A FIRST COURSE IN PROBABILITY AND MARKOV CHAINS
A First Course in Probability and Markov Chains
A First Course in Probability and Markov Chains

Giuseppe Modica

and

Laura Poggiolini

University of Firenze, Italy
# Contents

## Preface

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>xi</td>
</tr>
</tbody>
</table>

## 1 Combinatorics

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Binomial coefficients</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1</td>
<td>Pascal triangle</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2</td>
<td>Some properties of binomial coefficients</td>
<td>2</td>
</tr>
<tr>
<td>1.1.3</td>
<td>Generalized binomial coefficients and binomial series</td>
<td>3</td>
</tr>
<tr>
<td>1.1.4</td>
<td>Inversion formulas</td>
<td>4</td>
</tr>
<tr>
<td>1.1.5</td>
<td>Exercises</td>
<td>6</td>
</tr>
<tr>
<td>1.2</td>
<td>Sets, permutations and functions</td>
<td>8</td>
</tr>
<tr>
<td>1.2.1</td>
<td>Sets</td>
<td>8</td>
</tr>
<tr>
<td>1.2.2</td>
<td>Permutations</td>
<td>8</td>
</tr>
<tr>
<td>1.2.3</td>
<td>Multisets</td>
<td>10</td>
</tr>
<tr>
<td>1.2.4</td>
<td>Lists and functions</td>
<td>11</td>
</tr>
<tr>
<td>1.2.5</td>
<td>Injective functions</td>
<td>12</td>
</tr>
<tr>
<td>1.2.6</td>
<td>Monotone increasing functions</td>
<td>12</td>
</tr>
<tr>
<td>1.2.7</td>
<td>Monotone nondecreasing functions</td>
<td>13</td>
</tr>
<tr>
<td>1.2.8</td>
<td>Surjective functions</td>
<td>14</td>
</tr>
<tr>
<td>1.2.9</td>
<td>Exercises</td>
<td>16</td>
</tr>
<tr>
<td>1.3</td>
<td>Drawings</td>
<td>16</td>
</tr>
<tr>
<td>1.3.1</td>
<td>Ordered drawings</td>
<td>16</td>
</tr>
<tr>
<td>1.3.2</td>
<td>Simple drawings</td>
<td>17</td>
</tr>
<tr>
<td>1.3.3</td>
<td>Multiplicative property of drawings</td>
<td>17</td>
</tr>
<tr>
<td>1.3.4</td>
<td>Exercises</td>
<td>18</td>
</tr>
<tr>
<td>1.4</td>
<td>Grouping</td>
<td>19</td>
</tr>
<tr>
<td>1.4.1</td>
<td>Collocations of pairwise different objects</td>
<td>19</td>
</tr>
<tr>
<td>1.4.2</td>
<td>Collocations of identical objects</td>
<td>22</td>
</tr>
<tr>
<td>1.4.3</td>
<td>Multiplicative property</td>
<td>23</td>
</tr>
<tr>
<td>1.4.4</td>
<td>Collocations in statistical physics</td>
<td>24</td>
</tr>
<tr>
<td>1.4.5</td>
<td>Exercises</td>
<td>24</td>
</tr>
</tbody>
</table>

## 2 Probability measures

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Elementary probability</td>
<td>28</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Exercises</td>
<td>29</td>
</tr>
</tbody>
</table>
2.2 Basic facts  
2.2.1 Events  34  
2.2.2 Probability measures  36  
2.2.3 Continuity of measures  37  
2.2.4 Integral with respect to a measure  39  
2.2.5 Probabilities on finite and denumerable sets  40  
2.2.6 Probabilities on denumerable sets  42  
2.2.7 Probabilities on uncountable sets  44  
2.2.8 Exercises  46  
2.3 Conditional probability  51  
2.3.1 Definition  51  
2.3.2 Bayes formula  52  
2.3.3 Exercises  54  
2.4 Inclusion–exclusion principle  60  
2.4.1 Exercises  63  

3 Random variables  68  
3.1 Random variables  68  
3.1.1 Definitions  69  
3.1.2 Expected value  75  
3.1.3 Functions of random variables  77  
3.1.4 Cavalieri formula  80  
3.1.5 Variance  82  
3.1.6 Markov and Chebyshev inequalities  82  
3.1.7 Variational characterization of the median and of the expected value  83  
3.1.8 Exercises  84  
3.2 A few discrete distributions  91  
3.2.1 Bernoulli distribution  91  
3.2.2 Binomial distribution  91  
3.2.3 Hypergeometric distribution  93  
3.2.4 Negative binomial distribution  94  
3.2.5 Poisson distribution  95  
3.2.6 Geometric distribution  98  
3.2.7 Exercises  101  
3.3 Some absolutely continuous distributions  102  
3.3.1 Uniform distribution  102  
3.3.2 Normal distribution  104  
3.3.3 Exponential distribution  106  
3.3.4 Gamma distributions  108  
3.3.5 Failure rate  110  
3.3.6 Exercises  111  

4 Vector valued random variables  113  
4.1 Joint distribution  113
4.1.1 Joint and marginal distributions 114
4.1.2 Exercises 117

4.2 Covariance 120
   4.2.1 Random variables with finite expected value and variance 120
   4.2.2 Correlation coefficient 123
   4.2.3 Exercises 123

4.3 Independent random variables 124
   4.3.1 Independent events 124
   4.3.2 Independent random variables 127
   4.3.3 Independence of many random variables 128
   4.3.4 Sum of independent random variables 130
   4.3.5 Exercises 131

4.4 Sequences of independent random variables 140
   4.4.1 Weak law of large numbers 140
   4.4.2 Borel–Cantelli lemma 142
   4.4.3 Convergences of random variables 143
   4.4.4 Strong law of large numbers 146
   4.4.5 A few applications of the law of large numbers 152
   4.4.6 Central limit theorem 159
   4.4.7 Exercises 163

5 Discrete time Markov chains 168
   5.1 Stochastic matrices 168
      5.1.1 Definitions 169
      5.1.2 Oriented graphs 170
      5.1.3 Exercises 172
   5.2 Markov chains 173
      5.2.1 Stochastic processes 173
      5.2.2 Transition matrices 174
      5.2.3 Homogeneous processes 174
      5.2.4 Markov chains 174
      5.2.5 Canonical Markov chains 178
      5.2.6 Exercises 181
   5.3 Some characteristic parameters 187
      5.3.1 Steps for a first visit 187
      5.3.2 Probability of (at least) \( r \) visits 189
      5.3.3 Recurrent and transient states 191
      5.3.4 Mean first passage time 193
      5.3.5 Hitting time and hitting probabilities 195
      5.3.6 Exercises 198
   5.4 Finite stochastic matrices 201
      5.4.1 Canonical representation 201
      5.4.2 States classification 203
      5.4.3 Exercises 205
5.5 Regular stochastic matrices 206
  5.5.1 Iterated maps 206
  5.5.2 Existence of fixed points 209
  5.5.3 Regular stochastic matrices 210
  5.5.4 Characteristic parameters 218
  5.5.5 Exercises 220

5.6 Ergodic property 222
  5.6.1 Number of steps between consecutive visits 222
  5.6.2 Ergodic theorem 224
  5.6.3 Powers of irreducible stochastic matrices 226
  5.6.4 Markov chain Monte Carlo 228

5.7 Renewal theorem 233
  5.7.1 Periodicity 233
  5.7.2 Renewal theorem 234
  5.7.3 Exercises 239

6 An introduction to continuous time Markov chains 241
  6.1 Poisson process 241
  6.2 Continuous time Markov chains 246
    6.2.1 Definitions 246
    6.2.2 Continuous semigroups of stochastic matrices 248
    6.2.3 Examples of right-continuous Markov chains 256
    6.2.4 Holding times 259

Appendix A Power series 261
  A.1 Basic properties 261
  A.2 Product of series 263
  A.3 Banach space valued power series 264
    A.3.2 Exercises 267

Appendix B Measure and integration 270
  B.1 Measures 270
    B.1.1 Basic properties 270
    B.1.2 Construction of measures 272
    B.1.3 Exercises 279
  B.2 Measurable functions and integration 279
    B.2.1 Measurable functions 280
    B.2.2 The integral 283
    B.2.3 Properties of the integral 284
    B.2.4 Cavalieri formula 286
    B.2.5 Markov inequality 287
    B.2.6 Null sets and the integral 287
    B.2.7 Push forward of a measure 289
    B.2.8 Exercises 290
B.3 Product measures and iterated integrals 294
  B.3.1 Product measures 294
  B.3.2 Reduction formulas 296
  B.3.3 Exercises 297
B.4 Convergence theorems 298
  B.4.1 Almost everywhere convergence 298
  B.4.2 Strong convergence 300
  B.4.3 Fatou lemma 301
  B.4.4 Dominated convergence theorem 302
  B.4.5 Absolute continuity of integrals 305
  B.4.6 Differentiation of the integral 305
  B.4.7 Weak convergence of measures 308
  B.4.8 Exercises 312

Appendix C Systems of linear ordinary differential equations 313
  C.1 Cauchy problem 313
    C.1.1 Uniqueness 313
    C.1.2 Existence 315
  C.2 Efficient computation of $e^{Qt}$ 317
    C.2.1 Similarity methods 317
    C.2.2 Putzer method 319
  C.3 Continuous semigroups 321

References 324

Index 327
Preface

This book collects topics covered in introductory courses in probability delivered by the authors at the University of Florence. It aims to introduce the reader to typical structures of probability with a language appropriate for further advanced reading. The attention is mainly focused on basic structures.

There is a well established tradition of studies in probability due to the wide range of possible applications of related concepts and structures in science and technology. Therefore, an enormous amount of literature on the subject is available, including treatises, lecture notes, reports, journal papers and web pages. The list of references at the end of this book is obviously incomplete and includes only references used directly in writing the following pages. Throughout this book we adopt the language of measure theory (relevant notions are recalled in the appendices).

The first part of the book deals with basic notions of combinatorics and probability calculus: counting problems and uniform probability, probability measures, probability distributions, conditional probability, inclusion–exclusion principle, random variables, dispersion indexes, independence, and the law of large numbers are also discussed. Central limit theorem is presented without proof. Only a basic knowledge of linear algebra and mathematical analysis is required.

In the second part we discuss, as a first example of stochastic processes, Markov chains with discrete time and discrete states, including the Markov chain Monte Carlo method, and we introduce Poisson process and continuous time Markov chains with finite states. For this part, further notions in mathematical analysis (summarized in the appendices) are required: the Banach fixed point theorem, systems of linear ordinary differential equations, powers and power series of matrices.

We wish to thank all the students who have attended our courses. We also wish to thank our colleagues Matteo Focardi, Mariano Giaquinta, Paolo Maria Mariano, Andrea Mennucci, Marco Romito and Enrico Vicario who helped us with suggestions and comments. Special gratitude goes to Enrico Vicario for many helpful discussions on the applications.

Our special thanks also go to the editorial staff of Wiley for the excellent quality of their work.
We have tried to avoid misprints and errors. However, we would be very grateful to be notified of any errors or misprints and would be glad to receive any criticism or comments. Our e-mail addresses are:

giuseppe.modica@unifi.it    laura.poggiolini@unifi.it

We will try to keep up an errata corrige at the following web pages:

http://www.dma.unifi.it/~modica
http://www.dma.unifi.it/~poggiolini
http://www.dmi.unifi.it/~modica
http://www.dmi.unifi.it/~poggiolini
Combinatorics

Combinatorics deals with the cardinality of classes of objects. The first example that jumps to our minds is the computation of how many triplets can be drawn from 90 different balls. In this chapter and the next we are going to compute the cardinality of several classes of objects.

1.1 Binomial coefficients

1.1.1 Pascal triangle

Binomial coefficients are defined as

\[
\binom{n}{k} := \begin{cases} 
1 & \text{if } k = 0, \\
0 & \text{if } k > n, \\
\frac{n(n-1)\ldots(n-k+1)}{k!} & \text{if } n \geq 1 \text{ and } 1 \leq k \leq n.
\end{cases}
\]

Binomial coefficients are usually grouped in an infinite matrix

\[\mathbf{C} := (\mathbf{C}_k^n), \quad n, k \geq 0, \quad \mathbf{C}_k^n := \binom{n}{k}\]

called a Pascal triangle given the triangular arrangement of the nonzero entries, see Figure 1.1. Here and throughout the book we denote the entries of a matrix (finite or infinite) \(\mathbf{A} = (a_{ij})\) where the superscript \(i\) and the subscript \(j\) mean the \(i\)th row and the \(j\)th column, respectively. Notice that the entries of each row of \(\mathbf{C}\) are zero if the column index is large enough, \(\mathbf{C}_j^i = 0 \forall i, j \text{ with } j > i \geq 0\). We also recall the Newton binomial formula,

\[
(1 + z)^n = \binom{n}{0} + \binom{n}{1}z + \cdots + \binom{n}{n}z^n = \sum_{k=0}^{\infty} \binom{n}{k}z^k = \sum_{k=0}^{\infty} \mathbf{C}_k^n z^k.
\]
Thus formula can be proven with an induction argument on \( n \) or by means of Taylor formula.

1.1.2 Some properties of binomial coefficients

Many formulas are known on binomial coefficients. In the following proposition we collect some of the simplest and most useful ones.

**Proposition 1.1** The following hold.

- (i) \( \binom{n}{k} = \frac{n!}{k!(n-k)!} \) \( \forall k, n, 0 \leq k \leq n \).
- (ii) \( \binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1} \) \( \forall k, n, 1 \leq k \leq n \).
- (iii) **Stifel formula** \( \binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \) \( \forall k, n, 1 \leq k \leq n \).
- (iv) \( \binom{n}{j} \binom{j}{k} = \binom{n}{k} \binom{n-k}{j-k} \) \( \forall k, j, n, 0 \leq k \leq j \leq n \).
- (v) \( \binom{n}{k} = \binom{n}{n-k} \) \( \forall k, 0 \leq k \leq n \).
- (vi) the map \( k \mapsto \binom{n}{k} \) achieves its maximum at \( k = \left\lfloor \frac{n}{2} \right\rfloor \).
- (vii) \( \sum_{k=0}^{n} \binom{n}{k} = 2^n \forall n \geq 0 \).
- (viii) \( \sum_{k=0}^{n} (-1)^k \binom{n}{k} = \delta_{0,n} = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \neq 0. \end{cases} \)
- (ix) \( \binom{n}{k} \leq 2^n \forall k, n, 0 \leq k \leq n \).
Proof. Formulas (i), (ii), (iii), (iv) and (v) directly follow from the definition. (vi) is a direct consequence of (v). (vii) and (viii) follow from the Newton binomial formula \( \sum_{k=0}^{n} \binom{n}{k} z^k = (1 + z)^n \) choosing \( z = 1 \) and \( z = -1 \). Finally, (ix) is a direct consequence of (vii).

Estimate (ix) in Proposition 1.1 can be made more precise. For instance, from the Stirling asymptotical estimate of the factorial,

\[
\frac{n!}{n^n e^{-n} \sqrt{2\pi n}} \to 1 \quad \text{as} \quad n \to \infty
\]

one gets

\[
(2n)! = 4^n n^{2n} e^{-2n} \sqrt{4\pi n} (1 + o(1)),
\]

\[
(n!)^2 = n^{2n} e^{-2n} 2\pi n (1 + o(1)),
\]

so that

\[
\binom{2n}{n} = \frac{4^n}{\sqrt{\pi n}} \frac{1 + o(1)}{1 + o(1)}
\]

or, equivalently,

\[
\frac{\binom{2n}{n}}{4^n} \to 1 \quad \text{as} \quad n \to \infty.
\]

(1.1)

Estimate (1.1) is ‘accurate’ also for small values of \( n \). For instance, for \( n = 4 \), one has \( \binom{8}{4} = 70 \) and \( 4^4 \frac{1}{\sqrt{\pi}} \approx 72.2 \).

1.1.3 Generalized binomial coefficients and binomial series

For \( \alpha \in \mathbb{R} \) we define the sequence \( \binom{\alpha}{n} \) of generalized binomial coefficients as

\[
\binom{\alpha}{n} := \begin{cases} 
1 & \text{if } n = 0, \\
\frac{\alpha(\alpha-1)(\alpha-2) \cdots (\alpha-n+1)}{n!} & \text{if } n \geq 1.
\end{cases}
\]

Notice that \( \binom{\alpha}{k} \neq 0 \ \forall k \) if \( \alpha \notin \mathbb{N} \) and \( \binom{\alpha}{k} = 0 \ \forall k \geq \alpha + 1 \) if \( \alpha \in \mathbb{N} \). The power series

\[
\sum_{n=0}^{\infty} \binom{\alpha}{n} z^n
\]

is called the binomial series.

Proposition 1.2 (Binomial series) The binomial series converges if \( |z| < 1 \) and

\[
\sum_{n=0}^{\infty} \binom{\alpha}{n} z^n = (1 + z)^\alpha \quad \text{if} \quad |z| < 1.
\]
Proof. Since
\[
\left| \frac{\alpha_{n+1}}{\alpha_n} \right| = \frac{|\alpha - n|}{n+1} \to 1 \quad \text{as } n \to \infty,
\]
it is well known that \( \sqrt[n]{|a_n|} \to 1 \) as well; thus, the radius of the power series in (1.2) is 1.

Differentiating \( n \) times the map \( z \mapsto (1 + z)^\alpha \), one gets
\[
D^n((1 + z)^\alpha) = \alpha(\alpha - 1) \cdots (\alpha - n + 1)(1 + z)^{\alpha-n},
\]
so that the series on the left-hand side of (1.2) is the McLaurin expansion of \((1 + z)^\alpha\).

Another proof is the following. Let \( S(z) := \sum_{n=0}^{\infty} (\frac{\alpha}{n}) z^n \), \( |z| < 1 \), be the sum of the binomial series. Differentiating one gets
\[
(1 + z)S'(z) = \alpha S(z), \quad |z| < 1,
\]
hence
\[
\left( \frac{S(z)}{(1 + z)^\alpha} \right)' = \frac{(1 + z)S'(z) - \alpha S(z)}{(1 + z)^{\alpha+1}} = 0.
\]
Thus there exists \( c \in \mathbb{R} \) such that \( S(z) = c (1 + z)^\alpha \) if \( |z| < 1 \). Finally, \( c = 1 \) since \( S(0) = 1 \).

**Proposition 1.3** Let \( \alpha \in \mathbb{R} \). The following hold.

(i) \( \binom{\alpha}{k} = \frac{\alpha}{k} \binom{\alpha - 1}{k - 1} \) \( \forall k \geq 1 \).

(ii) \( \binom{\alpha}{k} = \binom{\alpha - 1}{k} + \binom{\alpha - 1}{k - 1} \) \( \forall k \geq 1 \).

(iii) \( \binom{-\alpha}{k} = (-1)^k \binom{\alpha + k - 1}{k} \) \( \forall k \geq 0 \).

Proof. The proofs of (i) and (ii) are left to the reader. Proving (iii) is a matter of computation:
\[
\binom{-\alpha}{k} = \frac{-\alpha(-\alpha - 1) \cdots (-\alpha - k + 1)}{k!} = (-1)^k \frac{\alpha(\alpha + 1) \cdots (\alpha + k - 1)}{k!}
\]
\[
= (-1)^k \binom{\alpha + k - 1}{k}.
\]
A few negative binomial coefficients are quoted in Figure 1.2.

### 1.1.4 Inversion formulas

For any \( N \), the matrix \( C_N := (C^n_k) \), \( n, k = 0, \ldots, N \), is lower triangular and all its diagonal entries are equal to 1. Hence 1 is the only eigenvalue of \( C_N \).
\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & \ldots \\
1 & -2 & 3 & -4 & 5 & -6 & 7 & -8 & 9 & -10 & \ldots \\
1 & -3 & 6 & -10 & 15 & -21 & 28 & -36 & 45 & -55 & \ldots \\
1 & -4 & 10 & -20 & 35 & -56 & 84 & -120 & 165 & -220 & \ldots \\
1 & -5 & 15 & -35 & 70 & -126 & 210 & -330 & 495 & -715 & \ldots \\
1 & -6 & 21 & -56 & 126 & -252 & 462 & -792 & 1287 & -2002 & \ldots \\
1 & -7 & 28 & -84 & 210 & -462 & 924 & -1716 & 3003 & -5005 & \ldots \\
1 & -8 & 36 & -120 & 330 & -720 & 1560 & -3432 & 6435 & -11440 & \ldots \\
1 & -9 & 45 & -165 & 495 & -1287 & 3003 & -6435 & 12870 & -24310 & \ldots \\
\end{pmatrix}
\]

Figure 1.2 The coefficients \(\binom{-n}{k}\).

with algebraic multiplicity \(N\). In particular \(C_N\) is invertible, its inverse is lower triangular, all its entries are integers and its diagonal entries are equal to 1.

**Theorem 1.4** For any \(n, k = 0, \ldots, N\), \((C_N^{-1})^n_k = (-1)^{n+k} \binom{n}{k}\).

**Proof.** Let \(B := (B^n_k), B^n_k := (-1)^{n+k} \binom{n}{k}\) so that both \(B\) and \(C_N B\) are lower triangular, i.e. \((C_N B)^n_k = 0\) if \(0 \leq n < k\). Moreover, (iv) and (viii) of Proposition 1.1 yield for any \(n \geq k\)

\[
(C_N B)_k^n = n \sum_{j=1}^{N} \left( \binom{n}{j}(-1)^{j+k} \binom{j}{k} \right) = \sum_{j=k}^{n} (-1)^{j+k} \binom{n}{j} \binom{j}{k} \\
= \left( \binom{n}{k} \right) \sum_{j=k}^{n} (-1)^{j+k} \binom{n-k}{j-k} = \left( \binom{n}{k} \right) \sum_{i=0}^{n-k} (-1)^{i} \binom{n-k}{i} \\
= \left( \binom{n}{k} \right) \delta_{0,n-k} = \delta_{n,k}.
\]

A few entries of the inverse of the matrix of binomial coefficients are shown in Figure 1.3. As a consequence of Theorem 1.4 the following inversion formulas hold.

**Corollary 1.5** Two sequences \(\{x_n\}, \{y_n\}\) satisfy

\[
y_n = \sum_{k=0}^{n} \binom{n}{k} x_k, \quad \forall n \geq 0
\]

if and only if

\[
x_n = \sum_{k=0}^{n} (-1)^{n+k} \binom{n}{k} y_k, \quad \forall n \geq 0.
\]
Similarly,

**Corollary 1.6** Two $N$-tuples or real numbers $\{x_n\}$ and $\{y_n\}$ satisfy

$$y_n = \sum_{k=n}^{N} \binom{k}{n} x_k, \quad \forall n, \ 0 \leq n \leq N,$$

if and only if

$$x_n = \sum_{k=n}^{N} (-1)^{n+k} \binom{k}{n} y_k, \quad \forall n, \ 0 \leq n \leq N.$$ 

### 1.1.5 Exercises

**Exercise 1.7** Prove Newton binomial formula:

(i) directly, with an induction argument on $n$;

(ii) applying Taylor expansion formula;

(iii) starting from the formula $D((1+z)^n) = n(1+z)^{n-1}$.

**Exercise 1.8** Differentiating the power series, see Appendix A, prove the formulas in Figure 1.4.

*Solution.* Differentiating the identity $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$, $|z| < 1$, we get for $|z|<1$

$$\sum_{k=0}^{\infty} k z^k = z \sum_{k=0}^{\infty} D(z^k) = z D\left( \sum_{k=0}^{\infty} z^k \right) = z D\left( \frac{1}{1-z} \right) = \frac{z}{(1-z)^2};$$
Let $z \in \mathbb{C}$, $|z| < 1$, and $n \in \mathbb{Z}$. We have the followings.

(i) \[ \sum_{k=0}^{\infty} k z^k = \frac{z}{(1 - z)^2}, \]

(ii) \[ \sum_{k=0}^{\infty} k^2 z^{k-1} = D\left(\frac{z}{(1 - z)^2}\right) = \frac{1 + z}{(1 - z)^3}, \]

(iii) \[ \sum_{k=0}^{\infty} \binom{n}{k} z^k = (1 + z)^n, \]

(iv) \[ \sum_{k=0}^{\infty} k \binom{n}{k} z^k = nz(1 + z)^{n-1}, \]

(v) \[ \sum_{k=0}^{\infty} k^2 \binom{n}{k} z^k = nz(1 + nz)(1 + z)^{n-2}. \]

Figure 1.4 The sum of a few series related to the geometric series.

\[
\sum_{k=0}^{\infty} k^2 z^{k-1} = \sum_{k=0}^{\infty} D(kz^k)
\]

\[= D\left(\sum_{k=0}^{\infty} k z^k\right) = D\left(\frac{z}{(1 - z)^2}\right) = \frac{1 + z}{(1 - z)^3}, \]

Moreover, for any non-negative integer $n$, differentiating the identities

\[ \sum_{k=0}^{\infty} \binom{n}{k} z^k = (1 + z)^n \quad \text{and} \quad \sum_{k=0}^{\infty} (-\binom{n}{k}) z^k = (1 + z)^{-n} \]

for any $|z| < 1$, we get

\[ \sum_{k=0}^{\infty} k \binom{n}{k} z^k = z \sum_{k=0}^{n} D\left(\binom{n}{k} z^k\right) = zD((1 + z)^n) = nz(1 + z)^{n-1}; \]

\[ \sum_{k=0}^{\infty} k^2 \binom{n}{k} z^k = z \sum_{k=0}^{n} D\left(k \binom{n}{k} z^k\right) = zD\left(\sum_{k=0}^{n} \binom{n}{k} z^k\right) = nz(1 + nz)(1 + z)^{n-2}. \]
1.2 Sets, permutations and functions

1.2.1 Sets

We recall that a finite set $A$ is an unordered collection of pairwise different objects. For example, the collection of objects are $1, 2, 3$ is a finite set which we denote as $A = \{1, 2, 3\}$; the collection $1, 2, 2, 3$ is not a finite set, and $\{1, 2, 3\}$ and $\{2, 1, 3\}$ are the same set.

If $A$ is a finite set with $n$ objects (or elements), we may enumerate the elements of $A$ so that $A = \{x_1, \ldots, x_n\}$. Therefore, for counting purposes, we can assume without loss of generality that $A = \{1, \ldots, n\}$. The number $n$ is the cardinality of $A$, and we write $|A| = n$.

**Proposition 1.9** Let $A$ be a finite set with $n$ elements, $n \geq 1$. There are $C^n_k = \binom{n}{k}$ subsets of $A$ with $k$ elements.

**Proof.** Different proofs can be done. We propose one of them. Let $d_{n,k}$ be the number of subsets of $A$ with $k$ elements. Obviously, $d_{n,1} = n$ and $d_{n,n} = 1$. For $2 \leq k \leq n - 1$, assume we have $n$ football players and we want to select a team of $k$ of them, including the captain of the team. We may proceed in the following way: first we choose the team of $k$-players: $d_{n,k}$ different teams can be selected. Then, among the team, we select the captain: $k$ different choices are possible: so there are $kd_{n,k}$ ways to select the team and the captain. However, we can proceed in another way: first we choose the captain among the $n$ players: there are $n$ different possible choices. Then we choose $k - 1$ players among the remaining $n - 1$ players: there are $d_{n-1,k-1}$ possible choices. Thus

$$d_{n,k} = \frac{n}{k}d_{n-1,k-1}$$

which by induction, gives

$$d_{n,k} = \frac{n}{k} \frac{n-1}{k-1} \cdots \frac{n-k+2}{2}d_{n-k+1,1}$$

$$= \frac{n}{k} \frac{n-1}{k-1} \cdots \frac{n-k+2}{2} \frac{n-k+1}{1} = \binom{n}{k}.$$

1.2.2 Permutations

Let $N$ be a finite set and let $n$ be its cardinality. Without loss of generality, we can assume $N = \{1, 2, \ldots, n\}$. A permutation of $N$ is an injective (and thus one-to-one) mapping $\pi : N \to N$. Since composing bijective maps yields another bijective map, the family of permutations of a set $N$ is a group with respect to the composition of maps; the unit element is the identity map; this group is called the group of permutations of $N$. It is denoted as $S_n$ or $P_n$. Notice that $P_n$ is a not a commutative group if $n \geq 3$.

Each permutation is characterized by its image-word or image-list, i.e. by the $n$-tuple $(\pi(1), \ldots, \pi(n))$. For instance, the permutation $\pi \in P_6$ defined by
\[ \pi(1) = 2, \pi(2) = 3, \pi(3) = 1, \pi(4) = 4, \pi(5) = 6 \text{ and } \pi(6) = 5 \] is denoted as

\[
\begin{pmatrix}
1 & 2 & 3 & 4 & 5 & 6 \\
2 & 3 & 1 & 4 & 6 & 5
\end{pmatrix}.
\]

or, in brief, with its image-word 231465.

The set of permutations of \( N = \{1, \ldots, n\} \) has \( n! \) elements,

\[ |P_n| = n! \]

In fact, the image \( \pi(1) \) of 1 can be chosen among \( n \) possible values, then the image \( \pi(2) \) of 2 can be chosen among \( n - 1 \) possible values and so on. Hence

\[ |P_n| = n(n - 1)(n - 2) \cdots 3 \cdot 2 \cdots 1 = n! \]

### 1.2.2.1 Derangements

Let \( \pi \in P_n \) be a permutation of \( N = \{1, \ldots, n\} \). A point \( x \in N \) is a fixed point of \( \pi \) if \( \pi(x) = x \).

We now compute the cardinality \( d_n \) of the set \( D_n \) of permutations without fixed points, also called derangements.

\[ D_n := \left\{ \pi \in P_n \mid \pi(i) \neq i \ \forall i \in N \right\}. \]

**Proposition 1.10** The cardinality of \( D_n \) is

\[ d_n = n! \sum_{j=0}^{n} (-1)^j \frac{1}{j!} \quad \forall n \geq 1. \]

**Proof.** If a permutation of \( N \) has \( j \) fixed points, \( 0 \leq j \leq n \), then it is a derangement of the other \( n - j \) points of \( N \). Thus, a permutation with \( j \) fixed points splits as a couple: the set of its fixed points and a derangement of \( n - j \) points. There are \( \binom{n}{j} \) different choices for the \( j \) fixed points and \( d_{n-j} \) derangements of the remaining \( n - j \) points, so that, the possible permutations of \( N \) with exactly \( j \) fixed points are \( \binom{n}{j} d_{n-j} \) (where \( d_0 = 1 \)). Thus

\[ |P_n| = \sum_{j=0}^{n} \binom{n}{j} d_{n-j} \forall n \geq 1, \text{ i.e.} \]

\[ n! = \sum_{j=0}^{n} \binom{n}{j} d_{n-j} \quad \forall n \geq 0. \quad (1.3) \]

The inversion formula of binomial coefficients, see Corollary 1.5, reads

\[ d_n = \sum_{j=0}^{n} (-1)^{n+j} \binom{n}{j} j! = n! \sum_{j=0}^{n} (-1)^j \frac{j!}{j!} \quad \forall n \geq 0. \]
0, 0, 1, 2, 9, 44, 265, 1854, 14,833, 133,496, 1334,961, 14,684,570, …

Figure 1.5 From the left, the numbers \(d_0, d_1, d_2, d_3, \ldots\) of derangements of 0, 1, 2, 3, … points.

**Corollary 1.11** The number \(d_n\) of derangements of \(n\) points is the nearest integer to \(n!/e\).

**Proof.** The elementary estimate between the exponential and its McLaurin expansion gives

\[
|e^x - \sum_{j=0}^{n} \frac{x^j}{j!}| \leq \frac{|x|^{n+1}}{(n+1)!}, \quad \forall x \leq 0;
\]

hence for \(x = -1\) we get

\[
\left| \frac{1}{e} - \sum_{j=0}^{n} \frac{(-1)^j}{j!} \right| \leq \frac{1}{(n+1)!},
\]

so that, from Proposition 1.10 one gets

\[
\left| d_n - \frac{n!}{e} \right| = n! \left| \sum_{j=0}^{n} \frac{(-1)^j}{j!} - \frac{1}{e} \right| \leq \frac{n!}{(n+1)!} = \frac{1}{n+1} \leq \frac{1}{3}
\]

for each \(n \geq 2\).

Figure 1.5 contains the first elements of the sequence \(\{d_n\}\).

### 1.2.3 Multisets

Another interesting structure is an unordered list of elements taken from a given set \(A\). This structure is called a **multiset** on \(A\). More formally, a multiset on a set \(A\) is a couple \((A, a)\) where \(A\) is a given set and \(a : A \rightarrow \mathbb{N} \cup \{+\infty\}\) is the *multiplicity function* which counts ‘how many times’ an element \(x \in A\) appears in the multiset. Clearly, each set is a multiset where each object has multiplicity 1. We denote as \(\{a^2, b^2, c^5\}\) or \(a^2b^2c^5\) the multiset on \(A := \{a, b, c\}\) where \(a\) and \(b\) have multiplicity 2 and \(c\) has multiplicity 5. The cardinality of a multiset \((A, a)\) is \(\sum_{x \in A} a(x)\) and is denoted by \(|(A, a)|\) or \(#(A, a)\). For instance, the cardinality of \(a^2b^2c^5\) is 9.

If \(B\) is a subset of \(A\), then \(B\) is also the multiset \((A, a)\) on \(A\) where

\[
a(x) = \begin{cases} 
1 & \text{if } x \in B, \\
0 & \text{if } x \notin B.
\end{cases}
\]
Given two multisets \((B, b)\) and \((A, a)\), we say that \((B, b)\) is included in \((A, a)\) if \(B \subset A\) and \(b(x) \leq a(x)\ \forall x \in B\). In this case, \((B, b) = (A, \mathring{b})\) where

\[
\mathring{b}(x) = \begin{cases} 
  b(x) & \text{if } x \in B, \\
  0 & \text{if } x \notin B.
\end{cases}
\]

**Proposition 1.12** Let \(A\) be a finite set, \(|A| = n\). Let \((A, a)\) be a multiset on \(A\) and let \(k\) be a non-negative integer such that \(k \leq a(x)\ \forall x \in A\). The multisets included in \((A, a)\) with \(k\) elements are

\[
{n + k - 1 \choose k}.
\]

**Proof.** Let \(A = \{1, \ldots, n\}\). A multiset \(S\) of cardinality \(k\) included in \((A, a)\) contains the element 1 \(x_1\) times, the element 2 \(x_2\) times, and so on, with \(x_1 + x_2 + \cdots + x_n = k\). Moreover, the \(n\)-tuple \((x_1, \ldots, x_n)\) characterizes \(S\). We can associate to a \(n\)-tuple \((x_1, \ldots, x_n)\) the binary sequence

\[
\underline{00} \ldots \underline{01} \underline{00} \ldots \underline{01} \underline{00} \ldots \underline{01} \underline{00} \ldots \underline{0},
\]

where the symbol 1 denotes the fact that we are changing the element of \(A\). This is a binary word of length \(n + k - 1\) with \(k\) zeroes.

The correspondence described above is a one-to-one correspondence between the set of multisets of cardinality \(k\) included in \((A, a)\) and the set of binary words of length \(n + k - 1\) with \(k\) zeroes. There are exactly

\[
{n + k - 1 \choose k}
\]

different words of this kind, so that the claim is proven.

**1.2.4 Lists and functions**

Given a set \(A\), a list of \(k\) objects from the set \(A\) or a \(k\)-word with symbols in \(A\) is an ordered \(k\)-tuple of objects. For instance, if \(A = \{1, 2, 3\}\), then the 6-tuples \((1,2,3,3,2,1)\) and \((3,2,1,3,2,1)\) are two different 6-words of objects in \(A\). In these lists, or words, repetitions are allowed and the order of the objects is taken into account. Since each element of the list can be chosen independently of the others, there are \(n\) possible choices for each object in the list. Hence, the following holds.

**Proposition 1.13** The number of \(k\)-lists of objects from a set \(A\) of cardinality \(n\) is \(n^k\).

A function \(f : X \to A\) is defined by the value it assumes on each element of \(X\): if \(f : \{1, \ldots, k\} \to A\), then \(f\) is defined by the \(k\)-list \((f(1), f(2), \ldots, f(k))\),
which we refer to as the image-list or image-word of \( f \). Conversely, each \( k \)-list \((a_1, a_2, \ldots, a_k)\) with symbols from \( A \) defines the function \( f : \{1, \ldots, k\} \to A \) given by \( f(i) := a_i \ \forall i \). If \(|A|\) is finite, \(|A| = n\), we have a one-to-one correspondence between the set \( \mathcal{F}_n^k \) of maps \( f : \{1, \ldots, k\} \to A \), and the set of the \( k \)-lists with symbols in \( A \). Therefore, we have the following.

**Proposition 1.14** The number of functions in \( \mathcal{F}_n^k \) is \( F_n^k := |\mathcal{F}_n^k| = n^k \).

### 1.2.5 Injective functions

We use the symbol \( \mathcal{I}_n^k \) to denote the set of injective functions \( f : \{1, \ldots, k\} \to A \), \(|A| = n\), \( k \leq n \). Let \( I_n^k = |\mathcal{I}_n^k| \). Obviously, \( I_n^k = 0 \) if \( k > n \). The image-list of an injective function \( f \in \mathcal{I}_n^k \) is a \( k \)-word of pairwise different symbols taken from \( A \). To form any such image-list, one can choose the first entry among \( n \) elements, the second entry can be chosen among \( n - 1 \) elements, \( \ldots \), the \( k \)th entry can be chosen among the remaining \( n - k + 1 \) elements of \( A \), so that we have the following.

**Proposition 1.15** The cardinality \( I_n^k \) of \( \mathcal{I}_n^k \) is

\[
I_n^k = |\mathcal{I}_n^k| = n(n - 1) \cdot \ldots \cdot (n - k + 1) = k! \binom{n}{k} = \frac{n!}{(n - k)!}.
\]

Some of the \( I_n^k \)'s are in Figure 1.6.

### 1.2.6 Monotone increasing functions

Let \( \mathcal{C}_n^k \), \( k \leq n \), be the set of strictly monotone increasing functions \( f : \{1, \ldots, k\} \to \{1, \ldots, n\} \). The image-list of any such function is an ordered \( k \)-tuple of strictly increasing—hence pairwise disjoint—elements of \( \{1, \ldots, n\} \). The \( k \)-tuple is thus identified by the subset of the elements of \( \{1, \ldots, n\} \) appearing in it, so that we have the following.

\[
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \ldots \\
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \ldots \\
0 & 0 & 2 & 6 & 12 & 20 & 30 & 42 & 56 & 72 & \ldots \\
0 & 0 & 0 & 6 & 24 & 60 & 120 & 210 & 336 & 504 & \ldots \\
0 & 0 & 0 & 0 & 24 & 120 & 360 & 840 & 1680 & 3024 & \ldots \\
0 & 0 & 0 & 0 & 0 & 120 & 720 & 2520 & 6720 & 15120 & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 720 & 5040 & 20160 & 60480 & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 5040 & 40320 & 181440 & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40320 & 362880 & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 362880 & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots 
\end{pmatrix}
\]

**Figure 1.6** The cardinality \( I_n^k \) of the set of injective maps \( \mathcal{I}_n^k \) for \( n, k \geq 0 \).
Proposition 1.16 The cardinality $C_n^k$ of $C_n^k$ is

$$C_n^k := |C_n^k| = \binom{n}{k} = C_k^n = (C^T)_n^k.$$ 

1.2.7 Monotone nondecreasing functions

Let $D_n^k$ be the class of monotone nondecreasing functions $f : \{1, \ldots, k\} \rightarrow \{1, \ldots, n\}$. The image-list of any such function is a nondecreasing ordered $k$-tuple of elements of $\{1, \ldots, n\}$, so that elements can be repeated. The functions in $D_n^k$ are as many as the multisets with cardinality $k$ included in a multiset $(A, a)$, where $A = \{1, \ldots, n\}$ and $a(x) \geq k \ \forall x \in A$. Thus, see Proposition 1.12, we have the following.

Proposition 1.17 The cardinality $D_n^k$ of $D_n^k$ is

$$D_n^k := |D_n^k| = |\mathcal{C}_{n+k-1}^k| = \binom{n+k-1}{k}.$$ 

Another proof of Proposition 1.17. Consider the map $\phi : D_n^k \rightarrow \mathcal{F}_{n+k-1}^k$ defined by $\phi(f)(i) := f(i) + i - 1 \ \forall i \in \{1, \ldots, k\}, \ \forall f \in D_n^k$. Obviously, if $f \in D_n^k$, then $\phi(f)$ is strictly monotone increasing, $\phi(f) \in C_{n+k-1}^k$. Moreover, the correspondence $\phi : D_n^k \rightarrow C_{n+k-1}^k$ is one-to-one, thus

$$D_n^k = |D_n^k| = |\mathcal{C}_{n+k-1}^k| = \binom{n+k-1}{k}.$$ 

Yet another proof of Proposition 1.17. We are now going to define a one-to-one correspondence between a family of multisets and $D_n^k$. Let $(A, a)$ be a multiset on $A = \{1, \ldots, n\}$ with $a(x) \geq k \ \forall k$. For any multiset $(S, n_S)$ of cardinality $k$ included in $(A, a)$, let $f_S : A \rightarrow \{0, \ldots, k\}$ be the function defined by 

$$f_S(x) := \sum_{y \leq x} n_S(y),$$

i.e. for each $x \in A$, $f_S(x)$ is the sum of the multiplicities $n_S(y)$ of all elements $y \in A, y \leq x$. $f_S$ is obviously a nondecreasing function and $f_S(n) = k$. Moreover, it is easy to show that the map

$$S \mapsto f_S$$

is a one-to-one correspondence between the family of the multisets included in $(A, a)$ of cardinality $k$ and the family of monotone nondecreasing functions from $\{1, \ldots, n\}$ to $\{0, \ldots, k\}$ such that $f(k) = 1$. In turn, there is an obvious one-to-one correspondence between this class of functions and the class
of monotone nondecreasing functions from $\{1, \ldots, n-1\}$ to $\{0, \ldots, k\}$. Thus, applying Proposition 1.17 we get

$$|\mathcal{D}_{n-1}^{|k+1|} = \binom{k + 1 + (n - 1) - 1}{n - 1} = \binom{n + k - 1}{k}.$$  

### 1.2.8 Surjective functions

The computation of the number of surjective functions is more delicate. Let $\mathcal{S}_n^k$ denote the family of surjective functions from $\{1, \ldots, k\}$ onto $\{1, \ldots, n\}$ and let

$$\mathcal{S}_n^k = \begin{cases} 1 & \text{if } n = k = 0, \\ 0 & \text{if } n = 0, k > 0 \\ |\mathcal{S}_n^k| & \text{if } n \geq 1. \end{cases}$$

Obviously, $\mathcal{S}_n^k = |\mathcal{S}_n^k| = 0$ if $k < n$. Moreover, if $k = n \geq 1$, then a function $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ is surjective if and only if $f$ is injective, so that $S_n^n = |\mathcal{S}_n^n| = I_n^n = n!$.

If $k > n \geq 1$, then $\mathcal{S}_n^k \neq \emptyset$. Observe that any function is onto on its range. Thus, for each $j = 1, \ldots, n$, consider the set $A_j$ of functions $f : \{1, \ldots, k\} \to \{1, \ldots, n\}$ whose range has cardinality $j$. We must have

$$n^k = |\mathcal{F}_n^k| = \sum_{j=1}^{n} |A_j|.$$  

There are exactly $\binom{n}{j}$ subsets of $\{1, \ldots, n\}$ with cardinality $j$ and there are $\mathcal{S}_j^k$ different surjective functions onto each of these sets. Thus, $|A_j| = \binom{n}{j} \mathcal{S}_j^k$ and

$$n^k = \sum_{j=1}^{n} \binom{n}{j} \mathcal{S}_j^k \quad \forall n \geq 1.$$  

Since we defined $\mathcal{S}_0^k = 0$, we get

$$n^k = \sum_{j=0}^{n} \binom{n}{j} \mathcal{S}_j^k \quad \forall n \geq 0. \tag{1.5}$$

Therefore, applying the inversion formula in Corollary 1.5 we conclude the following.

**Proposition 1.18** The cardinality $\mathcal{S}_n^k$ of the set $\mathcal{S}_n^k$ of surjective functions from $\{1, \ldots, k\}$ onto $\{1, \ldots, n\}$ is

$$\mathcal{S}_n^k = \sum_{j=0}^{n} (-1)^{n+j} \binom{n}{j} j^k = \sum_{j=0}^{n} (-1)^{j} \binom{n}{j} (n - j)^k \quad \forall n, k \geq 1.$$
We point out that the equality holds also if \( k \leq n \) so that

\[
\frac{1}{n} \sum_{j=0}^{n} (-1)^j \binom{n}{j} (n-j)^k = \frac{1}{n!} S_n^k = \begin{cases} 1 & \text{if } k = n, \\ 0 & \text{if } k < n. \end{cases}
\]

Another useful formula for \( S_n^k \) is an inductive one, obtained starting from \( S_n^n = n! \forall n \geq 0 \) and \( S_n^k = 0 \) for any \( k \) and \( n \) with \( k < n \).

**Proposition 1.19** We have

\[
\begin{align*}
S_n^k &= n(S_n^{k-1} + S_{n-1}^{k-1}) & \text{if } k \geq 1, n \geq 0, \\
S_n^n &= n!, \\
S_0^k &= 0 & \text{if } k \geq 1.
\end{align*}
\] (1.6)

**Proof.** Assume \( n \geq 1 \) and \( k \geq 1 \) and let \( f : \{1, \ldots, k\} \to \{1, \ldots, n\} \) be a surjective function. Let \( A \subset S_n^k \) be the class of functions such that the restriction \( f : \{1, \ldots, k-1\} \to \{1, \ldots, n\} \) of \( f \) is surjective and let \( B := S_n^k \setminus A \). The cardinality of \( A \) is \( nS_{n-1}^{k-1} \) because there are \( S_{n-1}^{k-1} \) surjective maps from \( \{1, \ldots, k-1\} \) onto \( \{1, \ldots, n\} \) and there are \( n \) possible choices for \( f(k) \). Since the maps on \( B \) have a range of \( (n-1) \) elements, we infer that there are \( nS_{n-1}^{k-1} \) maps of this kind. In fact, there are \( \binom{n}{n-1} = n \) subsets \( E \) of \( \{1, \ldots, n\} \) of cardinality \( n-1 \) and there are \( S_{n-1}^{k-1} \) surjective functions from \( \{1, \ldots, k-1\} \) onto \( E \). Therefore,

\[
S_n^k = |A| + |B| = nS_{n-1}^{k-1} + nS_{n-1}^{k-1}.
\]

i.e. (1.6).

Some of the \( S_n^k \)'s are in Figure 1.7.

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & 1 & 6 & 6 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & 1 & 14 & 36 & 24 & 0 & 0 & 0 & 0 & \ldots \\
0 & 1 & 30 & 150 & 240 & 120 & 0 & 0 & 0 & \ldots \\
0 & 1 & 62 & 540 & 1560 & 1800 & 720 & 0 & 0 & \ldots \\
0 & 1 & 126 & 1806 & 8400 & 16800 & 15120 & 5040 & 0 & \ldots \\
0 & 1 & 254 & 5796 & 40824 & 126000 & 191520 & 141120 & 40320 & \ldots \\
0 & 1 & 510 & 18150 & 186480 & 834120 & 1905120 & 2328480 & 1451520 & \ldots \\
\end{pmatrix}
\]

*Figure 1.7* The cardinality \( S_n^k \) of the set of surjective maps \( S_n^k \) for \( n, k \geq 0 \).
1.2.9 Exercises

Exercise 1.20 How many diagonals are there in a polygon having $n$ edges?

1.3 Drawings

A drawing or selection of $k$ objects from a population of $n$ is the choice of $k$ elements among the $n$ available ones. We want to compute how many of such selections are possible. In order to make this computation, it is necessary to be more precise, both on the composition of the population and on the rules of the selection as, for instance, if the order of selection is relevant or not. We consider a few cases:

- The population is made by pairwise different elements, as in a lottery: in other words, the population is a set.
- The population is a multiset $(A, a)$. In this case, we say that we are dealing with a drawing from $A$ with repetitions.
- The selected objects may be given an order. In this case we say that we consider an ordered selection. Unordered selections are also called simple selections.

Some drawing policies simply boil down to the previous cases:

- In the lottery game, numbers are drawn one after another, but the order of drawings is not taken into account: it is a simple selection of objects from a set.
- In ordered selections the $k$-elements are selected one after another and the order is taken into account.
- A drawing with replacement, i.e. a drawing from a set where each selected object is put back into the population before the next drawing is equivalent to a drawing with repetitions, i.e. to drawing from a multiset where each element has multiplicity larger than the total number of selected objects.

1.3.1 Ordered drawings

Ordered drawings of $k$ objects from a multiset $(A, a)$ are $k$-words with symbols taken from $A$.

1.3.1.1 Ordered drawings from a set

Each ordered drawing of $k$ objects from a set $A$ is a $k$-list with symbols in $A$ that are pairwise different. Thus the number of possible ordered drawings of $k$ elements from $A$ is the number of $k$-lists with pairwise different symbols in $A$. 