markov Chains and Mixing Times,” American Mathematical Society
- [LR] http://www.lightreading.com/document.asp?doc_id=4009&page_number=8

- [Mah] Hosam M. Mahmoud (2000) "Sorting: A Distribution Theory," Wiley-Interscience
- [Mal] C. L. Mallows (1957) "Non-Null Ranking Models I," *Biometrika* 44, 114-130
- [Mar] Barbara H. Margolius (2001) "Permutations with Inversions," *Journal of Integer Sequences*, Vol. 4, Article 01.2.4
- [MSE1] <http://math.stackexchange.com/questions/29534>
- [MSE2] <http://math.stackexchange.com/questions/117984>
- [O] Neil O'Connell (2003) "A Path-Transformation for Random Walks and the Robinson-Schensted Correspondence," *Trans. Amer. Math. Soc.* 355, 3669-3697
- [P] J. Pitman (2006) "Combinatorial Stochastic Processes," Springer-Verlag, Berlin
- [PW] James Gary Propp and David Bruce Wilson (1996) "Exact Sampling with Coupled Markov Chains and Applications to Statistical Mechanics," *Random Structures and Algorithms*, volume 9 number 1&2, pp. 223-252
- [R] Sheldon M. Ross (1985) "Introduction to Probability Models," 3rd Edition, Academic Press
- [SF] Robert Sedgewick and Philippe Flajolet (1995) "An Introduction to the Analysis of Algorithms", Addison-Wesley Professional
- [S] Shannon Starr (2009) "Thermodynamic Limit for the Mallows Model on S_n ," *J. Math. Phys.* 50 095208, DOI:10.1063/1.3156746
- [SM] Shannon Starr and Carl Mueller (2011) "The Length of the Longest Increasing Subsequence of a Random Mallows Permutation," *J. Theoret. Probab.* accepted and available at Online First. <http://arxiv.org/abs/1102.3402>
- [W] http://en.wikipedia.org/wiki/Pentagonal_numbers
- [XT] Ye Xia and David Tse (2008) "On the Large Deviations of Resequencing Queue Size: 2-M/M/1 Case," *IEEE Transactions on Information Theory*, vol. 54, no. 9, Page 4107-4118

Index

- Azuma Hoeffding inequality, 33
- backward ranks, 42
- bubblesort, 8
- Euler Pentagonal numbers, 10
- $f(n) \sim g(n)$, 25
- Gnedin Algorithm, 68
- harmonic number, 37
- inversion table, 5
- inversions, 7
- level, 11
- Mahonian, 41
- major index, 41
- Mallows measure, 38
- permutation, 5
- $\phi(n, i)$, 9
- Propp Wilson Algorithm, 68
- $\psi(n, i, l)$, 13
- q-factorial $[n]_q!$, 8
- q-number $[n]_q$, 8
- Rayleigh distribution, 27
- shifted truncated geometric distribution, 43
- staircase diagram, 9
- swap, 8
- transposition, 8
- truncated geometric distribution, 42
- uniform measure, 20