TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

D I P L O M A R B E I T

Rank-Based Strategies for the Hiring Problem

Ausgeführt am Institut für
Diskrete Mathematik und Geometrie
der Technischen Universität Wien

unter der Anleitung von
Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Alois Panholzer

durch
Clemens Tomsic
Roseggergasse 10
7350 Oberpullendorf

Datum

Unterschrift

MASTER THESIS

Rank-Based Strategies for the Hiring Problem

Author:
Clemens Tomsic

Supervisor:
Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Alois Panholzer

Abstract

This thesis analyzes the behavior of the hiring problem with the restriction to rank-based strategies. Given a sequence of candidates arriving to an interview, we have to make an immediate and irrevocable decision about hiring a candidate, when he is taking his interview. This decision has to be made on the relative rank of this candidate amongst all already interviewed candidates. For the decision on a hiring we define hiring strategies that may be described as an algorithm applied to the sequence of candidates. The main focus lies within the analysis of different strategies. Therefore we introduce hiring parameters that characterize the behavior of the single strategies. Concretely, we distinguish between parameters that give information on the hiring rate and parameters that analyze the quality of the hired stuff. We also consider a relaxation of the hiring problem, the so-called hiring and firing, where hired candidates can be replaced by other candidates during the hiring process.

When analyzing the strategies we are mainly interested in the probabilities and the expected values of the hiring parameters, such as the size of the hiring set, the waiting time until a certain number of candidates is recruited, the rank of the best discarded candidate or the number of replacements for hiring and firing. Therefore we use a wide range of tools like generating functions or complex analytic methods. We also determine the limit distributions of the parameters when the length of the sequence of applicants grows to infinity.

The investigated strategies are hiring above the minimum, hiring above the maximum, hiring above the m-th best, hiring above the median, as well as a probabilistic strategy. In all of the strategies a candidate is hired if and only if his relative rank is better than the rank of a certain threshold candidate. The threshold candidate may change during the hiring process and the strategies mainly differ in how the current threshold candidate is chosen. We can observe that each of the strategies has its own benefits, so it depends on the requirements which strategy is the best to use. We especially like to emphasize on the probabilistic strategy, which provides a connection between all the other strategies.

Zusammenfassung

Diese Arbeit behandelt das Rekrutierungsproblem mit der Einschränkung auf rangbasierte Strategien. Wir betrachten eine Folge von Kandidaten, die zu einem Vorstellungsgespräch eingeladen sind. Bei diesem Vorstellungsgespräch müssen wir eine sofortige und unwiderrufliche Entscheidung über die Einstellung des Kandidaten treffen. Dazu reihen wir die Kandidaten nach deren Beurteilung und treffen die Entscheidung anhand des relativen Ranges des Kandidaten unter allen bisher interviewten Kandidaten. Für die Entscheidung verwenden wir Strategien, die man als Algorithmen betrachten kann. Der größte Teil der Arbeit befasst sich mit der Analyse von verschiedenen Strategien zur Entscheidungsfindung. Dazu definieren wir Parameter, die das Verhalten der verschiedenen Strategien beschreiben. Diese Parameter messen einerseits die Frequenz der Einstellungen und andererseits die Qualität der eingestellten Kandidaten. Weiters betrachten wir eine Variante des Rekrutierungsproblems, bei welcher bereits eingestellte Kandidaten später durch neue Kandidaten ersetzt werden können.

Der Großteil der Arbeit beschäftigt sich mit der Analyse der exakten Wahrscheinlichkeitsverteilungen der einzelnen Parameter. Dazu benötigen wir viele Hilfsmittel aus verschiedenen Bereichen der Mathematik, wie zum Beispiel erzeugende Funktionen oder komplexe Analysis. Weiters sind wir aber auch an den Grenzverteilungen der Parameter interessiert, wenn die Anzahl der Bewerber gegen unendlich wächst.

Die betrachteten Strategien sind Rekrutierung über dem Minimum, Rekrutierung über dem Maximum, Rekrutierung über dem m-ten besten rekrutierten Kandidaten, Rekrutierung über dem Median und eine probabilistische Strategie. Jede dieser Strategien arbeitet nach der Regel, dass ein Kandidat nur dann eingestellt wird, wenn er besser als ein bestimmter Schwellwert, gegeben durch den Rang eines bereits rekrutierten Kandidaten, ist. Dieser Schwellwert, bzw. der Kandidat, kann sich mit der Zeit verändern. Der Hauptunterschied zwischen den einzelnen Strategien liegt in der Auswahl des Schwellwertes, bzw. dem zugehörigen Kandidaten. Es stellt sich heraus, dass jede der einzelnen untersuchten Strategien ihre eigenen Vor- und Nachteile hat, daher ist es von den gestellten Anforderungen abhängig, welche Strategie am geeignetsten ist. Besonders möchten wir auf die probabilistische Strategie hinweisen, welche eine Verbindung zwischen all den anderen untersuchten Strategien herstellt.

Acknowledgements

First of all I would like to thank my supervisor, Dr. Alois Panholzer, for the great support while working on this thesis.

Then I want to thank my parents, Susanna and Erwin Tomsic, for all their patience and moral as well as financial support during my education. I would not have been able to do all this without their support.

I especially want to thank my beloved grandmother, Theresia Töpfer, who always believed in me and inspired me in my studies.

Last but not least I want to thank my fellow companions, Birgit Ondra and Michael Wallner, for the awesome years during my studies.

Preface

The hiring problem is a recent research problem that belongs to the category of on-line selection under uncertainty. The process of the hiring problem is quite simple. We imagine a company hiring new employees. Thus, we consider a linear sequence of candidates arriving to an interview. The company has to make an immediate and irrevocable decision whether to hire a candidate or not. This implies that this decision has to be made just on the quality of the candidates seen so far without information about the future. For the decision of hiring a candidate or not the company ranks the candidates from best to worst and the decision has to be made on the relative rank of the currently interviewed candidate. This restriction transforms the problem into a discrete problem which gives us the chance to analyze it with combinatorial techniques.

There are several fascinating aspects about the hiring problem. First of all one may easily construct several different ways to make a decision about a hiring. In my opinion the most interesting aspect is the interaction between the different fields of mathematics when analyzing the problem. As we restrict ourselves to the combinatorial version of the hiring problem we may obviously apply combinatorial techniques. Other fascinating disciplines like complex analysis or higher probabilistic theory appear very often in our considerations as well.

In Chapter 1 we give a precise definition of the hiring problem and construct a framework for the analysis of hiring strategies. Furthermore we introduce hiring parameters, which are random variables that give information about the hiring process depending on the used strategy. The most fundamental parameter is the size of the hiring set, which indicates how many candidates get hired by a specific strategy. Other parameters like the waiting time or the distance between the last two hirings give information about how the hiring process is expected to evolve when the number of candidates increases, while the score of the best discarded candidate and the rank of the last hired candidate investigate the expected quality of the recruited candidates. We also introduce a modification of the hiring problem, the so-called hiring and firing, where we add a replacement mechanism. Thus, we relax the restriction that the decision of a hiring is irrevocable. If a candidate gets discarded by the strategy, although he is better than the least ranked already hired candidate, he replaces this candidate. The number of replacements gives a good measurement of how selective the hiring process is. We also state some important notations and theorems we will need from the different fields of mathematics in this chapter.

Chapter 2 analyzes a first rather simple strategy, the so-called hiring above the minimum. Here the first candidate gets hired and any further candidate gets hired if and only if he is better than the worst already hired candidate. Thus, there is a strong dependency on the quality of the first candidate. The hiring process is rather simple and can mostly be analyzed by using combinatorial considerations.

The main part of this work can be found in the Chapters 3 and 4 which form the category of Lake Wobegon Strategies. This type of strategy is defined by the property that the quality of the hiring set always increases when recruiting a new candidate. In Chapter 3 we consider hiring above the maximum and hiring above the m-th best, which are two related strategies. In these strategies we hire the first, respectively the first m candidates, and other candidates get hired if they are better than the best, respectively the m-th best, already hired candidate. The benefit of hiring above the m-th best is, that the parameter m can be chosen arbitrarily. Thus, one can adjust the hiring process to his personal conditions.

Chapter 4 analyzes the strategy hiring above the median. Here the first candidate gets hired and others get hired if their rank is better then the median score of all already hired candidates. The process is a more complicated one as one does not know the median score of all already hired candidates. However, by an approach using an automaton Helmi and Panholzer were able to solve this problem.

Chapter 5 gives a probabilistic hiring process and arose from personal correspondence with Dr. Alois Panholzer. In particular we hire the first candidate and others get hired if they are better than a threshold candidate. This threshold is a candidate from the current hiring set. When hiring a new candidate there is a chance of $p \in [0, 1]$ that the threshold changes to the very next better already hired candidate. Because of the probability p it is rather difficult to investigate this strategy. However, we could determine the asymptotic behavior of the hiring set. We especially investigated the connection of the probabilistic approach to the other strategies.

Clemens Tomsic

Contents

Abstract **i**

Zusammenfassung **ii**

Acknowledgements **iii**

Preface **iv**

1 Preliminaries **1**

1.1 The hiring problem . . . 1

1.2 Hiring parameters . . . 6

1.2.A Rate parameters . . . 6

1.2.B Quality parameters . . . 7

1.2.C Hiring and firing . . . 9

1.3 Notational conventions and asymptotics . . . 10

1.4 Formal power series and generating functions . . . 13

1.5 Complex analysis . . . 16

1.6 Probabilistic methods . . . 19

2 Hiring above the minimum **21**

2.1 Hiring above the minimum . . . 21

3 Hiring for the elite **30**

3.1 Hiring above the maximum . . . 31

3.2 Hiring above the m-th best . . . 37

4 Hiring above the median **60**

4.1 Hiring above the median . . . 60

5 Probabilistic hiring **86**
5.1 Probabilistic change of the threshold 86
5.2 Connection between the probabilistic and deterministic strategies 94
5.2.A Hiring above the maximum 94
5.2.B Hiring above the minimum 95
5.2.C Hiring above the median 95

Bibliography **97**

Chapter 1

Preliminaries

1.1 The hiring problem

The hiring problem is a recent research problem that was first introduced by Broder et al. in 2008 [3] and belongs to the category of on-line selection under uncertainty. It is closely related to the *secretary problem* (see [10, 11]) and the *Chinese restaurant process (CRP)* (see [13, 26]). Here we will formulate the hiring problem and set up a framework to consider several hiring strategies. In our model we will consider *rank-based* strategies only. This is a different approach than in [3] and allows us to consider the combinatorial aspects of the hiring problem. The benefit of this alternative formulation is the application of a variety of combinatorial methods like generating functions. In our notation we will mostly follow [1, 14, 15] who introduced this combinatorial approach.

Let us consider an expanding company that is searching for new employees. The applicants are invited to an interview and arrive one after another. After the interview the company has to take an irrevocable and immediate decision whether to hire the candidate or not. This implies that the company has no information about the future when interviewing a candidate so this decision must be based on the interviews so far.

Formally we may consider a sequence $X_1, X_2, \ldots, X_n$ of unknown length n concerning candidates invited to the interview. When taking the interview the i-th candidate is assigned a quality score Q_i. The quality scores may be interpreted as independent and identically distributed random variables. As we are mainly interested in combinatorial aspects we will not care about the quality score of a candidate itself but about the relative rank the candidate gets among all interviewed persons so far, i.e. the candidates are ranked from best to worst without ties, the best candidate gets rank n and the worst gets rank 1 - the higher the score of a candidate the higher the rank.

The central tool when ranking the candidates are permutations. In a first step when modeling this ranking could be a mapping which assigns each candidate to a rank. As we are mostly interested in the index of the candidate we may interpret this mapping as a permutation on the set of indices. The following definition issues this situation formally.

Definition 1.1.1. Let $X_1, X_2, \dots, X_n$ denote the sequence of the candidates arriving to the interview. Then the ranking of the applicants is denoted by

$$\sigma : \{X_1, \dots X_n\} \to \{1, \dots, n\}$$
$$X_i \mapsto \sigma(X_i).$$

We call $\sigma(X_i)$ the *rank* of candidate X_i. For convenience we will always denote the rank by $\sigma(i)$ as we are only interested in the index of the candidate, i.e. we consider σ as a permutation on the set $\{1, \dots, n\}$, what we will denote by $\sigma \in \mathcal{S}_n$. Furthermore, we will use sequential notation $\sigma = \sigma(1)\sigma(2)\dots\sigma(n)$ for permutations and we will write $|\sigma| = n$ for the length of the permutation.

□

For any candidate X_i the rank $\sigma(i)$ denotes the absolute position of the candidate's quality among all interviewed candidates. As we do not know the number of applicants when interviewing - and possibly hiring - a candidate, we are not interested in the absolute but the relative rank among the candidates $X_1, X_2, \dots, X_i$. Therefore it is useful to build up the permutation σ from step to step when interviewing candidates. This leads to the following definition which we get in a natural way.

Definition 1.1.2 (Permutation of the relative ranks). For a permutation σ on the ranks $\{1, 2, \dots, n\}$ we define the *permutation of the relative ranks* after $i \leq n$ interviewed candidates $\rho_i(\sigma)$ in the following inductive way

- $i = 1$: $\rho_1(\sigma) = 1$
- $i > 1$: $\rho_i(\sigma) = \rho_{i-1}(\sigma) \circ j$, where $j = |\{k \in \{1, 2, \dots i\} : \sigma(k) \leq \sigma(i)\}|$, i.e. denotes the *relative rank of* X_i amongst the first i candidates.

The operation $\circ$ is defined by

$$\circ : \mathcal{S}_{i-1} \times \{1, 2, \dots, i\} \to \mathcal{S}_i$$
$$(\tau \circ j)(k) = \begin{cases} \tau(k) + \mathbb{1}_{[\tau(k) \geq j]}, & 1 \leq k \leq i-1, \\ j, & k = i. \end{cases}$$

Starting with a permutation $\tau \in \mathcal{S}_{i-1}$ we obtain the permutation $\tau \circ j$ by increasing $\tau(k)$ by 1 if and only if $\tau(k) \geq j$, $1 \leq k \leq i-1$. For $k = i$ we have $(\tau \circ j)(i) = j$.

□

Note that the operation $\circ$ is well-defined as we get the relative ranks of the first i candidates at each step and $\rho_n(\sigma) = \sigma$. We can easily observe this in the following example which we will also consider for the hiring strategies later on.

Example 1.1.3. We want to reconstruct the permutation $\sigma = 463921785$ by considering the relative ranks after each interview. This yields

candidate	absolute rank	relative rank	$\rho_i(\sigma)$
1	4	1	1
2	6	2	12
3	3	1	231
4	9	4	2314
5	2	1	34251
6	1	1	453621
7	7	6	4537216
8	8	7	45382167
9	5	5	463921785

In the last step we reconstructed the permutation σ. This is a very important property for our purposes as we may compare the current candidate to all those who have already been interviewed but there is no information about the further candidates.

■

Note that in fact we do not know the permutation σ in the beginning but by interviewing candidates we know the relative rank at each step i and $\rho_i(\sigma)$. This allows us to build up the permutation step by step and in the end we receive σ. For convenience it is easier to consider a given permutation and reconstruct it like done in the example before.

When interviewing a candidate we will write $X_i = 0$ and say candidate i is *discarded* or $X_i = 1$ and call candidate i *hired*. This allows us to interpret a candidate as a random variable X_i in a natural way depending on whether the candidate gets hired or not. Given a permutation σ we are now able to define hiring strategies. A hiring strategy may be described as an algorithm that decides if we hire a candidate or not respecting the rules we have established in our formulation of the hiring problem.

Definition 1.1.4 (Hiring set, hiring strategy). Let $n \in \mathbb{N}$ and $\sigma \in \mathcal{S}_n$ denote the permutation of the ranks in a sequence of n candidates. We call the set of indices of hired candidates the *hiring set* and denote it by $\mathcal{H}(\sigma) = \{i : X_i = 1\}$. As the hiring set is generated while interviewing candidates we write $\mathcal{H}_i(\sigma)$ for the current hiring set at step i. Obviously we have $\mathcal{H}_n(\sigma) = \mathcal{H}(\sigma)$. A *hiring strategy* builds up the hiring set respecting the following rules.

$$i) \quad \mathcal{H}_0(\sigma) = \emptyset \tag{1.1}$$

$$ii) \quad \mathcal{H}_i(\sigma) \subseteq \{1, 2, \ldots, i\}, \quad \forall\, i \leq n \tag{1.2}$$

$$iii) \quad \mathcal{H}_i(\sigma) \setminus \{i\} = \mathcal{H}_{i-1}(\sigma) \tag{1.3}$$

$$iv) \quad \mathcal{H}_i(\sigma) = \mathcal{H}_i(\pi), \quad \forall\, \sigma, \pi \in \mathcal{S}_n : \ \sigma(j) = \pi(j) \quad \forall\, j \in \{1, 2, \ldots, i\} \tag{1.4}$$

These are exactly the rules we claimed when formulating the problem. The Constraints (1.1) and (1.2) state that no future candidate can be hired at step i. Condition (1.3) ensures that no previous candidate may get hired. All in all these three rules state that candidate i may only get hired in step i. It depends on the strategy itself whether the candidate is added to the hiring set (i.e. the candidate gets hired). Rule (1.4) simply guarantees that decisions are made without information about the future because all permutations σ, π that have the same relative ranks in at least the first i steps result in the same hiring sets at step i.

□

For our purposes the absolute ranks are not really important. We rather like to consider the relative ranks. This leads to the following tightening of (1.4).

Definition 1.1.5 (Rank-based strategy). In the setting of Definition 1.1.4 we call a hiring strategy *rank-based* if and only if for all permutations $\sigma \in \mathcal{S}_n$ and all i, $1 \leq i \leq n$, it holds that

$$\mathcal{H}_i(\sigma) = \mathcal{H}(\rho_i(\sigma)).$$

Obviously (1.4) holds for rank based strategies. In particular consider $\sigma, \pi \in \mathcal{S}_n : \sigma(j) = \pi(j) \quad \forall j \in \{1, 2, \ldots, i\}$. This yields $\mathcal{H}_i(\sigma) = \mathcal{H}(\rho_i(\sigma)) = \mathcal{H}(\rho_i(\pi)) = \mathcal{H}_i(\pi)$.

□

The definition above is not really a strict tightening but it gives us consistency with our notation when building up a permutation and allows us to consider the relative ranks at each step. Let us now consider a first very simple rank-based strategy.

Example 1.1.6 (Hiring by even index). Let $n \in \mathbb{N}$ and $X_1, X_2, \ldots, X_n$ denote a sequence of candidates. *Hiring by even index* hires candidates applying the following rule:

$$X_i = \begin{cases} 0, & i \equiv 1(2) \\ 1, & i \equiv 0(2) \end{cases}$$

This means that each candidate having an even index gets hired which yields

$$\mathcal{H}_i(\sigma) = \begin{cases} \{2, 4, \ldots i-1\}, & i \equiv 1(2) \\ \{2, 4, \ldots i\}, & i \equiv 0(2) \end{cases}$$

Obviously the strategy fulfills all the conditions given in Definition 1.1.4 and is actually rank-based as it is completely independent of the permutation σ.

■

The strategy given in the example above is actually not interesting at all and it is not sensible to observe its behavior. It just shows that it is easy to find strategies that serve our definition of a hiring strategy. The problem with 'hiring by even index' is that it is completely independent of the permutation at all, i.e. interviewing the candidates is absolutely senseless. This leads us to the next definition.

Definition 1.1.7 (Pragmatic strategy). Consider a permutation $\sigma \in \mathcal{S}_n$ of the ranks in a sequence of n candidates. Let $\mathcal{R}(\sigma) := |\{j : 1 \leq j \leq n+1 \ \wedge \ (n+1) \in \mathcal{H}(\sigma \circ j)\}|$ denote the number of ranks so that candidate X_{n+1} is hired right after σ, i.e. with respect to σ. Furthermore, we denote the indicator variable for rank j getting hired right after σ by

$$\mathcal{R}_j(\sigma) = \begin{cases} 1, & (n+1) \in \mathcal{H}(\sigma \circ \ j), \\ 0, & \text{else.} \end{cases}$$

A hiring strategy is called *pragmatic* if and only if the following holds for all permutations σ and all $1 \leq j \leq |\sigma| + 1$:

$$i) \quad \mathcal{R}_j(\sigma) = 1 \ \Rightarrow \ \mathcal{R}_k(\sigma) = 1 \quad \forall \ k \geq j, \tag{1.5}$$

$$ii) \quad \mathcal{R}(\sigma \circ j) \leq \mathcal{R}(\sigma) + \mathcal{R}_j(\sigma), \tag{1.6}$$

$$iii) \quad \mathcal{R}_j(\sigma) = 0 \ \Rightarrow \ \mathcal{R}(\sigma \circ j) = \mathcal{R}(\sigma). \tag{1.7}$$

Constraint (1.5) states whenever a strategy hires a candidate with score j it would hire a candidate with a higher score as well, while (1.6) simply sets up a boundary for the possible ranks to be hired, i.e. there must not be a big difference between the hiring criteria for consecutive interviews. Last but not least (1.7) causes that a discarded candidate does not affect the hiring criteria for the next candidate.

□

As one can easily see 'hiring by even index' is no pragmatic strategy. For this we consider a permutation σ satisfying $|\sigma| \equiv 0(2)$. Then $\mathcal{R}(\sigma) = \mathcal{R}_j(\sigma) = 0$ holds for $1 \leq j \leq |\sigma| + 1$ but $\mathcal{R}(\sigma \circ j) = |\sigma| + 2$ which contradicts (1.6).

Remark 1.1.8. The quantities $\mathcal{R}(\sigma)$ and $\mathcal{R}_j(\sigma)$ respectively characterize the hiring criteria of a hiring strategy. A strategy may be defined under specifications of $\mathcal{R}(\sigma)$ for each permutation σ.

■

Using these variables we obtain the first information about the hiring set and ranks that are hired for sure.

Theorem 1.1.9 (Archibald and Martínez [1], Theorem 2). For any pragmatic strategy and a permutation $\sigma \in \mathcal{S}_n$ of the ranks it holds that

$$\sigma(i) \geq n + 1 - \mathcal{R}(\sigma) \quad \Rightarrow \quad i \in \mathcal{H}(\sigma).$$

So the best $\mathcal{R}(\sigma)$ candidates are hired for sure.

Proof. The statement can be easily shown by induction on n:

- $n = 0$: As σ is the empty permutation the hiring set $\mathcal{H}(\sigma)$ contains the best $\mathcal{R}(\sigma)$ candidates.
- $n-1 \to n$: We consider a permutation $\sigma \in \mathcal{S}_n$ that is built up by $\rho_{n-1}(\sigma) \circ j$ for some $j \in \{1, 2, \ldots, n\}$,i.e. $\sigma(n) = j$. By inductive hypothesis $\mathcal{H}(\rho_{n-1}(\sigma))$ contains the best $\mathcal{R}(\rho_{n-1}(\sigma))$ candidates. As we are considering pragmatic strategies only we have that

$$X_n = \begin{cases} 1, & j \geq n + 1 - \mathcal{R}(\rho_{n-1}(\sigma)) \\ 0, & \text{else.} \end{cases}$$

  Case 1: $X_n = 1$. In this case it is true that $j \geq n + 1 - \mathcal{R}(\rho_{n-1}(\sigma))$, so X_n belongs to the best $\mathcal{R}(\rho_{n-1}(\sigma))$+1 candidates of $\mathcal{H}(\sigma)$. By Definition 1.1.7 we have $\mathcal{R}(\sigma) \leq \mathcal{R}(\rho_{n-1}(\sigma)) + 1$, which implies that $\mathcal{H}(\sigma)$ contains the best $\mathcal{R}(\sigma)$ candidates of σ.

  Case 2: $X_n = 0$. It is true that $j < n + 1 - \mathcal{R}(\sigma)$ and as a consequence the relative scores of the best $\mathcal{R}(\rho_{n-1}(\sigma))$ increase by one. By inductive hypothesis it follows that the best $\mathcal{R}(\rho_{n-1}(\sigma))$ are hired and (1.7) yields $\mathcal{R}(\sigma) = \mathcal{R}(\rho_{n-1}(\sigma))$.

Thus, the hiring set contains of at least the best $\mathcal{R}(\sigma)$ candidates.

□

We will only consider strategies that hire a candidate if and only if his relative rank is better than the relative rank of some tagged already hired candidate, called the *threshold candidate*. Depending on the strategy the threshold candidate may change in some specific way when a new candidate is hired. It is a routine task to prove this kind of strategies being pragmatic.

1.2 Hiring parameters

Before we start observing pragmatic hiring strategies and their properties we define *hiring parameters* in this section. Hiring parameters are essential quantities that characterize the behavior of a hiring strategy. On the one hand we are mainly interested in the exact distributions and the expectation of the quantities for the given strategy and on the other hand we will also investigate its asymptotic behavior. We will mainly consider the hiring parameters defined in [14, 15] and use the same notation. Most of the parameters were already investigated in [1, 3, 20] as well. Following [14] we may characterize three kinds of parameters. The first ones study the behavior of a hiring strategy, while the second ones give information about the quality measure of the hired candidates. The third category is actually concerned with a modification of the hiring problem, the so-called *hiring and firing*. We will use the same notation for the parameters for each hiring strategy and possibly add other indices to emphasize specific dependencies for certain strategies. For example, we will add the parameter m with *hiring above the m-th best*.

1.2.A Rate parameters

Our first and most important parameter is given by the *size of the hiring set* which is also called the *number of hired candidates*. This quantity represents the main focus of our considerations as it is not only our main interest but also important for all the other parameters. All the strategies we will consider hire the first candidate anyway no matter which score. This yields that there is always at least one hired candidate.

Definition 1.2.1 (Number of hired candidates). Consider a sequence $X_1, X_2, \ldots, X_n$ of candidates. Then we denote the *number of hired candidates* (or *size of the hiring set*) by h_n.

□

Like in [1] it is useful to consider the following bivariate generating function to compute the size of the hiring set h_n in most cases.

$$H(z,u) = \sum_{n=1}^{\infty} \sum_{\sigma \in \mathcal{S}_n} \frac{z^n}{n!} u^{h(\sigma)}, \tag{1.8}$$

where $h(\sigma) := |\mathcal{H}(\sigma)|$ denotes the number of hired candidates for the specific permutation σ concerning the investigated strategy. We will deal with this function and with generating functions in general in Section 1.4.

Another very interesting parameter closely related to h_n is the *hiring rate* which is in fact just the scaling of h_n and is given by $\frac{h_n}{n} \in [0, 1]$. As one may easily see this indicates how many candidates we hire in relation to the number n of interviewed candidates.

At the beginning we considered a company looking for new personnel. Companies often search for a certain number of new employees which leads us to the next parameter, the *waiting time*. This quantity gives information about how many interviews we probably have to make until a given number N of candidates get hired. Obviously this quantity is dependent on the hiring rate for the considered strategy. A high hiring rate will lead to a low waiting time and vice versa.

Definition 1.2.2 (Waiting time). For a sequence of candidates $X_1, X_2, \dots$ and the permutation of their ranks σ we denote the *waiting time* until N candidates get hired by W_N. Formally W_N can be described in the following way.

$$W_N = n \;\Leftrightarrow\; h_n(\sigma) = N \;\wedge\; X_n = 1$$

Here the first condition says that exactly N candidates are hired at step n and the second one says that candidate X_n gets hired. This means that for all $k < n$ there are less than N candidates hired, i.e. $h_k < N$. Obviously $N \geq n$ holds as no future candidates may get hired at any step.

□

For growing number n of candidates we expect a good strategy not to hire too many candidates. At the beginning we are not able to decide whether the candidate has a good absolute score but in each step we gain more information. That is why candidates taking their interview late will only get hired when having a good *absolute* rank. So the probability for a hiring will decrease gradually. The next two parameters give information about that situation.

Definition 1.2.3 (Index of the last hired candidate). For a sequence $X_1, X_2, \dots, X_n$ of candidates and the permutation of their ranks σ we denote the *index of the last hired candidate* by

$$L_n = \max \mathcal{H}(\sigma).$$

□

The index of the last candidate gives a measure for the quality of the recruited persons. A low index L_n could mean that there are many good candidates hired at the beginning, so it is hard for candidates arriving later to become hired. The last quantity describing the behavior of the strategy is the distance of the last two hirings. Together with the index of the last hired candidate we may deviate information about how the hiring rate changes in time.

Definition 1.2.4 (Distance between last two hirings). Consider a sequence of candidates given by $X_1, X_2, \dots, X_n$, and let $\sigma \in \mathcal{S}_n$ denote the permutation of their ranks. Then the *distance between the last two hirings* Δ_n is given by

$$\Delta_n = \begin{cases} 0, & h_n \leq 1 \\ L_n - \max(\mathcal{H}_n(\sigma) \setminus \{L_n\}), & \text{else.} \end{cases}$$

The distance is given by the number of candidates not hired between the two maximum indices in the hiring set plus one and tells us how many interviews we probably have to make until the next candidate will get hired.
Note: As there is no distance between two hirings for $h_n = 1$ we define $\Delta_n = 0$ in that case.

□

1.2.B Quality parameters

Quality parameters do not give information about the hiring set itself but about the (relative) ranks of the hired candidates. As the name says they form a good measure about the quality of the recruited staff. As we already debated before it will become more and more difficult to get hired for candidates who arrive later. The rank of the last hired candidate gives a good reflexion about this statement.

Definition 1.2.5 (Rank of the last hired candidate). Let L_n denote the index of the last hired candidate in a sequence consisting of n candidates whose ranks are described by a permutation $\sigma \in \mathcal{S}_n$. Then the *rank of the last hired candidate* is given by

$$R_n = \sigma(L_n).$$

□

We will often be confronted with a situation where a candidate holding a good quality score gets discarded although there exist candidates with a worse score that have already been hired. This is a situation we consider more detailed in Section 1.2.C. Our next parameter deals with the quality of discarded candidates or more precisely with their maximum score. A strategy that leads to a high score of the best discarded candidate obviously recruits only candidates having an excellent quality score. If the rank of the best discarded candidate is low on the other hand the strategy hires too many candidates.

Definition 1.2.6 (Score of the best discarded candidate). Consider the set of discarded candidates in a sequence of candidates $X_1, X_2, \dots, X_n$ which is given by $\mathcal{D}_n(\sigma) := \{1, 2, \dots, n\} \setminus \mathcal{H}_n(\sigma)$, whereat σ denotes the permutation of their ranks. Then we define the *score of the best discarded candidate* by

$$M_n = \begin{cases} 0, & h_n = n \\ \max\{\sigma(j) : \ j \in \mathcal{D}_n(\sigma)\}, & \text{else.} \end{cases}$$

Note: As there are no discarded candidates for $h_n = n$ we define $M_n = 0$ in that case.

□

A very interesting quantity is given by *the gap*. This quantity was already studied in [1,3,14,15] and gives a measure of the increase of the quality of the hiring set as it compares the difference between the relative rank of the last hired candidate and the maximum possible rank.

Definition 1.2.7 (Gap of the last hired candidate). In a sequence of candidates $X_1, X_2, \dots, X_n$ the *gap of the last hired candidate* is given by

$$g_n = 1 - \frac{R_n}{n}$$

which is the normalized difference between n and R_n.

□

Following [1] the expectation of the gap in a pragmatic hiring strategy can be generalized by the next theorem.

Theorem 1.2.8 (Archibald and Martínez [1], Theorem 3). For any pragmatic hiring strategy the following holds:

$$\mathbb{E}(g_n) = \frac{1}{2n}(\mathbb{E}(\mathcal{R}_n) - 1), \tag{1.9}$$

where $\mathbb{E}(\mathcal{R}_n) = [z^n] \sum_{k=1}^{\infty} \sum_{\sigma\in\mathcal{S}_k} \frac{\mathcal{R}(\sigma)z^k}{k!}$.

Proof. Theorem 1.1.9 yields $R_n \in \{n+1-\mathcal{R}(\sigma), n+2-\mathcal{R}(\sigma), \dots, n\}$. For a random permutation each of the $\mathcal{R}(\sigma)$ ranks are equally likely which yields

$$\mathbb{E}(R_n) = \mathbb{E}\left(\sum_{k=n-\mathcal{R}(\sigma)+1}^{n} \frac{k}{\mathcal{R}(\sigma)}\right) = \mathbb{E}\left(\frac{1}{\mathcal{R}(\sigma)}\left(\frac{n(n+1)}{2} - \frac{(n-\mathcal{R}(\sigma))(n+1-\mathcal{R}(\sigma))}{2}\right)\right)$$
$$= \mathbb{E}\left(n + \frac{1}{2} - \frac{1}{2}\mathcal{R}(\sigma)\right) = n + \frac{1}{2} - \frac{\mathbb{E}(\mathcal{R}_n)}{2}.$$

By definition of g_n we get $\mathbb{E}(g_n) = 1 - \frac{\mathbb{E}(R_n)}{n} = \frac{\mathbb{E}(\mathcal{R}_n)-1}{2n}$.

□

1.2.C Hiring and firing

In our original formulation of the hiring problem an irrevocable decision has to be made whether to hire a candidate right after the interview. As we will see it happens quite often that a candidate is discarded even though a candidate holding a worse relative rank has already got hired. *Hiring and firing* is a relaxation of the hiring problem caring about exactly the mentioned situation. Therefore we add a *replacement mechanism* to all strategies. This mechanism works in a very simple way. Given a sequence $X_1, X_2, \dots X_n$ of candidates and a permutation $\sigma \in \mathcal{S}_n$ of the relative ranks of the candidates, we denote the hiring variables for hiring and firing with $X_i^{[f]}$ whilst X_i denotes the hiring variable for the corresponding strategy without hiring and firing, i.e. $X_i^{[f]} = 1$ if the i-th candidate gets hired with hiring and firing, and $X_i^{[f]} = 0$ otherwise.

Hiring and firing performs the hiring process for a candidate in two steps.

1. *Hiring:* There are two possibilities for the candidate to get hired with hiring and firing.
   (a) $X_i = 1$, i.e. the candidate gets hired by the strategy anyway.
   (b) $X_i = 0 \wedge \exists\, j < i : X_j^{[f]} = 1 \wedge \sigma(j) < \sigma(i)$, i.e. the hiring strategy would not hire candidate i but there exists an already hired candidate whose rank is worse than the rank of the current candidate.
2. *Firing:* In contrast to the constraint for the normal hiring the decision for a hiring is not irrevocable. If a new candidate gets hired although $X_i = 0$, the worst already hired candidate gets fired. Note that the decision not to hire a candidate is irrevocable as well as firing is.

We can see directly that when hiring a candidate by applying (1b), the worst already hired candidate is replaced by the currently interviewed one. This yields $h_n = h_n^{[f]}$, so the number of hired candidates is not affected by the replacement mechanism. The huge benefit of this mechanism lies within the quality of the hired candidates. No matter which pragmatic strategy we use hiring and firing will always hire the best h_n candidates. Obviously the most interesting quantity for hiring and firing is the number of candidates hired by applying (1b), which we call the *number of replacements*.

Definition 1.2.9 (Number of replacements). Let $X_1, X_2, \ldots, X_n$ denote a sequence of candidates for some hiring process. Then the *number of replacements* is given by $f_n := |\{i :\ 1 \le i \le n \ \wedge\ X_i = 0 \ \wedge\ X_i^{[f]} = 1\}|$. This is exactly the number of candidates hired by applying our replacement algorithm.

□

The quantity f_n gives an analytic relationship between the quantities defined in the Sections 1.2.A and 1.2.B. A good hiring strategy already hires only candidates having a good quality. This means that the number of replacements ought to be low. On the other hand this could also mean that the strategy hires too many candidates. We will notice this behavior for *hiring above the minimum* which we consider in Chapter 2.

1.3 Notational conventions and asymptotics

In this section we will introduce some important notations we will need for our purposes and state some useful asymptotic identities.

As we are often interested in what happens when the length n of our sequence tends to infinity we first consider the asymptotic notation. We will mostly follow the notation and definitions in [7, Chapter A.2].

Let S be a set and $s_0 \in S$ a particular element of S. We assume a notion of neighborhood to exist on S. Let f, g denote two functions from $S \setminus \{s_0\}$ to $\mathbb{R}$ or $\mathbb{C}$.

- *$\mathcal{O}$-notation:*

  $$f(s) \underset{s \to s_0}{=} \mathcal{O}(g(s))$$

  if $\frac{f(s)}{g(s)}$ stays bounded as $s \to s_0$ in S. In other words, there exists a neighborhood V of s_0 and a real constant $C > 0$, so that

  $$|f(s)| \le C\ |g(s)|, \quad s \in V \setminus \{s_0\}$$

- *$\sim$-notation:*

  $$f(s) \underset{s \to s_0}{\sim} g(s)$$

  if $\frac{f(s)}{g(s)}$ tends to 1 as $s \to s_0$ in S.

- *o-notation:*

  $$f(s) \underset{s \to s_0}{=} o(g(s))$$

  if $\frac{f(s)}{g(s)}$ tends to 0 as $s \to s_0$ in S. In other words, for any arbitrary small $\varepsilon > 0$ there exists a neighborhood V_ε of s_0, so that

  $$|f(s)| \le \varepsilon\ |g(s)|, \quad s \in V_\varepsilon \setminus \{s_0\}.$$

For convenience we will often leave $s \to s_0$ if it is clear at which point s_0 we consider the asymptotics.

Definition 1.3.1 (Harmonic numbers). For $n, k \in \mathbb{N}$ the n-th *harmonic number* of order k is defined by

$$H_n^{(k)} := \sum_{i=1}^{n} \frac{1}{i^k}.$$

For $k = 1$ we simply write H_n instead of $H_n^{(1)}$.

□

Lemma 1.3.2. Asymptotically, for $n \to \infty$, the harmonic numbers have the following expansion:

$$\begin{aligned} H_n &= \log n + \gamma + \mathcal{O}\left(\frac{1}{n}\right), & \\ H_n^{(k)} &= \mathcal{O}(1), & k > 1, \end{aligned}$$

whereat $\gamma \doteq 0.5772$ denotes the *Euler-Mascheroni-constant.*

□

Another important kind of numbers which will play an important role when considering hiring above the m-th best, are the Stirling numbers of first kind.

Definition 1.3.3 (Stirling numbers). For $n, k \in \mathbb{N}$ the *(unsigned) Stirling numbers of first kind* $\left[{n \atop k}\right]$ are implicitly given by

$$z^{\overline{n}} = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} z^k,$$

where $z^{\overline{n}} = z(z+1)(z+2)\cdots(z+n-1)$ denotes the *rising factorials.*

□

As $z^{\overline{n}}$ is a polynomial of degree n we may interpret the Stirling numbers of first kind as the coefficients of $z^{\overline{n}}$ relating to the canonical basis $\{1, z, \ldots, z^n\}$ of the polynomial ring $\mathbb{R}[z] := \left\{\sum_{i=0}^{n} a_i z^i \mid n \in \mathbb{N} \ \wedge \ a_i \in \mathbb{R}, \ 0 \le i \le n\right\}$.

Theorem 1.3.4 (Stirling's formula). Consider a number $n \in \mathbb{N}$. Then for the factorial $n! := \prod_{i=1}^{n} i$ it holds that

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right). \tag{1.10}$$

□

Stirling's formula gives us some useful asymptotics for the binomial coefficients that we state below.

Corollary 1.3.5. Consider two natural numbers $n, k \in \mathbb{N}$. Then for $1 \leq k \leq n^{\frac{1}{2}+\varepsilon}$ and arbitrary $0 < \varepsilon < \frac{1}{6}$ it holds that

$$c(n,k) := \frac{\binom{n-k}{k}}{\binom{n}{k}} = e^{-\frac{k^2}{n}} \left(1 + \mathcal{O}\left(\frac{k^3}{n^2}\right)\right).$$

Furthermore, the numbers $c(n,k)$ are sub-exponentially small for $k \geq n^{\frac{1}{2}+\varepsilon}$.

Proof. We use the logarithmic version of Stirling's formula, Theorem 1.3.4, which is given by

$$\log(n!) = \left(n + \frac{1}{2}\right)\log(n) - n + \frac{1}{2}\log(2\pi) + \mathcal{O}\left(\frac{1}{n}\right) \tag{1.11}$$

to show the claim. Applying (1.11) to

$$c(n,k) := \frac{\binom{n-k}{k}}{\binom{n}{k}} = \frac{((n-k)!)^2}{n!\ (n-2k)!}$$

gives us

$$\begin{aligned}\log(c(n,k)) = 2\left(\left(n - k + \frac{1}{2}\right)\log(n-k)\right) - \left(n + \frac{1}{2}\right)\log(n) \\ - \left(n - 2k + \frac{1}{2}\right)\log(n-2k) + \mathcal{O}\left(\frac{1}{n}\right).\end{aligned}$$

Applying the decompositions

$$n - k = n\left(1 - \frac{k}{n}\right) \qquad \text{and} \qquad n - 2k = n\left(1 - \frac{2k}{n}\right),$$

together with the fact that $\log(1+x) = x - \frac{x^2}{2} + \mathcal{O}(x^3)$ yield

$$\begin{aligned}\log(c(n,k)) &= 2\left(n - k + \frac{1}{2}\right)\log\left(1 - \frac{k}{n}\right) - \left(n - 2k + \frac{1}{2}\right)\log\left(1 - \frac{2k}{n}\right) + \mathcal{O}\left(\frac{1}{n}\right) \\ &= -\frac{k^2}{n} + \mathcal{O}\left(\frac{k^3}{n^2}\right) + \mathcal{O}\left(\frac{1}{n}\right),\end{aligned}$$

which shows the claim. For $k \geq n^{\frac{1}{2}+\varepsilon}$ this result implies that

$$c(n,k) = \mathcal{O}\left(e^{-n^{2\varepsilon}}\right)$$

which means that the numbers $c(n,k)$ are sub-exponentially small in this case.

□

In the following we define two very famous functions that help us when investigating hiring above the median (see Chapter 4). These are firstly the *Riemann zeta function* and secondly the *gamma function*. Furthermore, we state their asymptotic expansions.

Definition 1.3.6 (Riemann zeta function). The *Riemann zeta function* is defined by

$$\begin{aligned}\zeta : \mathbb{C}_{>1} &\to \mathbb{C} \\ s &\mapsto \sum_{n=1}^{\infty} n^{-s},\end{aligned}$$

where $\mathbb{C}_{>1} := \{s \in \mathbb{C} \mid \Re(s) > 1\}$ denotes the set of all complex numbers s with real part $\Re(s) > 1$.

□

The Riemann zeta function can be continued analytically to $\mathbb{C} \setminus \{1\}$.

Definition 1.3.7 (Gamma function). Consider the set $\mathbb{C}_{>0} := \{z \in \mathbb{C} \mid \Re(z) > 0\}$ of all complex numbers z with positive real part. Then the complex-valued function

$$\begin{aligned} \Gamma : \mathbb{C}_{>0} &\to \mathbb{C} \\ z &\mapsto \int_0^\infty t^{z-1} e^{-t} dt \end{aligned}$$

is called the *gamma function*.

□

By using the identity $\Gamma(s+1) = s\,\Gamma(s)$, we can analytically continue the gamma function to $\mathbb{C} \setminus \{0, -1, -2, \ldots\}$. Furthermore, this provides the useful identity $n! = \Gamma(n+1)$. For detailed computations see Example 1.5.4. There is also an asymptotic expansion of the gamma function that is a generalization of Stirling's formula for the factorials.

Corollary 1.3.8 (Stirling's formula for the gamma function). For $\Re(z) \to \infty$, the gamma function has the following asymptotic expansion:

$$\Gamma(z) = \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z \left(1 + \mathcal{O}\left(\frac{1}{z}\right)\right). \tag{1.12}$$

□

Another important identity for the gamma function is given by the following lemma.

Lemma 1.3.9 (Legendre's duplication formula). For each $z \in \mathbb{C} \setminus \{-\frac{n}{2} \mid n \in \mathbb{N}\}$ it holds that

$$\Gamma(z) \cdot \Gamma\left(z + \frac{1}{2}\right) = \frac{\sqrt{\pi}}{2^{2z-1}} \cdot \Gamma(2z).$$

□

1.4 Formal power series and generating functions

In this section we introduce generating functions, that play a fundamental role for most combinatorial problems. We assume a basic knowledge of the field of analytic combinatorics and most of the stated results can be found in [7, 12].

Let us start with the introduction of formal power series which are closely related to polynomials that we already mentioned in Section 1.3.

Definition 1.4.1 (Formal power series). Let R denote a ring with unity. Then we call the sum

$$\sum_{n=0}^{\infty} a_n z^n = a_0 + a_1 z + a_2 z^2 + \dots$$

a *formal power series* with coefficients in R. We denote the set of all formal power series by $R[[z]]$. Furthermore, we define the sum of two formal power series $\sum_{n=0}^{\infty} a_n z^n$, $\sum_{n=0}^{\infty} b_n z^n$ by

$$\sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} (a_n + b_n) z^n,$$

and their product by

$$\left(\sum_{n=0}^{\infty} a_n z^n\right) \cdot \left(\sum_{n=0}^{\infty} b_n z^n\right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k}\right) z^n.$$

The set $R[[z]]$ together with the two defined operations $+$ and $\cdot$ forms a ring with unity again.

□

The main reason for considering a formal power series is that we want to gain information about its coefficients. The following definition gives us a powerful tool to investigate the coefficients.

Definition 1.4.2 (Coefficient extraction operator). Consider a ring with unity R and a formal power series $\sum_{n=0}^{\infty} a_n z^n$ with coefficients in R. Then we call the operator

$$\begin{aligned} [z^n]: \quad & R[[z]] \to R, \\ & \sum_{n=0}^{\infty} a_n z^n \mapsto a_n, \end{aligned}$$

the *coefficient extraction operator*.

□

Obviously the coefficient extraction operator is a linear operator. Furthermore, for a formal power series $A(z) \in R[[z]]$ it satisfies the following useful identity that holds for all $n \in \mathbb{N}$ and $1 \leq k \leq n$:

$$[z^{n-k}]\ A(z) = [z^n]\ z^k A(z).$$

Definition 1.4.3 (Derivative of a formal power series). Let R denote a ring with unity and $A(z) := \sum_{n=0}^{\infty} a_n z^n \in R[[z]]$ be a formal power series. Then we call the function

$$A'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}$$

the *formal derivative* of $A(z)$.

□

The formal derivative is a useful tool to deal with coefficients $a_n := nb_n$ for an arbitrary sequence of values $b_n \in R$. We only state at this point that it is possible to introduce a topology on formal power series, which is a complete topology. Thus, we may also consider and can often identify formal power series with analytic functions. A detailed description of this topology can be found in [7, 12].

Definition 1.4.4 (Generating function). For a given sequence $a_n \in R$ in a ring R we often call the corresponding formal power series $A(z) = \sum_{n=0}^{\infty} a_n z^n$ the *generating function* of the sequence a_n.

□

Very often the sequence a_n is given by a recursion. Generating functions are a useful tool for solving these recursions and receiving an explicit form for a_n. Thus, by multiplying the equation by z^n and summing over n we can transport the recursion into a functional equation or a (partial) differential equation for the corresponding generating function. When solving this equation we can apply the coefficient extraction operator to determine a_n.

In many cases we do not simply have a sequence with only one index. For these sequences we may generalize the notation of generating functions.

Definition 1.4.5 (Bivariate generating function (BGF)). Let $a_{n,k} \in R$ denote a double-indexed sequence. Then the function

$$A(z, u) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} a_{n,k} z^n u^k$$

is called the *bivariate generating function* of the sequence $a_{n,k}$. We can easily adapt this definition inductively if the sequence consists of more than two indices. The corresponding functions are called *multivariate generating functions* (MGF).

□

For our purposes BGFs play an important role for investigating random variables. In particular consider a sequence of discrete random variables X_n with values in $\mathbb{N}_0$. We get a sequence for the probabilities by $p_{n,k} = \mathbb{P}(X_n = k)$ in a natural way. We can also invert this construction. Starting from a sequence of non-negative values $a_{n,k}$ we get a random variable X_n by

$$\mathbb{P}(X_n = k) = p_{n,k} = \frac{a_{n,k}}{a_n},$$

where we defined $a_n = \sum_{k=0}^{\infty} a_{n,k}$. We can also define generating functions for random variables.

Definition 1.4.6 (Probability generating function (PGF)). Let X denote a discrete random variable with values in $\mathbb{N}_0$. Then we call the function

$$P(u) = \sum_{k=0}^{\infty} \mathbb{P}(X = k) u^k$$

the *probability generating function* of X.

□

The PGF of a random variable X can be used to gain more information about X. For instance, we have this useful identity:

$$\mathbb{E}(X) = \sum_{k=0}^{\infty} k\ \mathbb{P}(X = k) = P'(u)\big|_{u=1}.$$

The following theorem shows the connection between the PGF and the BGF of the sequence $p_{n,k}$ of the probabilities.

Theorem 1.4.7 ([7] Proposition III.1). Let $P(z,u) = \sum_{n,k=0}^{\infty} \mathbb{P}(X_n = k) z^n u^k$ denote the BGF of a sequence of discrete random variables X_n with values in $\mathbb{N}_0$. Then for all $n \in \mathbb{N}$ the PGF of X_n is given by

$$P_n(u) = \sum_{k=0}^{\infty} \mathbb{P}(X_n = k) u^k = \frac{[z^n]P(z,u)}{[z^n]P(z,1)}.$$

□

The following very useful lemma delivers a representation for the factorial moments of a sequence of random variables X_n.

Lemma 1.4.8. Consider a sequence of discrete random variables $(X_n)_{n\in\mathbb{N}}$ with values in $\mathbb{N}_0$ and the corresponding BGF $P(z,u) = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} \mathbb{P}(X_n = k) z^n u^k$. Then the integer factorial moments $\mathbb{E}(X_n^{\underline{r}})$ are, for $r \in \mathbb{N}$, given by

$$\mathbb{E}(X_n^{\underline{r}}) = r!\ [z^n u^r] P(z, u+1).$$

Proof. By Newton's binomial theorem it holds that

$$P(z, u+1) = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \binom{k}{\ell} \mathbb{P}(X_n = k) z^n u^{\ell}.$$

When applying the coefficient extraction operator we get

$$[z^n u^r]P(z, u+1) = [u^r] \sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \mathbb{P}(X_n = k) u^{\ell} = \sum_{k=0}^{\infty} \frac{k(k-1)\cdots(k-r+1)}{r!} \mathbb{P}(X_n = k),$$

and as a direct consequence we get that $\mathbb{E}(X_n^{\underline{r}}) = r![z^n u^r]P(z, u+1)$.

□

1.5 Complex analysis

In this section we will state some results from the field of complex analysis. We assume a basic knowledge of complex function theory, however detailed information about this topic, as well as the proofs of the stated theorems, can be found in [17, 27]. First of all we need some definitions.

Let $D \subseteq \mathbb{C}$ be open and $f : D \setminus \{w\} \to \mathbb{C}$ holomorphic. Then f can be represented by the so called *Laurent series about* w given by $f(z) = \sum_{n=-c}^{\infty} a_n(z-w)^n$, with an arbitrary $c \in \mathbb{N}$. The coefficient $a_{-1} =: Res(f, w)$ is called the *residuum of* f *at* w.

Consider a closed, continuous and piecewise continuous differentiable path $\gamma : [a, b] \to \mathbb{C}$. Then the *cycling number of* γ *about a point* $w \in \mathbb{C} \setminus \gamma([a, b])$ is given by

$$n(\gamma, w) := \frac{1}{2\pi i} \int_{\gamma} \frac{1}{\zeta - w} d\zeta.$$

With these definitions we are able to formulate the following theorem.

Theorem 1.5.1 (Residue theorem). Let $D \subseteq \mathbb{C}$ be open and $w_1, w_2, \ldots, w_n \in D$. Furthermore, consider a holomorphic function $f : D \setminus \{w_1, w_2, \ldots, w_n\} \to \mathbb{C}$ and closed, continuous and piecewise continuous differentiable paths $\gamma_k : [a_k, b_k] \to D \setminus \{w_1, w_2, \ldots, w_n\}$, $k = 1, 2, \ldots, m$ holding $\sum_{k=1}^{m} n(\gamma_k, z) = 0$ for all $z \in \mathbb{C} \setminus D$. Then it holds that

$$\sum_{k=1}^{m} \frac{1}{2\pi i} \int_{\gamma_k} f(\zeta) d\zeta = \sum_{j=1}^{n} \mathrm{Res}(f, w_j) \cdot \sum_{k=1}^{m} n(\gamma_k, w_j).$$

□

Another powerful tool for asymptotic analysis of complex functions is the *Mellin transform*. We will introduce the Mellin transform and state some properties of this type of integral transformation. Detailed explanations and proofs can be found in [6, 7].

Definition 1.5.2 (Mellin transform). For a function $f : [0, \infty) \to \mathbb{C}$ we call the complex function

$$\begin{aligned} f^{\star} : \mathbb{C} &\to \mathbb{C} \\ s &\mapsto \int_0^{\infty} f(x) x^{s-1} dx \end{aligned}$$

the *Mellin transform of* f.

□

The major use of the Mellin transform is for the asymptotic analysis of functions $f(x)$ either at $x = 0$ or for the limit $x \to \infty$. The asymptotics then can be extracted from the singularities of the Mellin transform. We also give some criteria for the existence of the Mellin transform in the following lemma.

Lemma 1.5.3. Let $f(x)$ be locally integrable and assume the following two conditions hold:

$$\begin{aligned} f(x) &\underset{x \to 0^+}{=} \mathcal{O}(x^u), \\ f(x) &\underset{x \to \infty}{=} \mathcal{O}(x^v). \end{aligned}$$

Then the Mellin transform $f^{\star}(s)$ of $f(x)$ exists in a strip $-u < \Re(s) < -v$.

□

Like for most important integral transformations there exists an inversion of the Mellin transform $f^\star$ of $f(x)$ which is given by

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f^\star(s) x^{-s} ds,$$

if f is continuous in an interval containing x. The *abscissa* c should be chosen in the *fundamental strip* of f, i.e. any c satisfying $-u < c < -v$ with u, v like in Lemma 1.5.3.

Example 1.5.4. Let us consider the continuous function $f(x) = e^{-x}$. Then the Mellin transform $f^\star(s)$ is given by

$$f^\star(s) = \Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx,$$

as the resulting integral is exactly the definition of $\Gamma(s)$. The integral converges for $\Re(s) > 0$, but when applying integration by parts one can easily show the representation

$$\Gamma(s+1) = s\ \Gamma(s),$$

or equivalently, $\Gamma(s) = \frac{\Gamma(s+1)}{s}$. Hence, using this identity, we can continue the gamma function analytically to $\mathbb{C} \setminus \{0, -1, -2, \ldots\}$ with poles at all points $z \in \{0, -1, -2, \ldots\}$. Considering the fact that

$$\Gamma(1) = \int_0^\infty e^{-x} dx = 1,$$

together with the functional equation for $\Gamma(s)$ leads us to

$$\mathrm{Res}(\Gamma(s), 0) = \lim_{s \to 0} s\ \Gamma(s) = \Gamma(1) = 1.$$

So we determined the residuum of $\Gamma(s)$ at $s = 0$. The functional equation of $\Gamma(s)$ gives a recursion for the residues at all the other poles $r_n :=\mathrm{Res}(\Gamma(s), -n)$:

$$r_n = \frac{r_{n-1}}{-n}, \quad n \geq 1.$$

One may easily solve this recursion with the starting value $r_0 = 1$ by

$$\mathrm{Res}(\Gamma(s), -n) = r_n = r_0 \cdot \prod_{i=0}^{n-1} \left(-\frac{1}{i+1} \right) = \frac{(-1)^n}{n!},$$

which gives us a representation for all residues of the continued gamma function.

As the original definition of the gamma function converges for $\Re(z) > 0$, this is also the fundamental strip for the existence of the inversion of the Mellin transform of $f(x) = e^{-x}$ which is given by

$$e^{-x} = \int_{c-i\infty}^{c+i\infty} \Gamma(s) x^{-s} ds.$$

■

1.6 Probabilistic methods

We will state some important theorems from the field of probability theory that will help us to determine the probabilities for the hiring parameters and their distributions as well. A basic understanding of measure and probabilistic theory is assumed. We will not give proofs for the statements but they can be found in the mentioned works. First of all we introduce an important notation that we will often use when determining limit distributions. Detailed explanations and properties can be found in [23, 24].

Definition 1.6.1 (Convergence in distribution). Let $(X_n)_{n\in\mathbb{N}}$ denote a sequence of random variables with density functions $f_n(x)$. We say the sequence X_n is *convergent in distribution* to a random variable X with density function $f_X(x)$ if and only if the following condition holds:

$$\lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \int_{-\infty}^{x} f_X(x)dx,$$

or equivalently, if for each point x of continuity of f_X it holds that

$$\lim_{n\to\infty} f_n(x) = f_X(x).$$

We will denote this convergence by

$$X_n \xrightarrow{(d)} X.$$

□

The following theorem has many applications in analytic combinatorics and helps to determine the limit distribution of a sequence on random variables. It is an application of the Central Limit Theorem (CLT).

Theorem 1.6.2 (Quasi-powers Theorem, [7] Theorem IX.8). Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of discrete random variables with values in $\mathbb{N}_0$ with probability generating functions $p_n(u)$. Assume that, uniformly in a fixed complex neighborhood of $u = 1$, for sequences $\beta_n, \kappa_n \to +\infty$, it holds that

$$p_n(u) = A(u) \cdot B(u)^{\beta_n} \left(1 + \mathcal{O}\left(\frac{1}{\kappa_n}\right)\right), \tag{1.13}$$

where both $A(u)$ and $B(u)$ are analytic at $u = 1$ with $A(1) = B(1) = 1$. Assume finally that $B(u)$ satisfies the so-called *variability condition*,

$$B''(1) + B'(1) - B'(1)^2 \neq 0.$$

Then the distribution of X_n is, after standardization, asymptotically Gaussian:

$$\mathbb{P}\left(\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \leq x\right) = \Phi(x) + \mathcal{O}\left(\frac{1}{\kappa_n} + \frac{1}{\sqrt{\beta_n}}\right) \tag{1.14}$$

where $\Phi(x)$ is the distribution function of a standard normal,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{w^2}{2}} dw.$$

□

Another useful theorem for determining limit distributions is given by Curtiss.

Theorem 1.6.3 (Curtiss, [4]). Let $F_n(x)$ and $G_n(x)$ be respectively the probability density function and the moment generating function of a variate X_n. If $G_n(\alpha)$ exists for $|\alpha| < \alpha_1$ and for all $n \geq n_0$, and if there exists a finite-valued function $G(\alpha)$ defined for $|\alpha| \leq \alpha_2 < \alpha_1,\ \alpha_2 > 0$, such that

$$\lim_{n\to\infty} G_n(\alpha) = G(\alpha),\ |\alpha| \leq \alpha_2,$$

then there exists a variate X with probability density function $F(x)$ such that

$$\lim_{n\to\infty} F_n(x) = F(x)$$

at each continuity point and uniformly in each finite or infinite interval of continuity of $F(x)$. The moment generating function of X exists for $|\alpha| \leq \alpha_2$ and is equal to $G(\alpha)$ in that interval.

□

The following theorem is a statement about the existence of the limit distribution of a sequence of random variables, when only the moments of the random variables in the sequence are known. It is a generalization of the second limit theorem and can be found in [8, 24].

Theorem 1.6.4 (Fréchet and Shohat, [8]). Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables that satisfy the following conditions:

i) There is a number $n_0 \in \mathbb{N}$ so that the moments $\mathbb{E}(X_n^r)$ exist for all $r \in \mathbb{N}_0$ and all $n \geq n_0$.

ii) For any $r \in \mathbb{N}_0$ the moments $\mathbb{E}(X_n^r)$ lie - if they exist - between two fixed boundaries independent of n.

Then there exists a subsequence $Y_i = X_{n_i},\ i \in \mathbb{N}$ with $n_i < n_{i+1}$ for all i, such that

a) $\lim_{i\to\infty} \mathbb{E}(Y_i^r) = m_r$ exists for all $r \in \mathbb{N}_0$.

b) The subsequence $(Y_i)_{i\in\mathbb{N}}$ converges in distribution to one fixed random variable X, i.e.

$$Y_i \xrightarrow{(d)} X.$$

c) The moments of X exist and satisfy that

$$\mathbb{E}(X^r) = m_r.$$

□

Chapter 2

Hiring above the minimum

In this section we will discuss a first simple but not trivial strategy. This strategy is called *hiring above the minimum*, where the first candidate gets hired and further applicants only get recruited if their rank is better than the candidate with the worst rank that got hired. Obviously this candidate is always the first candidate. This shows that hiring above the minimum coincides with the strategy *hiring above a fixed threshold* where we choose the first candidate as a threshold candidate and others only get hired if and only if they are better than the threshold.

2.1 Hiring above the minimum

Based on an idea mentioned in [1] we first consider the simple strategy *hiring above the minimum* which is also known as *hiring above a fixed threshold*. In this strategy we hire the first candidate X_1 who becomes the threshold candidate. Further candidates X_i get hired if and only if their score is better than the threshold. This implies that the first candidate stays the worst ranked among all hired candidates at each step, so a hiring happens if and only if the current candidate's rank is better than the minimum rank of all hired candidates so far. Formally the strategy may be described in the following way:

Strategy 2.1.1 (Hiring above the minimum). Let $\sigma \in \mathcal{S}_n$ denote the permutation of the ranks of the interviewed candidates after n interviews and $\mathcal{H}(\sigma)$ denote the set of hired candidates. In *hiring above the minimum* $\mathcal{H}(\sigma)$ is constructed applying these rules:

- $\mathcal{H}_1(\sigma) = \{1\}$.
- $\mathcal{H}_i(\sigma) = \begin{cases} \mathcal{H}_{i-1}(\sigma), & \sigma_i(i) < \sigma_i(1), \\ \mathcal{H}_{i-1}(\sigma) \cup \{i\}, & \sigma_i(i) > \sigma_i(1), \end{cases} \qquad i \geq 2.$

The value $\sigma_i(1) =: \tau_i$ is called the *threshold* at step i. For convenience we will often leave the index of the threshold value if it is clear which step we consider.

■

When considering a sequence of interviewed candidates the permutation gets built up like described in Section 1.1. Thus, the value of τ increases by one every time a candidate is discarded as we can see in the next example.

Example 2.1.2. Let $\sigma = 463921785$ be the permutation of the ranks in a sequence of interviews. We will now observe how the threshold changes when the permutation is built up after each interview.

n	rank	σ_n	hired	threshold
1	1	1	✓	1
2	2	12	✓	1
3	1	231		2
4	4	2314	✓	2
5	1	34251		3
6	1	453621		4
7	6	4537216	✓	4
8	7	45382167	✓	4
9	5	463921785	✓	4

As we can see the value of the threshold changes each time a candidate is discarded. For the hiring set of the permutation σ we get $\mathcal{H}(\sigma) = \{1, 2, 4, 7, 8, 9\}$.

■

As one would expect there is a strong dependency on the value of τ. Since we always consider random permutations of the candidates the first candidate could probably be the best one. In that case all the other candidates are discarded. On the other hand a low rank for τ would mean that almost each candidate gets hired. Like mentioned in [1] it is often useful to consider τ in dependency of n when investigating asymptotic behavior as the threshold will increase with growing n. Let us now investigate some quantities for hiring above the minimum.

Theorem 2.1.3. Consider the indicator variable X_n which evaluates to 1 if the n-th candidate is hired and to 0 otherwise. Then the following equalities hold

$$\mathbb{P}(X_n = 1 \mid \tau = k) = \frac{n-k}{n-1}, \qquad 1 \leq k \leq n \ \wedge \ n > 1, \tag{2.1}$$

$$\mathbb{P}(X_n = 1) = \begin{cases} 1, & n = 1, \\ \frac{1}{2}, & \text{else.} \end{cases} \tag{2.2}$$

Proof. Assuming $\tau := \sigma(1) = k$ we have $n - k$ out of the remaining $n - 1$ ranks getting hired. This leads to

$$\mathbb{P}(X_n = 1 \mid \tau = k) = \frac{n-k}{n-1}.$$

For the second formula let $\mathcal{T} = \{\sigma \in \mathcal{S}_n \mid \tau = k\}$. As each $\sigma \in \mathcal{T}$ has its first position fixed and the other elements may get permuted randomly we get $|\mathcal{T}| = (n-1)!$ out of $|\mathcal{S}_n| = n!$ permutations. This shows $\mathbb{P}(\tau = k) = \frac{(n-1)!}{n!} = \frac{1}{n}$. By definition of our strategy we get

$\mathbb{P}(X_1 = 1) = 1$ because the first candidate gets hired anyway. For $n > 1$ the law of complete probability yields

$$\begin{aligned}\mathbb{P}(X_n = 1) &= \sum_{k=1}^{n} \mathbb{P}(X_n = 1 \mid \tau = k)\ \mathbb{P}(\tau = k) = \sum_{k=1}^{n} \frac{n-k}{n-1}\ \frac{1}{n} \\ &= \frac{1}{n(n-1)} \left(\sum_{k=1}^{n} n - k \right) = \frac{1}{n(n-1)}\ \frac{(n-1)n}{2} = \frac{1}{2}\end{aligned}$$

□

Like mentioned before the restricted probability for hiring a candidate is strongly dependent on the rank of the threshold candidate. For a good strategy we do not want to have such a dependency. We could also observe that the probability for a hiring is $\frac{1}{2}$ for each candidate, i.e. independent of n. This result is very intentional because there are only two possibilities. Either the candidate is better than the threshold or worse. As a consequence we get the following theorem.

Theorem 2.1.4. The exact probability and the expectation for the size of the hiring set h_n and $1 \leq i \leq n$ are

$$\mathbb{P}(h_n = i) = \frac{1}{n}, \tag{2.3}$$

$$\mathbb{E}(h_n) = \frac{n+1}{2}. \tag{2.4}$$

Proof. For $n = 1$ the only candidate gets hired for sure. This yields $\mathbb{P}(h_1 = 1) = 1$. As each candidate gets hired if and only if he is better than the threshold, the size of the hiring set is directly correlating to τ. Thus, $\mathbb{P}(h_n = i) = \mathbb{P}(\tau = n - (i-1)) = \frac{1}{n}$.
For the expectation we get

$$\mathbb{E}(h_n) = \sum_{i=1}^{n} i\ \frac{1}{n} = \frac{n+1}{2}.$$

□

As mentioned in the proof above we can recognize a strong dependency on the threshold τ and the conditional probability for h_n is Diracian. This yields $\mathbb{E}(h_n \mid \tau = k) = n + 1 - k$ for the conditional expectation.

Remark 2.1.5. The asymptotic behavior of h_n itself is not very difficult to observe. For $n \to \infty$, h_n converges to a continuous uniform distribution. If we have a look at the conditioned distribution we can observe a different behavior. Following [1] we consider the asymptotic expansion of h_n respecting the relation between n and the rank τ of the threshold candidate. Concretely we observe the relation $\tau = \alpha n + o(n)$ for some α holding $0 < \alpha \leq 1$. For the density $\frac{\mathbb{E}(h_n)}{n}$ of the derivation of h_n this means

$$\frac{\mathbb{E}(h_n \mid \tau = \alpha n + o(n))}{n} = \frac{n + 1 - \alpha n - o(n)}{n} = 1 + \frac{1}{n} - \alpha + o(1) = 1 - \alpha + o(1).$$

■

Next we want to consider the waiting time W_N until N candidates get hired. The following remark is useful for this parameter and each (pragmatic) strategy.

Remark 2.1.6. For the waiting time W_N it holds that

$$\mathbb{P}(W_N = k) = \mathbb{P}(h_{k-1} = N-1)\ \mathbb{P}(X_k = 1 \mid h_{k-1} = N-1)$$

for each strategy.

■

Proof. The parameter of the waiting time tells us how long we have to wait until N candidates get recruited. More precisely $W_N = k$ means that X_k is the N-th hired candidate, i.e. exactly $N-1$ candidates get hired before candidate X_k. The probability that exactly $N-1$ candidates have got already recruited when X_k is taking his interview is given by $\mathbb{P}(h_{k-1} = N-1)$. In that case the conditioned probability for a hiring of candidate X_k is $\mathbb{P}(X_k = 1 \mid h_{k-1} = N-1)$. □

Using Remark 2.1.6 it is an easy task to compute the probability and expectation for the waiting time with hiring above the minimum.

Theorem 2.1.7. The probability distribution and expectation for the waiting time W_N are

$$\mathbb{P}(W_N = k) = \begin{cases} \delta_{k,1}, & N = 1, \\ 0, & k < N, \\ \frac{N-1}{(k-1)^2}, & \text{else}, \end{cases}$$

$$\mathbb{E}(W_N) = \begin{cases} 1, & N = 1, \\ \infty, & \text{else}. \end{cases}$$

Proof. At each step i the maximum number of hired candidates is i as no future candidates may get hired. This yields $\mathbb{P}(W_N = k) = 0$ for $k < N$. For $k \geq N$ Remark 2.1.6 as well as (2.1) and (2.3) yield

$$\begin{aligned} \mathbb{P}(W_N = k) &= \mathbb{P}(h_{k-1} = N-1)\ \mathbb{P}(X_k = 1 \mid h_{k-1} = N-1) \\ &= \mathbb{P}(h_{k-1} = N-1)\ \mathbb{P}(X_k = 1 \mid \tau = k-N+1) \\ &= \frac{1}{k-1}\ \frac{N-1}{k-1}. \end{aligned}$$

For the expectation we get

$$\begin{aligned} \mathbb{E}(W_1) &= \mathbb{P}(W_1 = 1) = 1 \\ \mathbb{E}(W_N) &= \mathbb{P}(W_N = 1) + \sum_{k=2}^{\infty} k\ \frac{N-1}{(k-1)^2}, \qquad N > 1, \end{aligned}$$

where $\mathbb{P}(W_N = 1) = \delta_{N,1}$. For $N = 1$ this yields $\mathbb{E}(W_1) = 1$ and for $N > 1$ we get

$$\mathbb{E}(W_N) = \sum_{k=2}^{\infty} k\ \frac{N-1}{(k-1)^2} \geq \sum_{k=2}^{\infty} k\ \frac{N-1}{k(k-1)} = (N-1)\left(\sum_{k=2}^{\infty} \frac{1}{(k-1)}\right) = \infty.$$

The last equality holds as the sum is the divergent harmonic sum.

□

When we take the expectation of h_n into account this result is not very surprising as the probability for only one hiring is actually very high.

The rank of the last hired candidate is another good measurement for the quality of a hiring strategy. However, for hiring above the minimum it is not such an important size as each hired candidate gets hired independent of the value of his index. We will see that this property does not hold for the other strategies.

Theorem 2.1.8. For the rank R_n of the last hired candidate in a sequence of length n it holds that

$$\mathbb{P}(R_n = i) = \begin{cases} \frac{1}{n}(H_{n-1} - H_{n-i}), & i < n, \\ \frac{1}{n}H_{n-1} + \frac{1}{n}, & i = n, \end{cases}$$

and

$$\mathbb{E}(R_n) = \frac{3n^2 + 3n - 2}{4n}.$$

Proof. Let us first observe the conditional probability for $R_n = i$ under $\tau = k$.

- *Case 1: $i < k$.* Here we get $\mathbb{P}(R_n = i \mid \tau = k) = 0$, because rank i does not get hired.
- *Case 2: $i = k$.* In this case the last hired candidate is the threshold, i.e. $\tau = n$. This yields $\mathbb{P}(r_n = i \mid \tau = k) = 1$ for $i = n$ and $\mathbb{P}(R_n = i \mid \tau = k) = 0$ otherwise.
- *Case 3: $i > k$.* The last hired candidate is not the threshold. Each of the other $n - k$ hired candidates could be hired as the last hired one, and each having the same probability. Thus, it holds that $\mathbb{P}(R_n = i \mid \tau = k) = \frac{1}{n-k}$.

Applying the law of complete probability we get for $i < n$:

$$\mathbb{P}(R_n = i) = \sum_{k=1}^{i-1} \mathbb{P}(R_n = i \mid \tau = k)\ \mathbb{P}(\tau = k) = \sum_{k=1}^{i-1} \frac{1}{n-k}\ \frac{1}{n} = \frac{1}{n}\ (H_{n-1} - H_{n-i})$$

In the case of $i = n$ we need to add the summand from *Case 2* multiplied by $\frac{1}{n}$.

For the expectation we get

$$\mathbb{E}(R_n) = \left(\sum_{i=1}^{n-1} i \frac{1}{n}(H_{n-1} - H_{n-i}) \right) + n\ \frac{1}{n}(H_{n-1} + 1).$$

Flipping the index $i \mapsto n-i$ then yields

$$\begin{aligned}\mathbb{E}(R_n) &= \frac{1}{n}\left(\sum_{i=1}^{n-1}(n-i)(H_{n-1}-H_i)\right) + H_{n-1} + 1 \\ &= \frac{1}{n}\left(\sum_{i=1}^{n-1} n\ H_{n-1}\right) - \frac{1}{n}\left(\sum_{i=1}^{n-1} n\ H_i\right) - \frac{1}{n}\left(\sum_{i=1}^{n-1} i\ H_{n-1}\right) + \frac{1}{n}\left(\sum_{i=1}^{n-1} i\ H_i\right) + H_{n-1} + 1 \\ &= (n-1)\ H_{n-1} - (n\ H_n - n) - \frac{1}{n}\ \frac{(n-1)n}{2}\ H_{n-1} + \frac{1}{n}\ \left(\frac{n(n-1)}{2}\ H_n - \frac{n(n-1)}{4}\right) \\ &+ H_{n-1} + 1,\end{aligned}$$

wherein we used identities for the harmonic numbers which can be found in [19, p.279] to compute the second and fourth sum. Simple algebraic manipulations and the fact that $H_{n-1} - H_n = -\frac{1}{n}$ then show that

$$\mathbb{E}(R_n) = \frac{n+1}{2}\ (H_{n-1} - H_n) + n + 1 - \frac{n(n-1)}{4n} = \frac{3n^2+3n-2}{4n}.$$

□

The next theorem is a statement about the index L_n of the last hired candidate. If we compare L_n to r_n we are not interested in the rank but the position of the last hiring. Therefore we want to remark the fact that a hiring of candidate i is independent of the candidates $X_1 \dots X_{i-1}$ if we do not care whether they are hired or not (i.e. we only restrict ourselves to $X_i = 1$ but $X_1, \dots X_{i-1}$ could have any of the two options 0 or 1). The law of complete probability then ensures $\mathbb{P}(X_i = 1) = \frac{1}{2}$. As a consequence the claim that candidates $j \in \{i+1, \dots, n\}$ are hired or discarded is independent of the candidates $X_1 \dots X_{i-1}$ as well, which can easily be shown by using the law of complete probability.

Theorem 2.1.9. For the exact probability of the index L_n of the last hired candidate it holds that

$$\mathbb{P}(L_n = i) = \begin{cases} \frac{1}{n}, & i = 1, \\ \frac{1}{(n-i+1)(n-i+2)}, & 1 < i \leq n. \end{cases}$$

The expectation of L_n is given by

$$\mathbb{E}(L_n) = n + 1 - H_n.$$

Proof. For $i = 1$ the threshold candidate is the last hired candidate, i.e. $\mathbb{P}(L_n = 1) = \frac{1}{n}$. In the case of $i > 1$ Theorem 2.1.3 provides $\mathbb{P}(X_i = 1) = \frac{1}{2}$. Furthermore, L_n is independent of the candidates $1 \dots i-1$. Thus, without loss of generality we may just delete the candidates $2, \dots, i-1$ and consider $i = 2$ in a permutation of length $n - i + 2$. For that special case it is true that

$$\begin{aligned}\mathbb{P}(L_n = i) &= \mathbb{P}(L_{n-i+2} = 2)\\ &= \mathbb{P}(X_2 = 1) \cdot \prod_{k=1}^{n-i} \mathbb{P}(X_{k+2} = 0 \mid X_2 = 1, X_3 = X_4 = \ldots = X_{k+1} = 0)\\ &= \frac{1}{2} \cdot \prod_{k=1}^{n-i} \frac{k}{k+2} = \frac{1}{(n-i+1)(n-i+2)}.\end{aligned}$$

For the expectation it holds that

$$\mathbb{E}(L_n) = \sum_{i=1}^{n} i\ \mathbb{P}(L_n = i) = \frac{1}{n} + \underbrace{\sum_{i=2}^{n} \frac{i}{(n-i+1)(n-i+2)}}_{=:A}.$$

Partial fraction decomposition applied on all summands of A yields

$$A = \sum_{i=2}^{n} \left(\frac{i}{n-i+1} - \frac{i}{n-i+2} \right) = n + 1 - \left(H_n + \frac{1}{n} \right)$$

and for the expectation we get $\mathbb{E}(L_n) = n + 1 - H_n$.

□

We can compute the distance between the last two hirings in a similar way. We only have to adapt our considerations to two runs of discarded candidates. The first run is between the second last hired candidate and the last hired one and the other run starts at the last hired candidate and leads to the end of the sequence.

Theorem 2.1.10. Let Δ_n denote the distance between the two last hirings (i.e. the number of discarded candidates between the two last hirings plus one). Then the exact probability for Δ_n is given by

$$\mathbb{P}(\Delta_n = i) = \begin{cases} \frac{1}{n}, & i = 0, \\ \frac{1}{i(i+1)}, & 1 \leq i \leq n-1. \end{cases}$$

For the expectation we get

$$\mathbb{E}(\Delta_n) = H_n - 1.$$

Proof. By definition we get $\mathbb{P}(\Delta_n = 0) = \frac{1}{n}$ as only the first candidate gets hired. For $h_n \geq 2$ we have to distinguish two cases.

- *Case 1: $h_n = 2$.* Assuming $\Delta_n = i$ the only two hired candidates are on position 1 (i.e. the threshold candidate) and on position $i + 1$. This yields

$$\mathbb{P}(\Delta_n = i \mid h_n = 2) = 1 \left(\prod_{k=2}^{i} \frac{k-1}{k} \right) \frac{1}{i+1} \left(\prod_{k=i+2}^{n} \frac{k-2}{k} \right) = \frac{1}{(n-1)n}$$

- *Case 2:* $h_n > 2$. With the same argument as in the proof of Theorem 2.1.9 the first of the last two hired candidates may again be the second candidate in a corresponding permutation of length $n-\ell$ for $\ell = 0, \ldots, n-i-2$ referring to the possible positions of the last hired candidate (i.e. the number of discarded candidates after the last hired one). Thus

$$\begin{aligned}\mathbb{P}(\Delta_n = i \mid L_{n-\ell} = i+1) &= \frac{1}{2}\left(\prod_{k=3}^{i+1}\frac{k-2}{k}\right)\frac{2}{i+2}\left(\prod_{k=i+3}^{n-\ell}\frac{k-3}{k}\right)\\ &= \frac{2}{(n-\ell)(n-\ell-1)(n-\ell-2)}.\end{aligned}$$

For the exact probability of Δ_n we get

$$\mathbb{P}(\Delta_n = i) = \frac{1}{(n-1)n} + \underbrace{\sum_{\ell=0}^{n-i-2}\frac{2}{(n-\ell)(n-\ell-1)(n-\ell-2)}}_{=:A}.$$

Again partial fraction decomposition applied on the summands of A yields

$$\begin{aligned}A &= \sum_{\ell=0}^{n-i-2}\left(\frac{1}{n-\ell} - \frac{2}{n-\ell-1} + \frac{1}{n-\ell-2}\right)\\ &= (H_n - H_{i+1}) - 2(H_{n-1} - H_i) + (H_{n-2} - H_{i-1})\\ &= -\frac{1}{(n-1)n} + \frac{1}{i(i+1)}\end{aligned}$$

and finally $\mathbb{P}(\Delta_n = i) = \frac{1}{i(i+1)}$. For the expectation it holds that

$$\mathbb{E}(\Delta_n) = \sum_{i=0}^{n-1} i\ \mathbb{P}(\Delta_n = i) = \sum_{i=1}^{n-1} i\ \frac{1}{i\ (i+1)} = H_n - 1.$$

□

To complete our considerations about hiring above the minimum we can easily see that the rank of the best discarded candidate is always $\tau - 1$. Furthermore, it is not sensible to consider hiring and firing for hiring above the minimum as exactly the best $n - \tau + 1$ candidates are hired. As we could see hiring above the minimum is no ideal hiring strategy due to the strong dependency on the rank of the first candidate and the infinity of the expectation of the waiting time W_N. A variant for hiring above a minimum could be that the threshold is not the first candidate but the median of all employees already working at the company.

Example 2.1.11 (Example 2.1.2, continued). Last but not least we compute all the investigated quantities for the permutation $\sigma = 463921785$. As we already mentioned before the hiring set is given by $\mathcal{H}(\sigma) = \{1, 2, 4, 7, 8, 9\}$. We get the following results

name	variable	value
Size of the hiring set	$h(\sigma)$	6
Waiting time for $N = 3$	$W_3(\sigma)$	4
Rank of the last hired candidate	$r(\sigma)$	5
Index of the last hired candidate	$L(\sigma)$	9
Distance between the last two hirings	$\Delta(\sigma)$	1

We also state the expected values for $n = 9$ to compare the general results to that specific case when considering σ.

name	variable	expectation
Size of the hiring set	h_9	5
Waiting time for $N = 3$	W_3	∞
Rank of the last hired candidate	r_9	$5.2\underline{5}$
Index of the last hired candidate	L_9	$7.1\underline{7}$
Distance between the last two hirings	Δ_9	$1.8\underline{3}$

Chapter 3

Hiring for the elite

"Welcome to Lake Wobegon, where all the women are strong,
all the men are good-looking and all the children are above average."
- Garrison Keillor

In this chapter we will consider two very selective strategies. Firstly hiring above the maximum which only hires candidates better than all already interviewed persons. This strategy hires the elite only but there may be many excellent candidates that are discarded on the other hand. Secondly we observe hiring above the m-th best. It is a generalization of hiring above the maximum which we get for $m = 1$. Nevertheless it is sensible to consider hiring above the maximum separately as it provides an interesting connection to *records* in permutations.

Together with hiring above the median, which we will consider in Chapter 4, we obtain a certain kind of strategies, the so-called *Lake Wobegon Strategies*. These strategies were first introduced in [3] who named them after the fictional town Lake Wobegon, appearing in the works of Garrison Keillor. Lake Wobegon Strategies only hire applicants that are better than the average of the already recruited staff. This implies that the quality of the hiring set increases every time a new candidate gets hired. This is where the name 'Lake Wobegon Strategies' comes from.

Another example for a score-based Lake Wobegon Strategy is *hiring above the mean* (see [3]). Following [25] this strategy is not only of theoretical interest as 'Google Search $^{\text{TM}}$' actually works with hiring above the mean.

As mentioned before, we are considering three strategies of this kind. Firstly hiring above the maximum which only hires candidates better than all already interviewed persons. This strategy hires the elite only but there may be many excellent candidates who are discarded on the other hand. Secondly we observe hiring above the m-th best. It is a generalization of hiring above the maximum which we get for $m = 1$. Nevertheless it is sensible to consider hiring above the maximum separately as it provides an interesting connection to *records* in permutations. The last considered strategy of this type is hiring above the median. It hires candidates that are better than the median of the already hired candidates only. According to [3] it is notable that hiring above the median delivers considerably different results than hiring above the mean does.

For all these strategies a new candidate gets recruited if and only if his relative rank is better than the rank of some other already hired threshold candidate similarly to hiring above the minimum in Chapter 2. The difference to hiring above the minimum is that the threshold is not the same candidate at each step. So the threshold candidate may change when a new candidate is hired. Note that according to Definition 1.1.7 the threshold candidate must not change if an applicant is rejected otherwise the strategy was not pragmatic.

3.1 Hiring above the maximum

The first Lake Wobegon Strategy we consider is *hiring above the maximum*. As the name says, hiring above the maximum only recruits a candidate if and only if he is the very best candidate seen so far, i.e. if a new applicant is better than the currently best candidate. Technically spoken we would not need a threshold candidate concerning this strategy as a hiring implies $\sigma_i(i) = i$ but obviously the two formulations are equivalent. In contrast to e.g. hiring above the minimum a benefit of this strategy is that the probability for a hiring is independent of the past candidates and the permutation of their relative ranks because obviously $\mathcal{R}(\sigma) = 1$ holds for each permutation σ of arbitrary length $|\sigma| \in \mathbb{N}$. As all our results are only the special case $m = 1$ in hiring above the m-th best (see Section 3.2) we will only consider $h_{n,1}$ the number of hired candidates, $W_{N,1}$ the waiting time as well as $\Delta_{n,1}$ the distance between the last two hirings. Due to consistency with Section 3.2 we denote the parameters by adding a second index m which will always be 1 in this section.

Strategy 3.1.1 (Hiring above the maximum). Let σ denote the permutation of ranks of the interviewed candidates after n interviews and $\mathcal{H}(\sigma)$ denote the set of hired candidates. In hiring above the maximum $\mathcal{H}(\sigma)$ is constructed applying these rules:

- $\mathcal{H}_1(\sigma) = \{1\}$.
- $\mathcal{H}_i(\sigma) = \begin{cases} \mathcal{H}_{i-1}(\sigma) \cup \{i\}, & \sigma_i(i) > \sigma_i(j) \quad \forall\, j < i, \\ \mathcal{H}_{i-1}, & \text{else,} \end{cases} \qquad \text{for } i \geq 2.$

Note that instead of $\sigma_i(i) > \sigma_i(j)\ \forall j < i$ we could also say $\sigma_i(i) = i$.

■

Before we start investigating the hiring parameters we first introduce the notation of *records* which are characteristic for permutations. We will use them to determine the parameters.

Definition 3.1.2 (Record, left-to-right-maximum). Let $\sigma \in \mathcal{S}_n$ denote a permutation of length n. An index $i \leq n$ is called *record* (or *left-to-right-maximum*) of σ if and only if

$$\sigma(i) > \sigma(j) \quad \forall\, j < i. \tag{3.1}$$

This means that the value of $\sigma(i)$ is the relative maximum of the first i indices. Note that 1 is always a record. So again the first candidate gets hired anyway and becomes the first threshold.

□

Remark 3.1.3. As one can easily see (3.1) equals exactly the constraint for a hiring in our strategy. This means each hired candidate corresponds to a record in the permutation of ranks and vice versa. As a direct consequence we get that the size of the hiring set after n interviews $h_{n,1}$ is given the number of records in the permutation of ranks.

■

Let $p_{n,k}$ denote the number of permutations $\sigma \in \mathcal{S}_n$ consisting of exactly k records. Obviously it holds that $\mathbb{P}(h_n = k) = \frac{p_{n,k}}{n!}$. We will now follow [18] to determine the value of $p_{n,k}$. A recursion for $p_{n,k}$ is given by

$$p_{n,k} = p_{n-1,k-1} + (n-1)\; p_{n-1,k}. \tag{3.2}$$

This recursion follows by considering $\sigma \circ j$ for a random permutation $\sigma \in \mathcal{S}_{n-1}$ and $1 \leq j \leq n$. We get two cases for the value of j:

- *Case 1: $j = n$.* Adding n to the permutation leads to an additional record. This yields the first summand.
- *Case 2: $j < n$.* In this case we get no additional record and only the length of the permutation increases by 1. There are $n-1$ possible values j in this case, which leads to the second term.

For fixed $n \in \mathbb{N}$ it is true that $p_{n,1} = (n-1)!$ for the initial values as it $\sigma(1) = n$ must hold for each $\sigma \in \mathcal{S}_n$ consisting of exactly 1 record. Additionally we define $p_{n,0} = \delta_{n,0}$ for technical issues. Solving this recursion leads us to the next theorem.

Theorem 3.1.4. For the exact distribution of the number h_n of hired candidates it holds that

$$\mathbb{P}(h_n = k) = \frac{\left[{n \atop k}\right]}{n!}. \tag{3.3}$$

For the expectation we get that

$$\mathbb{E}(h_n) = H_n \sim \log n + \gamma + \mathcal{O}\left(\frac{1}{n}\right). \tag{3.4}$$

Proof. Let $P_n(z) := \sum_{k=0}^{\infty} p_{n,k} z^k$ denote the vertical generating function of $p_{n,k}$. Multiplying Equation (3.2) by z^k and summing over all $k \geq 1$ then yields

$$\sum_{k=1}^{\infty} p_{n,k}\; z^k = z \sum_{k=0}^{\infty} p_{n-1,k}\; z^k + (n-1) \sum_{k=1}^{\infty} p_{n-1,k}\; z^k$$
$$P_n(z) = z\; P_{n-1}(z) + (n-1)\; P_{n-1}(z) = (z+n-1)\; P_{n-1}(z), \quad n \geq 1,$$

as we defined $p_{n,0} = \delta_{n,0}$. It is a routine task to solve this homogenous functional equation for $P_n(z)$ which yields

$$P_n(z) = \prod_{j=1}^{n-1} (z+n-j)\; \underbrace{P_1(z)}_{\equiv\, z} = (z)^{\overline{n}} = \sum_{k=0}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} z^k.$$

The last equation holds by Definition 1.3.3 and finally $\mathbb{P}(h_n = k) = \frac{p_{n,k}}{n!}$ shows (3.3).
By construction

$$H_n(u) := \frac{P_n(u)}{n!} = \frac{u(u+1)(u+2)\cdots(u+n-1)}{n!} = \binom{u+n-1}{n} \tag{3.5}$$

is the PGF for our parameter h_n with fixed $n \in \mathbb{N}$. We determine the expectation of the derivation following [7]. Therefore it is useful to consider the logarithmic derivation of the PGF.

$$\log(H_n(u))' = \frac{H_n(u)'}{H_n(u)} = \frac{P_n(u)'}{P_n(u)} \tag{3.6}$$

Simple manipulations and plugging

$$\log H_n(u) = \sum_{k=0}^{n-1} [\log(u+k)] - \log(n!)$$
$$\Rightarrow \log(H_n(u))' = \sum_{k=0}^{n-1} \left[\frac{1}{u+k}\right]$$

into (3.6) then yields

$$H_n(u)' = \log(H_n(u))' \cdot H_n(u) = \binom{u+n-1}{n} \cdot \left(\sum_{k=0}^{n-1} \frac{1}{u+k}\right). \tag{3.7}$$

Finally evaluation of (3.7) at $u = 1$ leads to

$$\mathbb{E}(h_n) = H_n(u)'\big|_{u=1} = \binom{n}{n} \cdot \left(\sum_{k=0}^{n-1} \frac{1}{k+1}\right) = H_n. \tag{3.8}$$

□

In the next step we determine the distribution of h_n for $n \to \infty$. Here it is useful to consider the BGF $H(z,u)$ of our parameter h_n starting from the PGF for fixed n given by $H_n(u)$. By Equation (3.5) it holds that

$$H(z,u) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{{n \brack k}}{n!} z^n u^k = \sum_{n=0}^{\infty} H_n(u) z^n = \sum_{n=0}^{\infty} \binom{u+n-1}{n} z^n$$

Applying the well known identity $(-1)^n \binom{-u}{n} = \binom{u+n-1}{n}$ for any $u \in \mathbb{C}$ and using Newton's generalized binomial theorem then yields

$$H(z,u) = \sum_{n=0}^{\infty} \binom{u+n-1}{n} z^n = \sum_{n=0}^{\infty} (-1)^n \binom{-u}{n} z^n = (1-z)^{-u}. \tag{3.9}$$

Last but not least we compute the horizontal OGF $H^{(k)}(z) := [u^k]H(z,u)$ for fixed $k \in \mathbb{N}$. According to [5] it is helpful to consider the extension

$$(1-z)^{-u} = e^{(-u)\log(1-z)} = e^{u \log\left(\frac{1}{1-z}\right)}.$$

When applying the power series expansion of e^z we get

$$e^{u\log(\frac{1}{1-z})} = \sum_{k=0}^{\infty} \frac{u^k \log\left(\frac{1}{1-z}\right)^k}{k!}$$

and finally equating the coefficients yields

$$H^{(k)}(z) = [u^k]H(z,u) = [u^k]e^{u\log(\frac{1}{1-z})} = \frac{\log\left(\frac{1}{1-z}\right)^k}{k!}. \tag{3.10}$$

Theorem 3.1.5 (Goncharov's Theorem). The distribution of h_n is asymptotically normal. In other words it holds that

$$\lim_{n\to\infty} \mathbb{P}(h_n \leq \log n + \gamma + x\sqrt{\log n}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-w^2}{2}} dw. \tag{3.11}$$

Proof. By Formula (3.9) it holds that $H(z,u) = (1-z)^{-u}$ for the BGF of the probabilities $\mathbb{P}(h_n = k)$. For fixed $n \in \mathbb{N}$ we get, for u in a small area around 1 (e.g. $|u-1| \leq \frac{1}{2}$),

$$H_n(u) = [z^n]H(z,u) = \binom{n+u-1}{n} = \frac{u(u+1)\cdots(u+n-1)}{n!} = \frac{\Gamma(u+n)}{\Gamma(u)\Gamma(n+1)}.$$

The equality $H_n(u) = \sum_{k=0}^{n} \mathbb{P}(h_n = k)u^k$ implies that $H_n(1) = 1$ which means that $H_n(u)$ is the PGF of the random variable h_n. Following [7], Stirling's formula for the gamma function, Theorem 1.3.8, then yields

$$H_n(u) = \frac{\Gamma(u+n)}{\Gamma(u)\Gamma(n)(n+1)} = \frac{\sqrt{2\pi}(n+u)^{n+u-\frac{1}{2}}e^{-(n+u)}}{\Gamma(u)\left(\sqrt{2\pi}n^{n-\frac{1}{2}}e^{-n}\right)(n+1)} \cdot \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Factoring out n in the numerator and the last factor of the denominator followed by pruning out factors then yields

$$\begin{aligned} H_n(u) &= \frac{n^{n+u-\frac{1}{2}}\left(1+\frac{u}{n}\right)^{n+u-\frac{1}{2}} e^{-u}}{\Gamma(u)\left(n^{n+\frac{1}{2}}\right)\left(1+\frac{1}{n}\right)} \cdot \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) \\ &= \frac{n^{u-1}e^{-u}e^{u+\mathcal{O}(\frac{1}{n})}}{\Gamma(u)\left(1+\frac{1}{n}\right)} \cdot \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) = \frac{n^{u-1}}{\Gamma(u)} \cdot \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right). \end{aligned}$$

The claim follows by applying the Quasi-powers Theorem 1.6.2 having $A(u) = \frac{1}{\Gamma(u)}$, $B(u) = e^{u-1}$, $\beta_n = \log n$ and $\kappa_n = n$ with analytic $A(u), B(u)$ having $A(1) = B(1) = 1$. Furthermore, the function $B(u)$ satisfies the variability condition.

□

As mentioned before we will now investigate the waiting time $W_{N,1}$ and the distance between the last two hirings $\Delta_{n,1}$ not only because these two parameters give a deeper insight how the strategy will perform for growing n but also as the results give a special case for $m = 1$.

The next theorem shows that the expectation for the waiting time with hiring above the maximum does not exist in contrast to $m > 1$ as we will see in Section 3.2.

Theorem 3.1.6. For the exact probabilities and the expectation of the waiting time $W_{N,1}$ until N candidates get hired with hiring above the maximum it holds that

$$\mathbb{P}(W_N = k) = \frac{\left[{k-1 \atop N-1}\right]}{k!}$$
$$\mathbb{E}(W_N) = \begin{cases} 1, & N = 1 \\ \infty & \text{else.} \end{cases}$$

Proof. In analogy to the proof of Theorem 2.1.7 it holds for the probability of $W_{N,1}$ that

$$\mathbb{P}(W_{N,1} = k) = \frac{1}{k}\ \mathbb{P}(h_{k-1,1} = N-1) = \frac{1}{k}\ \frac{\left[{k-1 \atop N-1}\right]}{(k-1)!} = \frac{\left[{k-1 \atop N-1}\right]}{k!}.$$

For the expectation it obviously holds that $\mathbb{E}(W_{1,1}) = 1$ as the first candidate always gets recruited. In the case of $N > 1$ we consider the BGF $H(z,u)$ of $h_{n,1}$ given by Equation (3.9) as we have

$$\mathbb{E}(W_{N,1}) = \sum_{k=1}^{\infty} k\ \mathbb{P}(W_{N,1} = k) = \sum_{k=1}^{\infty} \frac{\left[{k-1 \atop N-1}\right]}{(k-1)!} = [u^{N-1}]H(z,u)|_{z=1},$$

where the summand for $k = 1$ is 0 as $N > 1$. The last equation results from the BGF $H(z,u) = \sum\limits_{i,j=1}^{\infty} \frac{\left[{i \atop j}\right]}{i!} z^i u^j$. Furthermore, (3.10) yields

$$[u^{N-1}]\ H(z,u) = \sum_{i=1}^{\infty} \frac{\left[{i \atop N-1}\right]}{i!} z^i = \frac{1}{(N-1)!}\ \log\left(\frac{1}{1-z}\right)^{N-1}.$$

For $z = 1$ we have a pole of $[u^{N-1}]\ H(z,u)$, i.e. $\lim_{z\to 1} \frac{1}{(N-1)!}\ \log\left(\frac{1}{1-z}\right)^{N-1} = \infty$ which yields $\mathbb{E}(W_{N,1}) = \infty$ for $N > 1$.

□

Secondly we consider the distance between the last two hirings that comes up with different results than for $m > 1$.

Theorem 3.1.7 (Helmi, Martínez and Panholzer [14, Theorem 6 (ii)]). The exact probability and expectation of the distance $\Delta_{n,1}$ between the last two hirings are

$$\mathbb{P}(\Delta_{n,1} = d) = \begin{cases} \frac{1}{n}, & d = 0, \\ \frac{1}{n}(H_{n-1} - H_{d-1}), & 1 \le d \le n-1, \end{cases}$$
$$\mathbb{E}(\Delta_{n,1}) = \frac{n^2 + n - 2}{4n}.$$

Note that we defined $\Delta_{n,1} = 0$ for $h_{n,1} = 1$.

Proof. For the case $d = 0$, which is exactly the case that only one candidate gets hired, we get $\mathbb{P}(\Delta_{n,1} = 0) = \mathbb{P}(h_{n,1} = 1) = \frac{1}{n}$.
Let us now consider $d \geq 1$. Similarly to the proof of Theorem 2.1.10 the probability $\mathbb{P}(\Delta_{n,1} = d)$ is given by a hiring followed by $(d-1)$ discarded applicants, a hiring again and no further hirings until the end. The position of the first hiring depends on the number of discarded candidates at the end of the sequence. Obviously that number must not be larger than $n-d-1$ as there need to be $d-1$ discarded candidates between the last two hirings. Thus, we have

$$\mathbb{P}(\Delta_{n,1} = d) = \sum_{k=d+1}^{n} \frac{1}{k-d} \left(\prod_{l=k-d}^{k-1} \frac{l-1}{l} \right) \frac{1}{k} \left(\prod_{l=k+1}^{n} \frac{l-1}{l} \right)$$
$$= \sum_{k=d+1}^{n} \frac{1}{k-1} \frac{1}{n} = \frac{1}{n} \left(H_{n-1} - H_{d-1}\right).$$

The expectation is given by

$$\mathbb{E}(\Delta_{n,1}) = \sum_{d=0}^{\infty} d \; \mathbb{P}(\Delta_{n,1} = d) = \sum_{d=1}^{n-1} d \; \frac{1}{n}(H_{n-1} - H_{d-1}).$$

Splitting the sum and simple algebraic manipulations then yield

$$\mathbb{E}(\Delta_{n,1}) = \frac{1}{n} \frac{(n-1)n}{2} H_{n-1} - \frac{1}{n} \sum_{d=1}^{n-1} (d-1) \; H_{d-1} - \frac{1}{n} \sum_{d=1}^{n-1} H_{d-1}$$
$$= \frac{n-1}{2} H_{n-1} - \frac{(n-1)(n-2)}{2n} H_{n-1} + \frac{(n-1)(n-2)}{4} - \frac{n-1}{n} H_{n-1} + \frac{n-1}{n}.$$

The last equality follows by identities for harmonic summation which can be found in [19, p.279]. Finally simplifying the terms leads to

$$\mathbb{E}(\Delta_{n,1}) = \frac{n^2 + n - 2}{4n}.$$

□

Remark 3.1.8. As hiring above the m-th best always hires the first m candidates, the synthetic case when merely one hiring happens only appears with hiring above the maximum, i.e. $m = 1$. This implies that $W_{N,m}$ and $\Delta_{n,m}$ have to be investigated separately for $m = 1$ and $m > 1$. Another consequence of that special case is that - similarly to hiring above the minimum - $\mathbb{E}(W_{N,1}) = \infty$, for $N > 1$.

■

Example 3.1.9. Let us now investigate how the permutation $\sigma = 463921785$ that we also investigated with hiring above the minimum performs with hiring above the maximum.

n	rank	σ_n	hired
1	1	1	✓
2	2	12	✓
3	1	231	
4	4	2314	✓
5	1	34251	
6	1	453621	
7	6	4537216	
8	7	45382167	
9	5	463921785	

variable	value
$h_{9,1}(\sigma)$	3
$W_{1,1}(\sigma)$	1
$W_{2,1}(\sigma)$	2
$W_{3,1}(\sigma)$	4
$r_{9,1}(\sigma)$	9
$L_{9,1}(\sigma)$	4
$\Delta_{9,1}(\sigma)$	2
$M_{9,1}(\sigma)$	8
$f_{9,1}(\sigma)$	2

The left table shows that the number of hired candidates is far less than with hiring above the minimum. The hiring set is given by $\mathcal{H}(\sigma) = \{1, 2, 4\}$. This means that the second and third best candidate are not hired, so this strategy is very selective. In the right table we see hiring and firing replaces 2 out of 3 candidates and the waiting time increases vastly. This is caused by the early position of the best candidate.

■

3.2 Hiring above the m-th best

Hiring above the m-th best is a generalization of hiring above the maximum which we considered in Section 3.1. The results of this section are mainly based on [1,14]. Similarly to hiring above the maximum we hire the first m candidates and further candidates get recruited if and only if their rank is better than the m-th best already hired candidate. One huge contrast to hiring above the maximum is, that the case of only one hiring does not appear for $m \geq 2$. As a consequence we will achieve that for instance the expectation of the waiting time $W_{N,m}$ is finite.

Strategy 3.2.1 (Hiring above the m-th best). Let σ be the permutation of the ranks of the interviewed candidates after n interviews and $\mathcal{H}(\sigma)$ denote the set of hired candidates. In hiring above the m-th best $\mathcal{H}(\sigma)$ is constructed applying these rules:

- $\mathcal{H}_i(\sigma) = \{1, 2, \ldots, i\}, \qquad 1 \leq i \leq m.$
- $\mathcal{H}_i(\sigma) = \begin{cases} \mathcal{H}_{i-1}(\sigma) \cup \{i\}, & \sigma_i(i) \geq i - (m-1), \\ \mathcal{H}_{i-1}, & \text{else}, \end{cases} \qquad \text{for } i > m.$

This means that candidate X_i only gets hired if and only if less than m already hired candidates hold a better rank than X_i, which is fulfilled trivially if $i \leq m$.

■

The huge disadvantage when analyzing hiring above the m-th best is that we must not simply consider records in permutations. Thus, we need a generalization of records. We consider two main types of k-records in permutations. Most of the considerations can be found in [2].

Definition 3.2.2 (k-record). We call an index $1 \leq i \leq n$ a *Type1 k-record* of a permutation $\sigma \in \mathcal{S}_n$ if and only if $\sigma_i(i) = i - (k-1)$, hence $\sigma(i)$ is the k-th largest element beneath $\{\sigma(1), \ldots, \sigma(i)\}$. An index $1 \leq i \leq n$ is a *Type2 k-record* if and only if there exists $j \geq i$ satisfying $\sigma_j(i) = j - (k-1)$. Analogous this means that $\sigma(i)$ is the k-th largest value in $\{\sigma(1), \ldots, \sigma(j)\}$. For a permutation of length n we denote the number of Type1 k-records by $R^{[1]}_{n,k}$ and the number of Type2 k-records by $R^{[2]}_{n,k}$.

□

Obviously it holds for the number of recruited candidates with hiring above the m-th best that

$$h_{n,m} = R^{[2]}_{n,m} + m - 1. \tag{3.12}$$

Theorem 3.2.3 (Helmi, Martínez and Panholzer, [14, Theorem 3]). For the exact distribution of the number $h_{n,m}$ of hired candidates with hiring above the m-th best it holds that

$$\mathbb{P}(h_{n,m} = k) = \begin{cases} \delta_{n,k}, & m > n, \\ \frac{m!\ m^{k-m}}{n!} \left[{n-m+1 \atop k-m+1}\right], & m \leq k \leq n. \end{cases} \tag{3.13}$$

Proof. Obviously it holds for $n < m$ that $h_{n,m} = n$ which yields $\mathbb{P}(h_{n,m} = k) = \delta_{n,k}$ in that case. For $1 \leq m \leq n$ we will slightly adapt the combinatorial approach of [14, p.16f] to obtain a generalization of Theorem 3.1.4 which handles the case of $m = 1$. Therefore let $\mathcal{S}^{[m]}_{n,k} \subseteq \mathcal{S}_n$ denote the subset of permutations having length n that hire exactly k candidates with hiring above the m-th best. Then it obviously holds that $\mathbb{P}(h_{n,m} = k) = \frac{|\mathcal{S}^{[m]}_{n,k}|}{n!}$. Moreover, Theorem 3.1.4 implies that $|\mathcal{S}^{[1]}_{n,k}| = \left[{n \atop k}\right]$.

Consider a sequence $X_1, \ldots, X_{n-m+1}$ of candidates and a permutation $\sigma \in \mathcal{S}_{n-m+1}$ of their ranks. Furthermore, assume σ consists of exactly $k-m+1 \geq 1$ records, i.e. $\sigma \in \mathcal{S}^{[1]}_{n-m+1,k-m+1}$. By definition of $\circ$ we can write σ in a unique way by $\sigma = s_1 \circ s_2 \circ \cdots \circ s_{n-m+1}$ where $s_i = \sigma_i(i)$ denotes the relative rank of candidate X_i in step i. Hiring above the maximum hires exactly $k-m+1$ out of these $n-m+1$ candidates, namely exactly those having a relative rank satisfying $s_i = i$. By definition of the strategy candidate X_1 is always hired and the choice of the permutation σ determines which further candidates get hired too.

Starting from σ we now construct a permutation σ' that hires exactly the same candidates with hiring above the m-th best plus some further $m-1$ candidates whom we add in the beginning of the sequence. Therefore consider new candidates $Y_1, Y_2 \ldots Y_{m-1}$ together with candidate $Y_m := X_1$. Let $\tau \in \mathcal{S}_m$ denote the permutation of the ranks of these candidates. Again we may write the permutation $\tau = t_1 \circ t_2 \circ \cdots \circ t_m$ as a composition of the relative ranks, i.e. $t_j = \tau_j(j)$ for $1 \leq j \leq m$. Define $\sigma' \in \mathcal{S}_n$ by

$$\sigma' = t_1 \circ t_2 \circ \cdots \circ t_m \circ s'_2 \circ s'_3 \circ \cdots \circ s'_{n-m+1},$$

where it holds that

$$s'_i = \begin{cases} \ell, & s_i = i, \\ s_i & \text{else}, \end{cases}$$

with an arbitrary $\ell \in \{i, i+1, \ldots, i+m-1\}$. As candidate X_i is the $(i+m-1)$-th candidate in the sequence $Y_1, Y_2, \ldots, Y_m, X_2, X_3, \ldots, X_{n-m+1}$ the choice of s'_i ensures that additionally to $Y_1, Y_2 \ldots, Y_m$ the permutation σ' hires exactly the same $k-m$ (here X_1 is not counted as we hire him together with the candidates Y_i) candidates with hiring above the m-th best as σ did when using hiring above the maximum. For each hired candidate X_i we have m options to choose the relative rank s'_i. Furthermore, the choice of τ does not affect the hiring set as well as any possible choice for s'_i does not. This yields that we get $m! \cdot m^{k-m}$ different possibilities for σ' in order to hire the same candidates as σ.

We can easily invert this construction. Starting from a permutation $\sigma' \in \mathcal{S}_n$ given in the form $\sigma' = t_1 \circ t_2 \circ \cdots \circ t_{m-1} \circ s'_1 \circ s'_2 \circ \cdots \circ s'_{n-m+1}$ we simply delete the first $m-1$ candidates and set

$$s_i = \begin{cases} i, & s'_i \in \{i, i+1, \ldots, i+m-1\}, \\ s'_i, & \text{else.} \end{cases}$$

The composition $s_1 \circ \cdots \circ s_{m-n+1}$ finally represents σ again. As mentioned above we have $\left[{n-m+1 \atop k-m+1}\right]$ permutations σ of length $n-m+1$ with exactly $k-m+1$ hirings when using hiring above the maximum which then shows that

$$|\mathcal{S}^{[m]}_{n,k}| = m!\; m^{k-m}\; |\mathcal{S}^{[1]}_{n-m+1,k-m+1}| = m!\; m^{k-m} \left[{n-m+1 \atop k-m+1}\right]$$

and finally implies that

$$\mathbb{P}(h_{n,m} = k) = \frac{m!\; m^{k-m} \left[{n-m+1 \atop k-m+1}\right]}{n!}.$$

□

In our next step we compute the expectation and variance for $h_{n,m}$ too which was first published by Archibald and Martínez in [1]. We will give a different proof for these quantities for which the following simple fact about the hiring probability is quite useful. Obviously in a sequence of n candidates it holds that

$$\mathbb{P}(X_k = 1) = \begin{cases} 1, & 1 \le k \le m, \\ \frac{m}{k}, & \text{else.} \end{cases} \tag{3.14}$$

Corollary 3.2.4 (Archibald and Martínez, [1]). For the expectation and the variance of $h_{n,m}$ and $n \ge m$ it holds that

$$\mathbb{E}(h_{n,m}) = m(H_n - H_m + 1) = m(\log n - \log m + 1) + \mathcal{O}(1)$$
$$\mathbb{V}(h_{n,m}) = m(H_n - H_m) - m^2(H_n^{(2)} - H_m^{(2)}) = m\left(\log n - \log m - 1 + \frac{m}{n}\right) + \mathcal{O}(1).$$

Proof. We consider the fact that

$$h_{n,m} = X_1 + X_2 + \ldots + X_n.$$

Hence, together with (3.14) the following holds:

$$\mathbb{E}(h_{n,m}) = \sum_{k=1}^{n} \mathbb{E}(X_k) = \sum_{k=1}^{n} \mathbb{P}(X_k = 1) = \sum_{k=1}^{m} 1 + \sum_{k=m+1}^{n} \frac{k}{n} = m(H_n - H_m + 1).$$

For the variance we again consider the decomposition of $h_{n,m}$ into the single candidates which leads to

$$\mathbb{V}(h_{n,m}) = \mathbb{E}(h_{n,m}^2) - \mathbb{E}(h_{n,m})^2 = \sum_{i,j=1}^{n} \mathbb{P}(X_i X_j = 1) - \mathbb{E}(h_{n,m})^2.$$

For the product of the indicator variables it holds that

$$\mathbb{P}(X_i^2 = 1) = \begin{cases} 1, & i \le m, \\ \frac{m}{i}, & m < i \le n, \end{cases}$$

as well as for $i \neq j$

$$\mathbb{P}(X_i X_j = 1) = \begin{cases} 1, & 1 \le i, j \le m, \\ \frac{m}{i}, & m < i \le n \ \wedge \ 1 \le j \le m, \\ \frac{m}{j}, & 1 \le i \le m \ \wedge \ m < j \le n, \\ \frac{m^2}{i\,j}, & m < i, j \le n. \end{cases}$$

Thus, the squared random variable satisfies the following equation:

$$\begin{aligned} \mathbb{E}(h_{n,m}^2) &= \sum_{i=1}^{n} \mathbb{P}(X_i^2 = 1) + 2 \sum_{1 \le i < j \le n} \mathbb{P}(X_i X_j = 1) \\ &= \sum_{i=1}^{m} 1 + \sum_{i=m+1}^{n} \frac{m}{i} + 2 \left[\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} 1 + \sum_{i=1}^{m} \sum_{j=m+1}^{n} \frac{m}{j} + \sum_{i=m+1}^{n-1} \sum_{j=i+1}^{n} \frac{m^2}{i\,j} \right] \\ &= m + m(H_n - H_m) + 2 \left[\frac{(m-1)m}{2} + m^2(H_n - H_m) + m^2 \sum_{i=m+1}^{n-1} \frac{H_n - H_i}{i} \right] \\ &= m^2 + (m + 2m^2)(H_n - H_m) + 2m^2 H_n (H_{n-1} - H_m) - 2m^2 \sum_{i=m+1}^{n-1} \frac{H_i}{i}. \end{aligned}$$

For the last sum we make use of the following identity which can be found in [5, p.280]:

$$\sum_{i=1}^{n} \frac{H_i}{i} = \frac{1}{2} \left(H_n^2 + H_n^{(2)} \right) \tag{3.15}$$

Formula 3.15 together with $\mathbb{E}(h_{n,m})^2 = m^2(H_n - H_m)^2 + 2m^2(H_n - H_m) + m^2$ implies that

$$\begin{aligned} \mathbb{V}(h_{n,m}) = m^2 + (m + 2m^2)(H_n - H_m) + 2m^2 \left(H_{n-1}^2 + \frac{1}{n} H_{n-1} - H_n H_m \right) - \\ - m^2 \left(H_{n-1}^2 + H_{n-1}^{(2)} - H_m^2 - H_m^{(2)} \right) - m^2(H_n - H_m)^2 - 2m^2(H_n - H_m) - m^2, \end{aligned}$$

which we can simplify to

$$\mathbb{V}(h_{n,m}) = m(H_n - H_m) + m^2 H_{n-1}^2 + \frac{2m^2}{n} H_{n-1} - m^2 \left(H_{n-1}^{(2)} - H_m^{(2)}\right) - m^2 H_n^2.$$

Finally the fact that $m^2 H_n^2 = m^2 \left(H_{n-1}^2 + \frac{2}{n} H_{n-1} + \frac{1}{n^2}\right)$ leads us to

$$\mathbb{V}(h_{n,m}) = m(H_n - H_m) - m^2 \left(H_n^{(2)} - H_m^{(2)}\right).$$

□

For determining the asymptotic distribution of the standardized random variable $h_{n,m}^\star = \frac{h_{n,m} - \mathbb{E}(h_{n,m})}{\mathbb{V}(h_{n,m})}$ we first compute the PGF of $h_{n,m}$ which we do in a similar way as we did for $m = 1$ in Section 3.1. Analogous to Formula (3.2) defining $\mathbb{P}(h_{n,m} = k) = h_{n,k}^{[m]}$ and considering (3.14) leads to the recursion

$$h_{n,k}^{[m]} = \frac{m}{n}\, h_{n-1,k-1}^{[m]} + \frac{n-m}{n}\, h_{n-1,k}^{[m]}, \qquad n > m \;\wedge\; k \le n. \tag{3.16}$$

When solving this recursion by following [14] we obtain

$$h_{n,m}(u) = \sum_{k=0}^{\infty} h_{n,k}^{[m]} u^k = u^m \prod_{j=m+1}^{n} \frac{mu + (j-m)}{j} = u^m \, (mu)^{\overline{n-m}} \, \frac{m!}{n!}, \tag{3.17}$$

with respect to $h_{n,k}^{[m]} = \delta_{n,k}$ for $n \le m$. This can easily be shown in an analogous way as we did in the proof of Theorem 3.1.4. Also following [14] we can determine the BGF for $h_{n,k}^{[m]}$ by $h_m(z,u) = \sum_{n=m}^{\infty} \binom{n}{m} h_{n,m}(u) z^n$. Thus, we get the formula

$$h_m(z,u) = \sum_{n=m}^{\infty} \binom{n}{m} u^m \, (mu)^{\overline{n-m}} \, \frac{m!}{n!} \, z^n = \sum_{n=m}^{\infty} (zu)^m \, \frac{(mu)^{\overline{n-m}}}{(n-m)!} \, z^{n-m}.$$

Shifting the index and applying Newton's generalized binomial theorem finally yields the explicit form given by

$$h_m(z,u) = (uz)^m \sum_{n=0}^{\infty} \frac{(mu)^{\overline{n}}}{(n)!} \, z^n = (uz)^m \sum_{n=0}^{\infty} \binom{mu+n}{n} z^n = \frac{(zu)^m}{(1-z)^{mu+1}}.$$

Although m is fixed for any hiring process we determine the asymptotic behavior of the distribution depending on the relation between the parameter m which is called *rigidity* in [14], whom we will mostly follow in our proof, and the number n of candidates. This gives further insight about the characteristics of the hiring process when n might depend on m.

Corollary 3.2.5 (Helmi, Martínez and Panholzer, [14, Theorem 3]). Asymptotically, for $n \to \infty$, the limiting distribution of the number of hired candidates $h_{n,m}$ is, with respect to the relation between n and m, given by the following three cases:

(*i*) $n - m \gg \sqrt{n}$.
The standardized random variable $h^{\star}_{n,m}$ converges in distribution to a standard normal distribution, i.e.

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{h_{n,m} - m(\log n - \log m + 1)}{\sqrt{m\left(\log n - \log m - 1 + \frac{m}{n}\right)}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-w^2}{2}} dw.$$

Note: This is the most important case as it contains all fixed values for m. Furthermore, it is a generalization of Theorem 3.1.5.

(*ii*) $n - m \sim \alpha\sqrt{n} = \mathcal{O}(\sqrt{n})$, with $\alpha > 0$.
$n - h_{n,m}$ is asymptotically Poisson distributed with parameter $\frac{\alpha^2}{2}$, i.e.

$$n - h_{n,m} \xrightarrow{(d)} \text{Poisson}\left(\frac{\alpha^2}{2}\right).$$

(*iii*) $n - m = o(\sqrt{n})$.
$n - h_{n,m}$ converges in distribution to 0, i.e. $n - h_{n,m} \xrightarrow{(d)} 0$.

Proof.

(*i*) $n - m \gg \sqrt{n}$.
For fixed m the result can be obtained in a similar way like we did for $m = 1$ in Theorem 3.1.5. Another aproach for m fixed can be found in [1]. For $m \to \infty$ we follow [14] and consider the standardized random variable

$$h^{\star}_{n,m} = \frac{h_{n,m} - m(\log n - \log m + 1)}{\sqrt{m\left(\log n - \log m - 1 + \frac{m}{n}\right)}}.$$

The moment generating function is then given by

$$\mathbb{E}\left(e^{h^{\star}_{n,m}s}\right) = e^{-\frac{\mu}{\sigma}s} \cdot h_{n,m}\left(e^{\frac{s}{\sigma}}\right),$$

where $\mu := m(\log n - \log m + 1)$ and $\sigma^2 := m(\log n - \log m - 1 + \frac{m}{n})$. Applying Stirling's formula to the probability generating function, given in (3.17), yields

$$\begin{aligned}\log(h_{n,m}(u)) = {} & m(u-1)(\log n - \log m) \\ & + (n + m(u-1))\log\left(1 + \frac{m(u-1)}{n}\right) + \mathcal{O}(1-u) + \mathcal{O}\left(\frac{1}{m}\right).\end{aligned}$$

When plugging the moment generating function in this asymptotic formula we get

$$\log\left(\mathbb{E}\left(e^{h^{\star}_{n,m}s}\right)\right) = \frac{s^2}{2} + \mathcal{O}\left(\frac{m\left(1-\frac{m}{n}\right)^2}{\sigma^3}\right) + \mathcal{O}(\sigma^{-1}) + \mathcal{O}(m^{-1}),$$

which implies that $\mathbb{E}(h^{\star}_{n,m}) \longrightarrow e^{\frac{s^2}{2}}$, pointwise for each $s \in \mathbb{R}$, provided that $n - m \gg \sqrt{n}$. Since $e^{\frac{s^2}{2}}$ is the moment generating function of a standard normal distribution, the stated central limit theorem follows with Curtiss' theorem 1.6.3.

(ii) $n - m = \mathcal{O}(\sqrt{n})$.
We consider the random variable $\overline{h}_{n,m} := n - h_{n,m}$, which represents the number of discarded applicants. By Theorem 3.2.3 the explicit form of $h_{n,m}$ with $m \leq k \leq n$ is given by

$$\mathbb{P}(h_{n,m} = k) = \frac{m!m^{k-m}}{n!} \begin{bmatrix} n-m+1 \\ k-m+1 \end{bmatrix}.$$

When defining $\ell = n - m$ and $i = n - k$ we obtain

$$\mathbb{P}(n - h_{n,m} = n - k) = \mathbb{P}\left(\overline{h}_{n,m} = i\right) = \frac{(n-\ell)!(n-\ell)^{\ell-i}}{n!} \begin{bmatrix} \ell+1 \\ \ell-i+1 \end{bmatrix}.$$

Applying the asymptotic identity

$$\begin{bmatrix} \ell+1 \\ \ell-i+1 \end{bmatrix} = \frac{\ell^{2i}}{i!2^i} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

which can be obtained by extracting coefficients from an asymptotic expansion of the BGF of the Stirling numbers, together with

$$\frac{(n-\ell)!(n-\ell)^{\ell-i}}{n!} \sim n^{-i}\, e^{-\frac{\ell^2}{2n}},$$

for fixed i, which can be shown in a similar way as Corollary 1.3.5, finally yields the asymptotic expansion

$$\mathbb{P}\left(\overline{h}_{n,m} = i\right) \sim \frac{1}{i!} \left(\frac{\ell^2}{2n}\right)^i e^{-\frac{\ell^2}{2n}}$$

which holds for $\ell = \mathcal{O}(\sqrt{n})$. This is exactly the probability density function of a Poisson distribution having $\lambda = \frac{\ell^2}{2n} = \frac{(n-m)^2}{2n}$.

(iii) $n - m = o(\sqrt{n})$.
As m is very large in the case of $n - m = o(\sqrt{n})$ nearly every candidate gets hired. Again we consider the number of discarded candidates given by $\overline{h}_{n,m} := n - h_{n,m}$. One may easily see that the moment generating function of $\overline{h}_{n,m}$, which is, for $n \to \infty$, asymptotically given by

$$\mathbb{E}\left(e^{(n-h_{n,m})s}\right) \sim \frac{(n-m)^2}{2n}(e^s - 1)$$

converges to 0. In other words we get $\mathbb{P}(h_{n,m} = n) = 1 - o(1)$.

□

Let us now consider the waiting time which is an essential tool for the choice of m. Imagine a company that wants to recruit a certain number of applicants. The waiting time tells them how many applicants they have to interview until they reach their favored number. When adjusting the parameter m they can administer this number. This shows that the free choice of the parameter m is one of the benefits of this strategy.

Theorem 3.2.6. For the exact probability of the waiting time $W_{N,m}$ until N candidates are recruited when applying hiring above the m-th best it holds that

$$\mathbb{P}(W_{N,m} = k) = \begin{cases} \delta_{N,k}, & N \leq m, \\ \frac{m!\ m^{N-m}}{k!} \left[{k-m \atop N-m}\right], & m < N \leq k, \\ 0, & \text{else.} \end{cases}$$

For $m \geq 2$ the expectation of $W_{N,m}$ is given by

$$\mathbb{E}(W_{N,m}) = \begin{cases} N, & N \leq m, \\ m \ \left(\frac{m}{m-1}\right)^{N-m}, & \text{else.} \end{cases}$$

Note: We already discussed the expectation for $m = 1$ in Theorem 3.1.6.

Proof. We again follow [14] to show the statements. Obviously the case of $N \leq m$ holds because the first m candidates get hired for sure and as the number of hired candidates cannot outrun the number of interviewed applicants the case of $N > k$ holds as well. For the case of $m < N \leq k$ we use the fact stated in Remark 2.1.6 which shows that

$$\mathbb{P}(W_{N,m} = k) = \frac{m}{k}\ \mathbb{P}(h_{k-1,m} = N-1) = \frac{m}{k}\ \frac{m!\ m^{N-1-m}}{k-1!} \begin{bmatrix} k-m \\ N-m \end{bmatrix}.$$

For $2 \leq m < N$ the expectation is then given by

$$\mathbb{E}(W_{N,m}) = \sum_{k=N}^{\infty} k\ \frac{m!\ m^{N-m}}{k!} \begin{bmatrix} k-m \\ N-m \end{bmatrix} = m!\ m^{N-m} \sum_{k=N}^{\infty} \frac{1}{(k-1)!} \begin{bmatrix} k-m \\ N-m \end{bmatrix}.$$

Shifting the index and applying the identity

$$\sum_{i=k}^{\infty} \frac{\left[{i \atop k}\right]}{i!\binom{s+i}{i}} = \frac{1}{s^k}$$

which can be found in [21] and does not converge for $s = 0$ and $k > 0$ finally leads to

$$\begin{aligned} \mathbb{E}(W_{N,m}) &= m!\ m^{N-m} \sum_{k=N-m}^{\infty} \frac{1}{(k+m-1)!} \begin{bmatrix} k \\ N-m \end{bmatrix} \\ &= m^{N-m+1} \sum_{k=N-m}^{\infty} \frac{\left[{k \atop N-m}\right]}{k!\binom{k+m-1}{k}} = m\ \left(\frac{m}{m-1}\right)^{N-m}. \end{aligned}$$

Obviously $\mathbb{E}(W_{N,m}) = N$ holds for $N \leq m$ by definition of the strategy.

□

Remark 3.2.7. In contrast to hiring above the minimum and hiring above the maximum the expectation for the waiting time exists when setting $m \geq 2$. The reason for this fact is that the *bad* case where only one candidate gets hired does not appear for this setting. On the opposite the probability for only recruiting one out of n candidates with hiring above the minimum/maximum is given by $\frac{1}{n}$, which is very high.

This shows that it is more profitable to set $m \geq 2$ for a not too selective but sensible hiring process.

∎

Remark 3.2.8. As the probabilities for $W_{N,m}$ are strongly dependent on the probabilities of the number of recruited candidates $h_{n,m}$ we get similar results for the limiting distribution as well.

∎

Theorem 3.2.9. The exact probability of the index $L_{n,m}$ of the last recruited candidate in a sequence of n applicants with hiring above the m-th best is given by

$$\mathbb{P}(L_{n,m} = i) = \begin{cases} \delta_{i,n}, & i < m, \\ \frac{\binom{i-1}{m-1}}{\binom{n}{m}}, & m \leq i \leq n. \end{cases}$$

For the expectation and variance of $L_{n,m}$ we get

$$\mathbb{E}(L_{n,m}) = \begin{cases} n, & n < m, \\ \frac{m(n+1)}{m+1}, & \text{else}, \end{cases}$$

$$\mathbb{V}(L_{n,m}) = \begin{cases} 0, & n < m, \\ \frac{mn^2-(m^2-m)n-m^2}{(m+2)(m+1)^2}, & \text{else}. \end{cases}$$

Proof. Obviously we get $\mathbb{P}(L_{n,m} = i) = \delta_{i,n}$ for $i < m$ as the first m candidates always get hired. So the i-th candidate may only be the last hired candidate if and only if he is the last applicant, i.e. $i = n$. Let us now consider the non-trivial case of $m \leq i \leq n$. To get the probability $\mathbb{P}(L_{n,m} = i)$ we only need to consider the sequence where the i-th candidate gets hired and all further candidates are discarded. The fact that $\mathbb{P}(X_j = 1) = \frac{m}{j}$ and $\mathbb{P}(X_j = 0) = 1 - \frac{m}{j}$ for $m \leq j$ together with the independence of the random variables X_j leads to

$$\begin{aligned}\mathbb{P}(L_{n,m} = i) &= \mathbb{P}(X_i = 1) \prod_{j=i+1}^{n} \mathbb{P}(X_j = 0) = \frac{m}{i} \prod_{j=i+1}^{n} \left(1 - \frac{m}{j}\right) \\ &= \frac{m(i-1)!(n-m)!}{n!(i-m)!} = \frac{m(i-1)!(n-m)!}{n!(i-m)!} \cdot \frac{(m-1)!}{(m-1)!} = \frac{\binom{i-1}{m-1}}{\binom{n}{m}}.\end{aligned}$$

For the expectation and the variance the probability $\mathbb{P}(L_{n,m} = i) = \delta_{i,n}$ causes $\mathbb{E}(L_{n,m}) = n$ and $\mathbb{V}(L_{n,m}) = 0$ for $n < m$. When considering $n \geq m$ we get the following for the expectation:

$$\mathbb{E}(L_{n,m}) = \sum_{i=m}^{n} i \frac{\binom{i-1}{m-1}}{\binom{n}{m}} = \frac{1}{\binom{n}{m}} \sum_{i=m}^{n} \frac{i!}{(m-1)!(i-m)!} = \frac{m}{\binom{n}{m}} \sum_{i=m}^{n} \binom{i}{m}.$$

Using the well-known identity

$$\sum_{i=0}^{n} \binom{i}{m} = \binom{n+1}{m+1} \tag{3.18}$$

which can also be found in [19, p.160] then leads to

$$\mathbb{E}(L_{n,m}) = \frac{m \binom{n+1}{m+1}}{\binom{n}{m}} = \frac{m(n+1)}{m+1}.$$

For the variance it holds

$$\mathbb{V}(L_{n,m}) = \mathbb{E}(L_{n,m}^2) - \mathbb{E}(L_{n,m})^2 = \sum_{i=m}^{n} i^2 \frac{\binom{i-1}{m-1}}{\binom{n}{m}} - \left(\frac{m(n+1)}{m+1}\right)^2,$$

whereat we first need to compute $\mathbb{E}(L_{n,m}^2)$ which is given by

$$\begin{aligned}\sum_{i=m}^{n} i^2 \frac{\binom{i-1}{m-1}}{\binom{n}{m}} &= \frac{1}{\binom{n}{m}} \sum_{i=m}^{n} \frac{i^2\ (i-1)!}{(m-1)!(i-m)!} = \frac{1}{\binom{n}{m}} \sum_{i=m}^{n} \left(\frac{(i+1)!}{(m-1)!(i-m)!} - \frac{i!}{(m-1)!(i-m)!}\right) \\ &= \frac{1}{\binom{n}{m}} \sum_{i=m}^{n} \left(\frac{(i+1)!}{(m-1)!(i-m)!}\right) - \frac{m(n+1)}{m+1}.\end{aligned}$$

The last equation holds because the sum over the second term in brackets is exactly the same we considered for the expectation of $L_{n,m}$. For the remaining sum expanding factors and again using the identity (3.18) shows that

$$\frac{1}{\binom{n}{m}} \sum_{i=m}^{n} \frac{(i+1)!}{(m-1)!(i-m)!} = \frac{m(m+1)}{\binom{n}{m}} \sum_{i=m}^{n} \binom{i+1}{m+1} = \frac{m(m+1)\binom{n+2}{m+2}}{\binom{n}{m}} = \frac{m(n+2)(n+1)}{m+2}.$$

Putting all the parts together finally yields that

$$\mathbb{V}(L_{n,m}) = \frac{m(n+2)(n+1)}{m+2} - \frac{m(n+1)}{m+1} - \frac{m^2(n+1)^2}{(m+1)^2} = \frac{mn^2 - (m^2-m)n - m^2}{(m+2)(m+1)^2}.$$

□

In the following we will also determine the limiting distribution of $L_{n,m}$. According to [14], whom we will again follow in our considerations, this can in contrast to the number of recruited candidates $h_{n,m}$ easily be done by applying Stirling's formula (1.10). Like we did for $h_{n,m}$ we will again distinguish between different dependencies of m on the number of interviewed persons n. Concretely we observe the following regions which seem to be suitable dependencies. We will consider these regions for all further parameters in this section.

(i) m fixed.
This is the standard case. In a common hiring process m is fixed in the beginning.

(ii) $m \to \infty,\ m = o(n)$.
In this case we consider the main region, where m grows by time but it is rather small compared to n.

(iii) $m \sim \alpha n$, where it holds that $0 < \alpha < 1$.
Here we have an asymptotically fixed dependency of m when n grows to infinity.

(iv) $n - m = o(n)$.
This is a rather trivial case as m grows very fast, which implies that nearly each applicant is recruited in the hiring process. We do not await surprising results for this case.

When computing the limit distributions with respect to the relation between n and m we will, for convenience, consider $n - L_{n,m}$ instead of $L_{n,m}$ in some cases, which does not make a huge difference and gives us the distance of the last hired candidate to the end of the sequence. Furthermore, we only consider the non-trivial cases.

Corollary 3.2.10. The limiting distribution of $L_{n,m}$ for $n \to \infty$ is depending on the relation between m and n given by

(i) m fixed: Suitably normalized, $L_{n,m}$ is asymptotically Beta distributed with parameters m and 1, i.e.

$$\frac{L_{n,m}}{n} \xrightarrow{(d)} \mathrm{Beta}(m, 1).$$

(ii) $m \to \infty$, $m = o(n)$: Suitably normalized, $n - L_{n,m}$ is asymptotically exponential distributed with parameter 1, i.e.

$$\frac{m}{n}(n - L_{n,m}) \xrightarrow{(d)} \mathrm{Exp}(1).$$

(iii) $m \sim \alpha n$, where it holds that $0 < \alpha < 1$: $n - L_{n,m}$ is asymptotically geometrically distributed with success probability α, i.e.

$$n - L_{n,m} \xrightarrow{(d)} \mathrm{Geom}(\alpha).$$

(iv) $n - m = o(n)$: $n - L_{n,m}$ converges in distribution to 0, i.e.

$$n - L_{n,m} \xrightarrow{(d)} 0.$$

Proof.

(i) m fixed. The fact that $n^{\underline{m}} \sim n^m$ directly yields

$$\begin{aligned}\mathbb{P}(L_{n,m} = i) &= \frac{m\ i!(n-m)!}{i(i-m)!\ n!} = \frac{m}{i} \cdot \frac{i^{\underline{m}}}{n^{\underline{m}}} \\ &\sim \frac{m}{i} \cdot \frac{i^m}{n^m} = \frac{m}{n} \cdot \left(\frac{i}{n}\right)^{m-1}.\end{aligned}$$

This implies that $\frac{L_{n,m}}{n} \xrightarrow{(d)} L$, where L has the density function $f(x) = m\ x^{m-1}$, having $0 < x < 1$ as the range of i lies within $0 \le i \le n$. This is exactly the density function of a Beta distribution having parameters $\alpha = m$ and $\beta = 1$.

(ii) $m \to \infty$ but $m = o(n)$. Let $i = n - k$, where $k = o(n)$, then it holds that

$$\mathbb{P}(L_{n,m} = n - k) = \mathbb{P}(n - L_{n,m} = k) = \frac{m}{n}\, e^{-\frac{km}{n}} \left(1 + \mathcal{O}\left(\frac{k^2 m}{n}\right) + \mathcal{O}\left(\frac{km^2}{n}\right)\right),$$

which can be shown similarly to Corollary 1.3.5. Thus, we get $\frac{m}{n}(n - L_{m,n}) \xrightarrow{(d)} L$, where L has density function $f(x) = e^{-x}$ which is the density function of an exponential distribution having $\lambda = 1$.

(iii) $m \sim \alpha n$ with $0 < \alpha < 1$. Considering the fact that

$$\frac{n!}{(n-k)!} = n^k \cdot \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right), \tag{3.19}$$

for each $k \in \mathbb{N}$ fixed, leads to

$$\mathbb{P}(L_{n,m} = n - k) = \frac{m(n-m)!(n-k-1)!}{n!(n-k-m)!} = \frac{m}{n-k} \cdot \frac{(n-m)^k \cdot \left(1 + \mathcal{O}\left(\frac{1}{n-m}\right)\right)}{n^k \cdot \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right)}. \tag{3.20}$$

The specification $m \sim \alpha n$ of this case now shows, for each $k \in \mathbb{N}$ fixed, that

$$\mathbb{P}(L_{n,m} = n - k) = \frac{m}{n} \cdot \left(1 - \frac{m}{n}\right)^k \cdot \left(1 + \mathcal{O}\left(\frac{1}{n-m}\right)\right) = \alpha\ (1-\alpha)^k \cdot \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Hence, $n - L_{n,m}$ is asymptotically a geometric distribution with success probability $p = \alpha$.

(iv) $n - m = o(n)$. This case is rather trivial as the strategy hires almost all applicants. Consider the probability that the n-th candidate gets hired shows that

$$\mathbb{P}(L_{n,m} = n) = \frac{m}{n} = \frac{n - o(n)}{n} = 1 + o(1),$$

which directly yields $n - L_{n,m} \xrightarrow{(d)} 0$.

□

The next parameter we consider is the distance between the last two hirings $\Delta_{n,m}$. We already considered the special case of $m = 1$ which is the only one where the pathological case of $\Delta_{n,m} = 0$ appears.

Theorem 3.2.11. Consider the distance $\Delta_{n,m}$ between the last two hirings with hiring above the m-th best in a sequence of n candidates. Assuming $m \geq 2$, for the probability of $\Delta_{n,m}$ it holds that

$$\mathbb{P}(\Delta_{n,m} = d) = \begin{cases} \delta_{d,1}, & m > n, \\ \frac{1}{m-1}\left(\frac{m^2}{n} - \frac{1}{\binom{n}{m}}\right), & m \leq n \text{ and } d = 1, \\ \frac{m}{\binom{n}{m}} \sum\limits_{j=m+d}^{n} \frac{1}{j-m}\binom{j-d-1}{m-1}, & 2 \leq d \leq n - m. \end{cases}$$

For the expectation we get

$$\mathbb{E}(\Delta_{n,m}) = \frac{m(n+1)}{(m+1)^2} - \frac{1}{(m+1)\binom{n}{m}}\left(m^2+2m+1-\frac{2}{m-1}\right).$$

Proof. We here only consider $m \geq 2$, for $m = 1$ see Theorem 3.1.7. Let us first investigate the case $d \geq 2$ where we need to consider the last two hirings and $(d-1)$ discarded candidates between them. Denote j as the index of the last hired candidate. Then index $i = j - d$ of the second last candidate must satisfy $i \geq m$ as the first m candidates get hired and $d \geq 2$. Thus, we get the probability

$$\begin{aligned}\mathbb{P}(\Delta_{n,m} = d) &= \sum_{j=m+d}^{n} \mathbb{P}(X_{j-d} = 1) \cdot \prod_{\ell=j-d}^{j-1} [\mathbb{P}(X_\ell = 0)] \cdot \mathbb{P}(X_j = 1) \cdot \prod_{\ell=j+1}^{n} [\mathbb{P}(X_\ell = 0)] \\ &= \sum_{j=m+d}^{n} \frac{m}{j-d} \cdot \prod_{\ell=j-d}^{j-1}\left(1-\frac{m}{\ell}\right) \cdot \frac{m}{j} \cdot \prod_{\ell=j-d}^{j-1}\left(1-\frac{m}{\ell}\right).\end{aligned}$$

Pruning out factors finally shows the claim.

For $d = 1$ we additionally have to consider that the last two hirings could possibly be the $(m-1)$-th respectively the m-th candidate, which is exactly given by $\mathbb{P}(L_{n,m} = m)$. Hence, the probability for $\Delta_{n,m} = 1$ is given by

$$\begin{aligned}\mathbb{P}(\Delta_{n,m} = 1) &= \frac{m}{\binom{n}{m}} \sum_{j=m+1}^{n} \frac{1}{j-m}\binom{j-2}{m-1} + \mathbb{P}(L_{n,m} = m) \\ &= \frac{m}{\binom{n}{m}(m-1)} \sum_{j=m+1}^{n} \binom{j-2}{m-2} + \frac{1}{\binom{n}{m}}.\end{aligned}$$

Now using Formula (3.18) leads to

$$\mathbb{P}(\Delta_{n,m} = 1) = \frac{m}{\binom{n}{m}(m-1)}\left(\binom{n-1}{m-1} - 1\right) + \frac{1}{\binom{n}{m}} = \frac{1}{m-1}\left(\frac{m^2}{n} - \frac{1}{\binom{n}{m}}\right).$$

In the case of $n < m$ we trivially get $\Delta_{n,m} = 1$, as the last two candidates get hired.

The expectation can be computed straightforwardly and fortunately switching the order of summation leads to a form that can be simplified:

$$\mathbb{E}(\Delta_{n,m}) = \frac{1}{m-1}\left(\frac{m^2}{n} - \frac{1}{\binom{n}{m}}\right) + \frac{m}{\binom{n}{m}} \cdot \sum_{d=2}^{n-m} d \sum_{j=m+d}^{n} \frac{1}{j-m}\binom{j-d-1}{m-1} \tag{3.21}$$

$$= \frac{1}{m-1}\left(\frac{m^2}{n} - \frac{1}{\binom{n}{m}}\right) + \frac{m}{\binom{n}{m}} \cdot \sum_{j=m+2}^{n} \frac{1}{j-m} \underbrace{\sum_{d=2}^{j-m} d \binom{j-d-1}{m-1}}_{=:S} \tag{3.22}$$

When considering S, one can easily see that for $d > j - m$ all summands would evaluate to zero. Thus, we can extend the sum which yields

$$S = \sum_{d=1}^{j} \binom{d}{1}\binom{j-d-1}{m-1} - \binom{j-2}{m-1} = \binom{j}{m+1} - \binom{j-2}{m-1},$$

whereat the last equation is an application of a variant of the Chu-Vandermonde identity, which can be found in [19, p.169]. Plugging this result into the sum in (3.22) leads to

$$\begin{aligned}\sum_{j=m+2}^{n}\frac{1}{j-m}\sum_{d=2}^{j-m}d\binom{j-d-1}{m-1} &= \sum_{j=m+2}^{n}\frac{1}{j-m}\left[\binom{j}{m+1}-\binom{j-2}{m-1}\right]\\ &= \sum_{j=m+2}^{n}\left[\frac{1}{m+1}\binom{j}{m}-\frac{1}{m-1}\binom{j-2}{m-2}\right]\\ &= \frac{1}{m+1}\left[\binom{n+1}{m+1}-\binom{m}{m}-\binom{m+1}{m}\right]-\frac{1}{m-1}\binom{n-1}{m-1}\\ &= \frac{1}{m+1}\binom{n+1}{m+1}-\frac{m+2}{m+1}-\frac{1}{m-1}\binom{n-1}{m-1}.\end{aligned}$$

Finally, plugging this result into (3.22) shows that

$$\begin{aligned}\mathbb{E}(\Delta_{n,m}) &= \frac{1}{m-1}\left(\frac{m^2}{n}-\frac{1}{\binom{n}{m}}\right)+\frac{m}{\binom{n}{m}}\left(\frac{1}{m+1}\binom{n+1}{m+1}-\frac{m+2}{m+1}-\frac{1}{m-1}\binom{n-1}{m-1}\right)\\ &= \frac{1}{m-1}\left(\frac{m^2}{n}-\frac{1}{\binom{n}{m}}\right)+\frac{m(n+1)}{(m+1)^2}-\frac{m(m+2)}{(m+1)\binom{n}{m}}-\frac{m^2}{(m-1)n},\end{aligned}$$

and some simple algebraic manipulations show the claim.

□

By using Stirling's formula we can also determine the asymptotics for the distance between the last two hirings where we again consider the four different relations between n and m, like we did for $L_{n,m}$.

Corollary 3.2.12. For the limiting distribution of $\Delta_{n,m}$ when n tends to infinity it holds with respect to the relation between m and n that

(i) m fixed: Suitably normalized, $\Delta_{n,m}$ converges in distribution to a continuous random variable which is characterized by its density function, i.e. $\Delta_{n,m} \xrightarrow{(d)} X_m$, where X_m has the density function

$$f_m(x) = m^2\left((-1)^m x^{m-1}\log x + (-1)^{m-1}H_{m-1}x^{m-1} + \sum_{k=0}^{m-2}\frac{(-1)^k}{m-1-k}\binom{m-1}{k}x^k\right),$$

having $0 < x < 1$.

(ii) $m \to \infty$, $m = o(n)$: Suitably normalized, $\Delta_{n,m}$ is asymptotically exponential distributed with parameter 1, i.e.

$$\frac{m}{n}\Delta_{n,m} \xrightarrow{(d)} \mathrm{Exp}(1).$$

(iii) $m \sim \alpha n$, where it holds that $0 < \alpha < 1$: $\Delta_{n,m} - 1$ is asymptotically geometrically distributed with success probability α, i.e.

$$\Delta_{n,m} - 1 \xrightarrow{(d)} \mathrm{Geom}(\alpha).$$

(iv) $n-m=o(n)$: $\Delta_{n,m}-1$ converges in distribution to 0, i.e.

$$\Delta_{n,m}-1 \xrightarrow{(d)} 0.$$

Proof. The proof of the asymptotics for $\Delta_{n,m}$ is more difficult than for the index of the last hired candidate as we have no closed form for the probabilities. However, using Stirling's formula together with some adaptions to the sum shows the claims directly.

(i) m fixed. We start with the probabilities for $\Delta_{n,m}$. Applying (3.19) to the binomial coefficients first shows directly

$$\begin{aligned}\mathbb{P}(\Delta_{n,m}=d) &= \frac{m}{\binom{n}{m}} \sum_{j=m+d}^{n} \frac{1}{j-m}\binom{j-d-1}{m-1} \\ &= \frac{m\ m!}{n^m\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right)} \sum_{j=m+d}^{n} \frac{(j-d)^{m-1}}{(j-m)\ (m-1)!}\left(1+\mathcal{O}\left(\frac{1}{j-d}\right)\right) \\ &\sim \frac{m^2}{n}\int_{\frac{d}{n}}^{1} \frac{1}{t}\left(t-\frac{d}{n}\right)^{m-1} dt,\end{aligned}$$

whereat the asymptotic formula yields from the substitution $t=\frac{j}{n}$. Hence, $\Delta_{n,m} \xrightarrow{(d)} X_m$, having the density function $f_m(x)=m^2\int_x^1 \frac{1}{t}(t-x)^{m-1}dt$. When applying Newton's binomial theorem we can get a more explicit form of the density function given by

$$\begin{aligned}f_m(x) &= m^2\int_x^1 \frac{1}{t}\sum_{k=0}^{m-1}\binom{m-1}{k}(-1)^k x^k t^{m-1-k} dt \\ &= m^2\left(\underbrace{\sum_{k=0}^{m-2}\binom{m-1}{k}\frac{(-1)^k x^k}{m-1-k}(1-x^{m-1-k})}_{=:S} + (-1)^m x^{m-1}\log x\right)\end{aligned}$$

For S we get

$$S=\sum_{k=0}^{m-2}\binom{m-1}{k}\frac{(-1)^k x^k}{m-1-k} + x^{m-1}\underbrace{\sum_{k=0}^{m-2}\binom{m-1}{k}\frac{(-1)^{k-1}}{m-1-k}}_{=:S_2},$$

whilst substituting $i=m-1-k$ in S_2 and using an identity for the resulting sum, that can be found in [19, p.281], leads to

$$S_2=(-1)^{m-1}\sum_{i=1}^{m-1}\binom{m-1}{i}\frac{(-1)^{i-1}}{i} = \frac{(-1)^{m-1}}{(m-1)!}\begin{bmatrix} m \\ 2\end{bmatrix} = (-1)^{m-1}H_{m-1},$$

which shows the claim.

(ii) $m \to \infty, m = o(n)$. Doing some simplifications to $\mathbb{P}(\Delta_{n,m})$ leads to

$$\begin{aligned}\mathbb{P}(\Delta_{n,m} = d) &= m^2 \sum_{j=d}^{n-m} \frac{(n-j-1)!\,(n-m)!}{(n-m-j+d)\,(n-j-m)!\,n!} \\ &= m^2 \sum_{j=d}^{n-m} \frac{1}{n-m-j+d}\,\frac{1}{n} e^{\frac{-jm}{n}} \left(1 + \mathcal{O}\left(\frac{j^2 m}{n^2}\right) + \mathcal{O}\left(\frac{jm^2}{n^2}\right)\right) \\ &\sim \frac{m}{n} \int_{\frac{dm}{n}}^{\infty} e^{-t} dt = \frac{m}{n} e^{\frac{-md}{n}}.\end{aligned}$$

where the main contribution for this local approximation is for $d = \mathcal{O}\left(\frac{n}{m}\right)$. Thus, the random variable $\frac{m}{n}\Delta_{n,m} \xrightarrow{(d)} X$, where X has the density function $f(x) = e^{-x}, x > 0$ which is the density function of an exponential distribution having $\lambda = 1$.

(iii) $m \sim \alpha n$. Again using (3.18) shows that

$$\begin{aligned}\mathbb{P}(\Delta_{n,m} = d) &= m^2 \sum_{j=d}^{n-m} \frac{(n-j-1)!\,(n-m)!}{(n-m-j+d)\,(n-j-m)!\,n!} \\ &= m^2 \sum_{j=d}^{\infty} \frac{(n-m)^j}{(n-m)n^{j+1}} \left(1 + \mathcal{O}\left(\frac{1}{n-m}\right)\right) \\ &\sim \frac{m}{n}\left(1 - \frac{m}{n}\right)^{d-1} = \alpha(1-\alpha)^{d-1}.\end{aligned}$$

To conclude, $\Delta_{n,m} - 1$ has asymptotically a geometric distribution with success probability $p = \alpha$.

(iv) $n - m = o(n)$. As almost each candidate gets recruited $\Delta_{n,m}$ takes small values in this case. Concretely for $d = 1$ we have

$$\mathbb{P}(\Delta_{n,m} = 1) = \frac{m^2}{(m-1)n} + \mathcal{O}\left(\frac{1}{n}\right) = \frac{(n - o(n))^2}{(n - o(n) - 1)n} + \mathcal{O}\left(\frac{1}{n}\right) = 1 + o(1).$$

which directly shows that $\Delta_{n,m} \xrightarrow{(d)} 1$.

□

Let us now consider another measurement for the selectivity of the hiring process given by the rank of the best discarded candidate $M_{n,m}$. We first only consider the non-trivial case of $1 \le m \le n$. According to [14] it is useful to use a simple trick when computing the characteristics of this parameter. Therefore we introduce the auxiliary variable $a_{n,m,j}$, $0 \le j \le n-m$, which denotes the probability that at least all of the $m+j$ highest ranked candidates are hired. If we know the values of $a_{n,m,k}$ one may obviously get the probabilities for $M_{n,m}$ directly by

$$\mathbb{P}(M_{n,m} = b) = a_{n,m,n-m-b} - a_{n,m,n+1-m-b}. \tag{3.23}$$

So our first task is to determine $a_{n,m,k}$. We can easily get a recurrence for $a_{n,m,k}$ when considering the probabilities for a hiring. One may achieve a sequence of length n where the best $m+j$ candidates get hired either by hiring a candidate after a sequence of length $n-1$ where the best $m+j-1$ candidates are hired or discarding a candidate after a sequence of length $n-1$ where the best $m+j$ candidates have already been hired. However, in the second case the discarded applicant must not belong to the best $m+j$ candidates at step n. Thus, we get the following recurrence for $1 \le j \le n-m$:

$$a_{n,m,j} = \frac{m}{n}\, a_{n-1,m,j-1} + \left(1 - \frac{j+m}{n}\right) a_{n-1,m,j}, \tag{3.24}$$

where the starting values are given by $a_{n,m,0} = 1$, since it holds that $\mathcal{R}(\sigma) = m$ for each permutation of the ranks $\sigma \in \mathcal{S}_n$ when applying hiring above the m-th best.

Lemma 3.2.13. The solution of Recursion (3.24) for the variable $a_{n,m,j}$ is given by

$$a_{n,m,j} = \frac{m!\, m^j}{(m+j)!}.$$

Proof. When solving this recurrence it is applicable to modify it by introducing

$$b_{n,m,j} = \frac{n!}{m!(n-m-j)!}\, a_{n,m,j},$$

which directly leads to the simpler recurrence

$$b_{n,m,j} = m b_{n-1,m,j-1} + b_{n-1,m,j}.$$

An approach with using generating functions then yields the PDE

$$(1-z)B(z,u,v) = zuvB_u(z,u,v) + \frac{z^2uv}{(1-z(1+u))^2},$$

where

$$B(z,u,v) = \sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{j=1}^{n-m} b_{n,m,j} z^n u^m v^j$$

is the multivariate generating function of $b_{n,m,j}$. They used a trick in [14] to get the coefficients $b_{n,m,j}$ without solving this PDE explicitly. Therefore consider the bivariate generating function $b_m(z,v) = [u^m]B(z,u,v)$, which satisfies

$$(1-z)b_m(z,v) = mzvb_m(z,v) + mv\left(\frac{z}{1-z}\right)^{m+1},$$

whereat the inhomogeneity can be computed in the following way:

$$\begin{aligned}
[u^m]\frac{z^2uv}{(1-z(1+u))^2} &= z^2v\,[u^{m-1}](1-z(1+u))^{-2} = z^2v\,[u^{m-1}]\sum_{k=0}^{\infty}\binom{-2}{k}(-1)^k z^k(1+u)^k \\
&= z^2v\sum_{k=0}^{\infty}\binom{k+1}{k} z^k\,[u^{m-1}](1+u)^k = z^2v\sum_{k=0}^{\infty}(k+1)z^k\binom{k}{m-1} \\
&= z^2vm\sum_{k=m-1}^{\infty}\binom{k+1}{m} z^k = z^{m+1}\sum_{k=0}^{\infty}\binom{m+k}{m} z^{m-1+k} \\
&= mv\left(\frac{z}{1-z}\right)^{m+1}.
\end{aligned}$$

The last identity is an application of Newton's generalized binomial theorem and can be found in [19, p.335]. Some simple algebraic manipulations then lead to

$$b_m(z,v) = \frac{mvz^{m+1}}{(1-z-mzv)(1-z)^{m+1}}.$$

Finally extracting the coefficients from the closed form of $b_m(z,v)$, by again applying the generalized binomial theorem, yields

$$\begin{aligned} b_{n,m,j} &= [z^n v^j] b_m(z,v) = [z^n] m \left(\frac{z}{1-z}\right)^{m+1} [v^j] \frac{v}{1-z-mzv} \\ &= [z^n] \left(\frac{z}{1-z}\right)^{m+1} \frac{(mz)^j}{(1-z)^{j-1}} = m^j [z^{n-m-j}] \frac{1}{(1-z)^{m+j+1}} = m^j \binom{n}{m+j}, \end{aligned}$$

and resubstituting finally gives us the claim for $a_{n,m,j}$:

$$\begin{aligned} a_{n,m,j} &= \frac{m!(n-m-j)!}{n!} b_{n,m,j} = \frac{m!(n-m-j)!}{n!} m^j \binom{n}{m+j} \\ &= \frac{m!m^j}{(m+j)!}, \quad 0 \le j \le n-m. \end{aligned}$$

□

With the exact values of the auxiliary variables $a_{n,m,j}$ we are now able to analyze the behavior of the random variable $M_{n,m}$ that denotes the rank of the best discarded candidate.

Theorem 3.2.14. Consider the rank of the best discarded candidate in a sequence of n applicants given by $M_{n,m}$. Then for the exact distribution of $M_{n,m}$ it holds that

$$\mathbb{P}(M_{n,m} = b) = \begin{cases} \delta_{b,0}, & n < m, \\ \frac{m!}{n!} m^{n-m}, & b = 0, \ 1 \le m \le n, \\ \frac{m!}{(n-b+1)!} (n-m-b+1) m^{n-m-b}, & 1 \le b \le n-m, \ 1 \le m \le n. \end{cases}$$

The expectation of $M_{n,m}$ is given by

$$\mathbb{E}(M_{n,m}) = \begin{cases} 0, & n < m, \\ n-m+1-\sum_{j=0}^{n-m} \frac{m!m^j}{(m+j)!} = n-m+\mathcal{O}(\sqrt{m}), & \text{else}, \end{cases}$$

where the asymptotic expansion holds uniformly for $1 \le m \le n$ and $n \to \infty$.

Note: We defined $M_{n,m} = 0$ if no candidate is discarded, i.e. $h_{n,m} = n$.

Proof. The case of $n < m$ is as usual a trivial one, as all applicants get recruited, so the rank of the best discarded candidate is 0. For $1 \le m \le n$ the case of $M_{n,m} = 0$ is exactly the same as $h_{n,m} = n$ which directly yields $\mathbb{P}(M_{n,m} = 0) = \frac{m!}{n!} m^{n-m}$. For the last case, (3.23) together with Lemma 3.2.13 show that

$$\begin{aligned} \mathbb{P}(M_{n,m} = b) &= a_{n,m,n-m-b} - a_{n,m,n+1-m-b} = \frac{m!m^{n-m-b}}{(n-b)!} - \frac{m!m^{n+1-m-b}}{(n+1-b)!} \\ &= \frac{m!}{(n+1-b)!}(n-m-b+1)m^{n-m-b}. \end{aligned}$$

For $1 \le m \le n$, the expectation is given by

$$\begin{aligned}\mathbb{E}(M_{n,m}) &= \sum_{b=0}^{n-m} b(a_{n,m,n-m-b} - a_{n,m,n+1-m-b}) = (n-m)a_{n,m,0} + \sum_{b=0}^{n-m-1} a_{n,m,n-m-b} \\ &= n-m+1+\sum_{b=0}^{n-m} \frac{m!\, m^j}{(m+j)!}.\end{aligned}$$

For the asymptotic expansion we use a result that can be found in [18], namely

$$\sum_{j=0}^{\infty} \frac{m!\, m^j}{(m+j)!} = \sqrt{\frac{\pi m}{2}} + \frac{1}{3} + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

and as a consequence we get for $n \to \infty$ that

$$\mathbb{E}(M_{n,m}) = n - m + \mathcal{O}(\sqrt{m}).$$

□

Corollary 3.2.15. The limit distribution of $M_{n,m}$ is, for $n \to \infty$ and respecting the relation between n and m, given by the following cases:

i) m fixed: $n - m - M_{n,m}$ converges in distribution to a discrete random variable, which is characterized by its density function M_m, i.e. $n - m - M_{n,m} \xrightarrow{(d)} M_m$, where M_m has the density function

$$\mathbb{P}(M_m = b) = \frac{(b+1)m^b m!}{(m+b+1)!}, \qquad b \in \mathbb{N}.$$

ii) $m \to \infty$, but $n - m \gg \sqrt{m}$: Suitably normalized, $n - m - M_{n,m}$ is asymptotically Rayleigh distributed with parameter $\sigma = 1$, i.e.

$$\frac{n - m - M_{n,m}}{\sqrt{m}} \xrightarrow{(d)} \text{Rayleigh}(1).$$

iii) $n - m \sim \alpha\sqrt{m}$, with $\alpha > 0$: Suitably normalized, $n - m - M_{n,m}$ converges in distribution to the minimum of α and a Rayleigh distributed random variable with parameter $\sigma = 1$, i.e.

$$\frac{n - m - M_{n,m}}{\sqrt{m}} \xrightarrow{(d)} \min(\alpha, \text{Rayleigh}(1)).$$

iv) $n - m = o(\sqrt{n})$: $M_{n,m}$ converges in distribution to 0, i.w. $M_{n,m} \xrightarrow{(d)} 0$.

Proof.

i) m fixed. The claim is a direct consequence of the exact distribution of $n - m - M_{n,m}$ for $n \to \infty$.

ii) $m \to \infty$, but $n - m \gg \sqrt{m}$. Applying Stirling's formula leads to the local expansion

$$\mathbb{P}(n - m - M_{n,m} = b) = \frac{b}{m} e^{-\frac{b^2}{2m}} \cdot \left(1 + \mathcal{O}\left(\frac{b}{m}\right) + \mathcal{O}\left(\frac{b^3}{m^2}\right)\right),$$

and as a direct consequence we have $\frac{n-m-M_{n,m}}{\sqrt{m}} \xrightarrow{(d)} \widehat{R}$, where $\widehat{R}$ has the density function $\widehat{f}(x) = xe^{\frac{-x^2}{2}}$, $x > 0$, which is exactly the density function of a Rayleigh distribution with parameter $\sigma = 1$.

iii) $n - m \sim \alpha\sqrt{m}$, with $\alpha > 0$. We consider the random variable $\widehat{M}_{n,m} := n - m - M_{n,m}$ and compute its distribution function $\mathbb{P}\left(\widehat{M}_{n,m} \leq k\right)$. For $0 \leq k \leq n - m - 1$ we get

$$\begin{aligned}\mathbb{P}\left(\widehat{M}_{n,m} \leq k\right) &= \sum_{i=0}^{k} \mathbb{P}\left(\widehat{M}_{n,m} = i\right) = \sum_{i=0}^{k} \frac{(i+1)m!m^i}{(m+i+1)!} \\ &= \sum_{i=0}^{k} \left(\frac{m!m^i}{(m+i)!} - \frac{m!m^{i+1}}{(m+j+1)!}\right) = 1 - \frac{m!m^{k+1}}{(m+k+1)!},\end{aligned}$$

since the sum telescopes. Obviously it is true that $\mathbb{P}\left(\widehat{M}_{n,m} \leq n - m\right) = 1$. Hence, the distribution function of $\widehat{M}_{n,m}$ is given by

$$\mathbb{P}\left(\widehat{M}_{n,m} \leq k\right) \begin{cases} 1 - \frac{m!m^{k+1}}{(m+k+1)!}, & 0 \leq k \leq n - m - 1, \\ 1, & k \geq n - m. \end{cases}$$

For the region $n - m \sim \alpha\sqrt{m}$, $\alpha > 0$, we get the asymptotic expansion

$$\mathbb{P}\left(\widehat{M}_{n,m} \leq x\right) \longrightarrow 1 - e^{-\frac{x^2}{2}}, \qquad 0 \leq x < \alpha,$$

which again is a consequence of Stirling's formula and the substitution $k = x\sqrt{m}$. For $x \geq \alpha$ the distribution function of $\widehat{M}_{n,m}$ converges to 1, i.e. $\mathbb{P}\left(\frac{\widehat{M}_{n,m}}{\sqrt{m}} \leq x\right) \longrightarrow 1$. Thus, we have the asymptotic behavior

$$\frac{\widehat{M}_{n,m}}{\sqrt{m}} \xrightarrow{(d)} \widehat{M}_\alpha,$$

where $\widehat{M}_\alpha$ has the distribution function

$$\widehat{F}_\alpha(x) = \begin{cases} 1 - e^{-\frac{x^2}{2}}, & 0 \leq x < \alpha, \\ 1, & x \geq \alpha, \end{cases}$$

which is the composition of a Rayleigh distribution with parameter $\sigma = 1$ and a Diracian distribution at $x = \alpha$. Thus, $\widehat{F}_\alpha(x)$ is the distribution function of $\min(\text{Rayleigh}(1), \alpha)$.

iv) $n - m = o(\sqrt{n})$. In this case almost each candidate gets hired. Defining $k = n - m$ and applying Stirling's formula shows that

$$\mathbb{P}(M_{n,m} = 0) = \frac{m!m^{n-m}}{n!} = \frac{(n-k)!(n-k)^k}{n!} = e^{-\frac{k^2}{2n}} \cdot \left(1 + \mathcal{O}\left(\frac{k}{n}\right) + \mathcal{O}\left(\frac{k^3}{n^2}\right)\right).$$

Using the fact $k = o(\sqrt{n})$ finally gives us

$$\mathbb{P}(M_{n,m} = 0) = e^{o(1)} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) = 1 + o(1).$$

Thus, $M_{n,m}$ converges, for $n \to \infty$, in distribution to 0.

□

At the end of our considerations about hiring above the m-th best we consider the adaption hiring and firing of the hiring problem. Let us recap the hiring criteria with hiring and firing. A candidate gets hired if and only if his rank is beneath the m best candidates seen so far (normal hiring). Or he is better than the worst already hired candidate whom he replaces in this case (firing of the worst so far hired candidate). As mentioned before the size of the hiring set is not affected by this adaption. The parameter we investigate with hiring and firing is the number $f_{n,m}$ of replacements. By following [14] we will give the expectation of $f_{n,m}$ but also the probability for a replacement at step $j \leq n$. Additionally we were able to establish a recursion for $f_{n,m}$ which we unfortunately could not solve with standard techniques.

Let us first introduce the sequence of indicator variables $Y_{j,m}$ for a replacement which is given by

$$Y_{j,m} = \begin{cases} 0, & \text{there is no replacement at step } j, \\ 1, & \text{else.} \end{cases}$$

With the law of complete probability we may now easily determine the probability for a replacement at step j which is given by $\mathbb{P}(Y_{j,m} = 1)$.

Lemma 3.2.16. The probability for a replacement at the j-th candidate in a sequence of length n when performing hiring and firing with hiring above the m-th best is given by

$$\mathbb{P}(Y_{j,m} = 1) = \begin{cases} 0, & j \leq m, \\ \frac{m}{j}(H_{j-1} - H_m), & \text{else.} \end{cases} \tag{3.25}$$

Proof. Obviously the number of replacements, as well as the probability for a replacement, depends on the size of the hiring set. For the conditioned probability of $Y_{n,m}$ it holds that

$$\mathbb{P}(Y_{j,m} = 1 \mid h^{[f]}_{j-1,m} = k) = \begin{cases} 0, & i \leq m, \\ \frac{k-m}{j}, & j > m,\ m \leq k \leq j. \end{cases}$$

This equation follows in a natural way by definition of hiring and firing. Therefore let k denote the size of the hiring set at step $j-1$. Because of the replacement mechanism the hiring set contains exactly the ranks $j-k, j-(k-1), \dots, j-1$. The possible ranks for a replacement when hiring X_j are $j-k, j-(k-1), \dots, j-m-1$. For all higher ranks X_j would get recruited without a replacement. Thus, the probability for a replacement is given by

$$\mathbb{P}(Y_{j,m} = 1 \mid h^{[f]}_{j-1,m} = k) = \frac{(j-m-1)-(j-k-1)}{j} = \frac{k-m}{j}.$$

The law of complete probability together with Theorem 3.2.3 implies that

$$\begin{aligned}\mathbb{P}(Y_{i,m}=1) &= \sum_{k=m}^{j-1} \mathbb{P}(Y_{j,m}=1 \mid h_{j-1,m}=k)\mathbb{P}(h_{j-1,m}=k) \\ &= \sum_{k=1}^{j-m} \frac{k-1}{j} \frac{m!\, m^{k-1}}{(j-1)!} \begin{bmatrix} j-m \\ k \end{bmatrix}. \end{aligned} \tag{3.26}$$

Using the logarithmic deviation of the BGF of the Stirling numbers given by

$$\frac{\partial}{\partial u}\left[\frac{1}{(1-z)^u}\right] = \log\left(\frac{1}{1-z}\right)\frac{1}{(1-z)^u} = \sum_{n=0}^{\infty}\sum_{k=0}^{n} k \begin{bmatrix} n \\ k \end{bmatrix} \frac{z^n}{n!} u^{k-1},$$

and applying Newton's generalized binomial theorem as well as the identity $\log(1+x) = \sum_{i=1}^{\infty}(-1)^{i+1}\frac{x^i}{i}$ shows the two identities

$$\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} u^k = n!\binom{n+u-1}{n}, \qquad \sum_{k=0}^{n} k \begin{bmatrix} n \\ k \end{bmatrix} u^{k-1} = n!\binom{n+u-1}{n}\sum_{j=1}^{n}\frac{1}{j+u-1}.$$

Applying these identities evaluated at $u=m$ to (3.26) gives us

$$\begin{aligned}\mathbb{P}(Y_{j,m}=1) &= \sum_{k=1}^{j-m} \frac{k-1}{j} \frac{m!\, m^{k-1}}{(j-1)!} \begin{bmatrix} j-m \\ k \end{bmatrix} = \frac{m!}{j!}\left[\sum_{k=0}^{j-m} k \begin{bmatrix} j-m \\ k \end{bmatrix} m^{k-1} - \frac{1}{m}\sum_{k=0}^{j-m} \begin{bmatrix} j-m \\ k \end{bmatrix} m^k\right] \\ &= \frac{m!}{j!}\left[(j-m)!\binom{j-1}{j-m}\cdot\left(\sum_{k=1}^{j-m}\frac{1}{k+m-1}\right) - \frac{(j-m)!}{m}\binom{j-1}{j-m}\right] \\ &= \frac{m}{j}(H_{j-1}-H_m)\end{aligned}$$

□

With the probabilities of the indicator variables it is an easy task to compute the expectation of $f_{n,m}$ by exploitation of its linearity.

Theorem 3.2.17. For the expectation of the number $f_{n,m}$ of replacements when using the replacement mechanism with hiring above the m-th best it holds that

$$\mathbb{E}(f_{n,m}) = \frac{m}{2}\left((H_n - H_m)^2 - (H_n^{(2)} - H_m^{(2)})\right).$$

Proof. Formula (3.25) together with the linearity of the expectation shows that

$$\begin{aligned}\mathbb{E}(f_{n,m}) &= \sum_{j=m+1}^{n} \mathbb{E}(Y_{j,m}) = \sum_{j=m+1}^{n} \mathbb{P}(Y_{j,m}=1) \\ &= \sum_{j=m+1}^{n} \frac{m}{j}(H_{j-1}-H_m) = m\sum_{j=m+1}^{n}\frac{H_{j-1}}{j} - mH_m(H_n-H_m).\end{aligned} \tag{3.27}$$

For the last sum (3.15) implies that

$$\sum_{j=m+1}^{n} \frac{H_{j-1}}{j} = \sum_{j=m+1}^{n} \frac{H_j}{j} - \left(H_n^{(2)} - H_m^{(2)}\right) = \frac{1}{2}\left(H_n^2 - H_n^{(2)} - H_m^2 + H_m^{(2)}\right). \tag{3.28}$$

Finally plugging (3.28) into (3.27) yields in

$$\begin{aligned} \mathbb{E}(f_{n,m}) &= \frac{m}{2}\left(H_n^2 - H_n^{(2)} - H_m^2 + H_m^{(2)} - 2H_nH_m + 2H_m^2\right) \\ &= \frac{m}{2}\left((H_n - H_m)^2 - \left(H_n^{(2)} - H_m^{(2)}\right)\right). \end{aligned}$$

□

When computing the exact probabilities for $f_{n,m}$ we came up against a recursion which we could not solve with standard techniques. Although we have no solution for this recursion we will remark our approach.

Remark 3.2.18. Let $f_{n,m,k} = \mathbb{P}(f_{n,m} = k)$ denote the probability for k replacements in a sequence of length n. Then the following recursion holds:

$$f_{n,m,k} = \frac{m}{n}(H_{n-1} - H_m)f_{n-1,m,k-1} + \left(1 - \frac{m}{n}(H_{n-1} - H_m)\right) f_{n-1,m,k}, \quad n > m,\ 0 \leq k \leq n.$$

For the starting values we get

$$\begin{aligned} f_{m+1,m,k} &= \delta_{0,k}, \\ f_{n,m,0} &= \prod_{i=m+1}^{n}\left(1 - \frac{m}{i}(H_{i-1} - H_m)\right). \end{aligned}$$

The recursion can be derived from the two cases that appear at step $n-1$. Either we have had $k-1$ replacements until step $n-1$ and another one at step n or we have already had k replacements until step $n-1$ and no additional one. The probabilities for the cases are a direct consequence of Lemma 3.2.16.

For the starting values we can make a similar consideration. For $n = m+1$ there are no replacements for sure, as the first m candidates get recruited by the strategy. So X_{m+1} does not get hired by the strategy if and only if his rank is 1. But in this case he has a lower rank than all already hired candidates, so we have no replacement at step $m+1$.

The second type of starting values follows from the probability that there is no replacement until the n-th candidate.

■

Chapter 4

Hiring above the median

The last Lake Wobegon strategy we discuss is *hiring above the median*. Here the hiring process recruits the first candidate for sure. Further candidates get recruited if and only if their relative rank is higher than the median score of all already hired candidates. Concretely this means that in a sequence of n candidates with the permutation of their ranks $\sigma \in \mathcal{S}_n$ candidate X_i gets hired if and only if $\sigma_i(i) > \mathfrak{m}(\{\sigma_i(j) \ : \ j \in \mathcal{H}_i(\sigma)\})$. Therefore we define the *median* for any set $\mathcal{C} := \{c_1, c_2, \ldots, c_k\}$ that holds $c_1 < c_2 < \ldots < c_k$ by $\mathfrak{m}(\mathcal{C}) := c_{\lceil \frac{k}{2} \rceil}$. Like we did for all the other strategies we call the person who currently holds the median rank of all previously hired candidates the threshold candidate. Obviously any applicant gets hired if and only if his score is better than the score of the current threshold candidate. This shows that hiring above the median is a pragmatic hiring strategy.

The handicap when analyzing the parameters is that we have to distinguish between two cases when recruiting a new candidate: Either the median value of the hiring set changes or not, which affects the threshold candidate for the next interview. We did not have to deal with that problem when analyzing hiring above the minimum, where the threshold never changes, or hiring above the m-th best, where the threshold changes each time we hire a candidate. Last but not least we want to state that the hiring process always increases the quality of the hiring set as we only hire candidates whose rank is better than the median rank of the previously recruited candidates, which characterizes hiring above the median as a Lake Wobegon strategy.

4.1 Hiring above the median

When analyzing this strategy we will mostly follow the considerations of [15] who solved the problem we mentioned above very skillfully by considering an automaton. This automaton then leads to a system of recurrences for the two different cases. First of all let us give a more precise description of this strategy.

Strategy 4.1.1 (Hiring above the median). Consider a sequence of n candidates and the permutation of their ranks $\sigma \in \mathcal{S}_n$. Hiring above the median builds up the hiring set $\mathcal{H}(\sigma)$ in the following way:

- $\mathcal{H}_1(\sigma) = \{1\}$.
- $\mathcal{H}_i(\sigma) = \begin{cases} \mathcal{H}_{i-1}(\sigma) \cup \{i\}, & \sigma(i) > \mathfrak{m}(\{\sigma_i(j) \ : \ j \in \mathcal{H}_{i-1}(\sigma)\}), \\ \mathcal{H}_{i-1}, & \text{else}, \end{cases} \qquad 2 \leq i \leq n.$

Note that the threshold candidate changes after exactly every second hiring.

■

Let us first consider an example to get a better insight to our strategy and the behavior of how the threshold changes from step to step.

Example 4.1.2. We consider the hiring process with hiring above the median on the permutation $\sigma = 463921785$. The underlined numbers in the table mark the previously hired candidates in each step.

candidate	rel. rank	σ_i	thres. cand.	thres. rank	hired	hiring set
X_1	1	1	-	-	✓	$\{1\}$
X_2	2	$\underline{1}$ - 2	X_1	1	✓	$\{1,2\}$
X_3	1	$\underline{23}$ - 1	X_1	2		$\{1,2\}$
X_4	4	$\underline{23}1$ - 4	X_1	2	✓	$\{1,2,4\}$
X_5	1	$\underline{34}2\underline{5}$ - 1	X_2	4		$\{1,2,4\}$
X_6	1	$\underline{45}3\underline{6}2$ - 1	X_2	5		$\{1,2,4\}$
X_7	6	$\underline{45}3\underline{7}21$ - 6	X_2	5	✓	$\{1,2,4,7\}$
X_8	7	$\underline{45}3\underline{8}21\underline{6}$ - 7	X_2	5	✓	$\{1,2,4,7,8\}$
X_9	5	$\underline{46}3\underline{9}21\underline{78}$ - 5	X_7	7		$\{1,2,4,7,8\}$

For a possible next candidate X_{10} the threshold would not change as candidate X_9 is discarded, but even if X_9 was hired the threshold candidate would still be X_7 as the median only increases with every second hiring.

■

Let us first consider the parameter h_n which gives us the number of recruited candidates after n interviews. In [15] they therefore introduced auxiliary variables to compute the exact probabilities for h_n. Their approach bases on the observation that in a set of k different values (which in our case denote the ranks of the hired applicants) $r_1 < r_2 < \ldots < r_k$ we can describe the median value $r_{\lceil \frac{k}{2} \rceil}$ as the ℓ-th largest element of this set, i.e. $r_{k+1-\ell}$, with $\ell = \left\lceil \frac{k+1}{2} \right\rceil$. It is obvious that the size of the hiring set is closely related to the index ℓ of the median in the increasing sequence of ranks of the hired candidates. Let us assume we hired k candidates in our hiring process, i.e. $h(\sigma) = k$, where $\sigma \in \mathcal{S}_n$ denotes the permutation of the ranks. Then it holds that

$$\ell = \begin{cases} \frac{k+1}{2}, & k \equiv 1(2), \\ \frac{k}{2} + 1, & k \equiv 0(2), \end{cases}$$

and vice versa. So there is a direct connection between ℓ and the size of the hiring set $h_n = k$. This simple observation is the background of the approach in [15]. As we have to respect the parity of k, we introduce two auxiliary variables $a^{[1]}_{n,\ell}$ and $a^{[2]}_{n,\ell}$ that describe the probabilities of the size of the hiring set.

$$\mathbb{P}(h_n = k) = \begin{cases} a^{[1]}_{n,\frac{k+1}{2}}, & k \equiv 1(2), \\ a^{[2]}_{n,\frac{k}{2}+1}, & k \equiv 0(2). \end{cases} \tag{4.1}$$

Figure 4.1 shows the automaton that models the connection between the two auxiliary variables and gives the probabilities how the state may change when a new candidate is interviewed. Furthermore, it shows how the parameters n and ℓ change. The two nodes give the parity of k, i.e. if $a^{[1]}_{n,\ell}$ or $a^{[2]}_{n,\ell}$ is concerned. If the currently interviewed applicant is discarded, the size of the hiring set k does not change and as a consequence its parity does not change either. This case is modeled by the loops at each of the two nodes. On the other hand hiring a candidate changes the parity of k and so we have to go to the other state in the automaton. As we will see the numbers $a^{[1]}_{n,\ell}$ and $a^{[2]}_{n,\ell}$ play an important role for all of the investigated quantities.

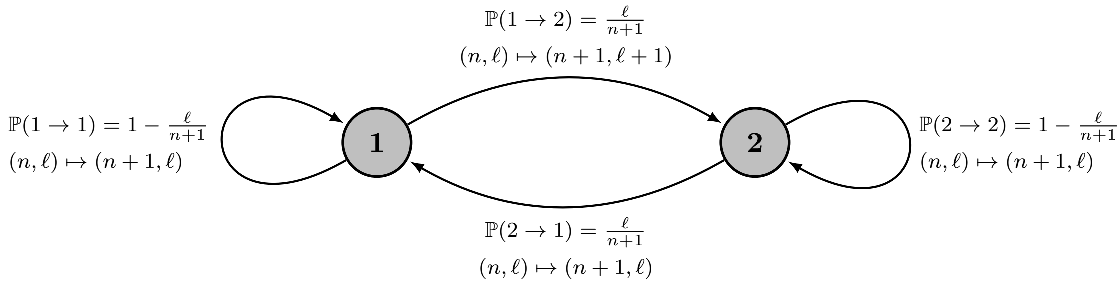

Figure 4.1: The automaton describing the relation and transition between the two variables $a^{[1]}_{n,\ell}$ and $a^{[2]}_{n,\ell}$.

As we can see the threshold changes together with the transition from state 1 to state 2. Starting from the automaton given in Figure 4.1 we receive a linked system of recursions for $a^{[1]}_{n,\ell}$ and $a^{[2]}_{n,\ell}$ given by

$$a^{[1]}_{n,\ell} = \frac{\ell}{n}\ a^{[2]}_{n-1,\ell} + \left(1 - \frac{\ell}{n}\right)\ a^{[1]}_{n-1,\ell}, \tag{4.2}$$

$$a^{[2]}_{n,\ell} = \frac{\ell-1}{n}\ a^{[1]}_{n-1,\ell-1} + \left(1 - \frac{\ell}{n}\right)\ a^{[2]}_{n-1,\ell}, \tag{4.3}$$

for $n \geq 2$ and $1 \leq \ell \leq n$ with the starting values $a^{[1]}_{1,1} = 1$ and $a^{[2]}_{1,1} = 0$. Outside the range we define $a^{[1]}_{n,\ell} = a^{[2]}_{n,\ell} = 0$.

Before we start solving this system of recursions we go back to Example 4.1.2 and reconstruct the hiring process with focus on the automaton.

Example 4.1.3 (Example 4.1.2, continued). Again we consider the permutation $\sigma = 463921785$ and observe the transitions in the automaton given in Figure 4.1 during the hiring process.

X_i	$\sigma_i(i)$	σ_i	hired	ℓ	state
X_1	1	1	✓	- $\mapsto$ 1	$\to$ 1
X_2	2	$\underline{1}$ - 2	✓	1 $\mapsto$ 2	1 $\to$ 2
X_3	1	$\underline{23}$ - 1		2	2
X_4	4	$\underline{23}1$ - 4	✓	2	2 $\to$ 1
X_5	1	$\underline{34}2\underline{5}$ - 1		2	1
X_6	1	$\underline{45}3\underline{6}2$ - 1		2	1
X_7	6	$\underline{45}3\underline{7}21$ - 6	✓	2 $\mapsto$ 3	1 $\to$ 2
X_8	7	$\underline{45}3\underline{8}21\underline{6}$ - 7	✓	3	2 $\to$ 1
X_9	5	$\underline{46}3\underline{9}21\underline{78}$ - 5		3	1

As we can observe the value of ℓ changes exactly in the contrary hirings as the threshold candidate does. The reason for this lies within the fact that when the threshold candidate changes, he changes to the very next better ranked candidate. As we only hire applicants better than the threshold we get one better candidate in this case but also change the threshold to the next candidate. So the threshold still stays the ℓ-th best hired candidate. On the other hand when the threshold does not change we have additionally one more better hired candidate and ℓ increases by one.

■

Let us now try to find a solution to the system of recursions (4.2) and (4.3). Obviously when we find a solution for $a^{[1]}_{n,\ell}$ we immediately get a solution for $a^{[2]}_{n,\ell}$ too and vice versa. We first try to find a recurrence for $a^{[1]}_{n,\ell}$ that is independent of the values $a^{[2]}_{n,\ell}$. By substituting (4.3) into (4.2) inductively one finally obtains such an independent recursion given by

$$a^{[1]}_{n,\ell} = \left(1 - \frac{\ell}{n}\right) a^{[1]}_{n-1,\ell} + \sum_{m=1}^{n-2} a^{[1]}_{m,\ell} \, \frac{\ell - 1}{m+1} \, \frac{\ell}{n} \prod_{j=m+1}^{n-2} \left(1 - \frac{\ell}{j+1}\right). \tag{4.4}$$

We could also obtain this recursion by considering our automaton again. For resulting in state 1 at step n we could have either been in state 1 at step $n-1$ and have stayed there, which is given by the first summand, or we have been in state 2 at step $n-1$, which yields the sum in the second term. Note that we initially start in state 1. So if we come back to 1 from state 2 we must have jumped to state two at some step m before, then stayed there until step $n-1$ and then got back to 1.

Following [15] it is useful to multiply (4.4) with $(n-\ell)\binom{n}{\ell}$ and introduce auxiliary variables

$$b^{[1]}_{n,\ell} := \binom{n}{\ell} \, a^{[1]}_{n,\ell},$$

which then yields following recursion

$$(n-\ell) b^{[1]}_{n,\ell} = (n-\ell) b^{[1]}_{n-1,\ell} + \sum_{m=1}^{n-2} (\ell - 1) b^{[1]}_{m,\ell-1}. \tag{4.5}$$

Applying the method of generating functions to this recursion by firstly multiplying it with $z^n u^\ell$ leads to

$$(n-\ell) b^{[1]}_{n,\ell} z^n u^\ell = z(n-\ell) b^{[1]}_{n-1,\ell} z^{n-1} u^\ell + \sum_{m=1}^{n-2} z^{n-m} u(\ell-1) b^{[1]}_{m,\ell-1} z^m u^{\ell-1}.$$

Secondly summing over $\ell \geq 2$ and applying some simple algebraic manipulations yield

$$nb_n^{[1]}(u)z^n - ub_n^{[1]\prime}(u)z^n = nb_{n-1}^{[1]}(u)z^n - ub_{n-1}^{[1]\prime}(u)z^n + \sum_{m=1}^{n-2} u^2 b_m^{[1]\prime}(u)z^n$$

where $b_n^{[1]}(u)$ is the horizontal generating function of the bivariate generating function of the sequence $b_{n,\ell}^{[1]}$

$$B^{[1]}(z,u) := \sum_{n=1}^{\infty}\sum_{\ell=1}^{n} b_{n,\ell}^{[1]}\, z^n u^\ell.$$

Finally summing over $n \geq 1$ leads to

$$zB_z^{[1]}(z,u) - uB_u^{[1]}(z,u) = zB^{[1]}(z,u) + z^2 B_z^{[1]}(z,u) - zuB_u^{[1]}(z,u) + \frac{z^2u^2}{1-z} B_u^{[1]}(z,u)$$

and some simple algebraic manipulations finally yield the linear partial differential equation (PDF)

$$z(1-z)B_z^{[1]}(z,u) + \left(zu - u - \frac{z^2u^2}{1-z}\right) B_u^{[1]}(z,u) - zB^{[1]}(z,u) = 0. \tag{4.6}$$

This linear PDE can be solved by applying the *method of characteristics*, which can be found in [22]. This method introduces a function f that satisfies $f(z,u,B^{[1]}) = 0$ and transforms (4.6) into a system of ordinary differential equations given by

$$\begin{aligned} \dot{z} &= z(1-z), \\ \dot{u} &= zu - u - \frac{u^2z^2}{1-z}, \\ \dot{B}^{[1]} &= zB^{[1]}. \end{aligned}$$

For solving this autonomic system of differential equations one may consider the derivatives

$$\frac{du}{dz} = \frac{zu - u - \frac{u^2z^2}{1-z}}{z(1-z)} = -\frac{u}{z} - \frac{u^2 z}{(1-z)^2}, \tag{4.7}$$

$$\frac{dB^{[1]}}{dz} = \frac{B^{[1]}}{1-z}. \tag{4.8}$$

The first ODE (4.7) is a Bernoulli differential equation, which can be solved by applying the substitution $x = u^{-1}$. This leads to the inhomogeneous ODE

$$x' = \frac{dx}{dz} = \frac{x}{z} - \frac{z}{1-z}, \tag{4.9}$$

which can be solved with standard techniques. For the homogeneous differential equation, applying separation of the variables leads to the solution

$$x = C_1 \cdot z,$$

with an arbitrary real constant C_1. For the inhomogeneous equation, variation of the constant shows that

$$x = C_1(z) \cdot z, \tag{4.10}$$
$$x' = C_1(z) + C_1'(z) \cdot z. \tag{4.11}$$

Plugging (4.10) and (4.11) into (4.9) then yields

$$C_1'(z) = \frac{1}{(1-z)^2},$$

and as a solution we get

$$C_1(z) = \frac{1}{1-z}.$$

Hence, the solution of (4.9) is given by

$$x = \frac{z}{1-z} + z \cdot C_1,$$

or equivalently,

$$C_1 = \frac{x}{z} - \frac{1}{1-z} = \frac{1}{uz} - \frac{1}{1-z} = \frac{1-z-uz}{uz(1-z)}.$$

The solution of (4.8) is given by

$$B = \frac{C_2}{1-z},$$

or equivalently,

$$C_2 = B^{[1]} \cdot (1-z),$$

and can be obtained by applying separation of the variables. Thus, we achieve the representation

$$f(z, u, B^{[1]}) = 0 = G\left((1-z)B^{[1]}, \frac{1-z-zu}{uz(1-z)}\right)$$

with an arbitrary differentiable function G. Finally applying the fundamental theorem of implicit functions, which can be found in [16], shows that

$$(1-z)B^{[1]} = F\left(\frac{1-z-zu}{uz(1-z)}\right).$$

Hence, we have the general solution

$$B^{[1]}(z,u) = \frac{1}{1-z}\, F\left(\frac{1-z-zu}{zu(1-z)}\right), \tag{4.12}$$

where $F(x)$ is an arbitrary differentiable function which we have to determine by applying the initial conditions. Following [15] it is quite useful to consider the modification

$$\widetilde{B}(z,u) := B^{[1]}\left(zu, \frac{1}{u}\right) = \sum_{n=1}^{\infty} \sum_{\ell=0}^{n-1} b^{[1]}_{n,n-\ell} z^n u^\ell. \tag{4.13}$$

The results in (4.12) and (4.13) directly yield

$$\widetilde{B}(z,u) = \frac{1}{1-zu}\ F\left(\frac{1-zu-z}{z(1-zu)}\right). \tag{4.14}$$

Due to the fact that the median only increases after every second hiring together with the parity of the hiring set one easily gets $a_{n,n}^{[1]} = 0$ for $n \geq 2$. Moreover, the initial condition gives us $a_{1,1}^{[1]} = 1$. As an immediate consequence we have $b_{n,n}^{[1]} = \delta_{n,1}$ for $n \geq 1$. Hence, putting this result into (4.13) gives us the initial condition $\widetilde{B}(z,0) = z$, which together with (4.14) shows that

$$z = \widetilde{B}(z,0) = F\left(\frac{1-z}{z}\right). \tag{4.15}$$

Subsituting $x = \frac{1-z}{z}$ in (4.15) then fully determines the function $F(x)$ by

$$F(x) = \frac{1}{1+x},$$

and as a consequence we finally get the generating function $B^{[1]}(z,u)$ by the following formula:

$$B^{[1]}(z,u) = \frac{1}{1-z}\ F\left(\frac{1-z-zu}{zu(1-z)}\right) = \frac{1-z-zu}{(1-z)\ \left(1+\frac{1-z-zu}{zu(1-z)}\right)} = \frac{zu}{1-z-z^2u}.$$

In a last step we have to extract the coefficients $b_{n,\ell}^{[1]}$ from the closed form of $B^{[1]}(z,u)$ which can easily be done by using Newton's generalized binomial theorem. Thus, we get the solution for $b_{n,\ell}^{[1]}$ by

$$\begin{aligned} b_{n,\ell}^{[1]} &= [z^n u^\ell] B^{[1]}(z,u) = [z^n u^\ell] \frac{zu}{1-z-z^2u} = [z^n]\frac{1}{1-z}[u^\ell]\frac{u}{1-\frac{z^2}{1-z}u} \\ &= [z^n]\frac{1}{1-z}\left(\frac{z^2}{1-z}\right)^{\ell-1} = [z^n]\frac{z^{2\ell-1}}{(1-z)^\ell} = [z^{n-2\ell+1}](1-z)^{-\ell} \\ &= \binom{-\ell}{n-2\ell+1}(-1)^\ell = \binom{n-\ell}{\ell-1}, \end{aligned}$$

and as an immediate consequence this result yields

$$a_{n,\ell}^{[1]} = \frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}}. \tag{4.16}$$

For determining the values of $a_{n,\ell}^{[2]}$ plugging this result into (4.2) gives us

$$\frac{\binom{n+1-\ell}{\ell-1}}{\binom{n+1}{\ell}} = \frac{\ell}{n+1} a_{n,\ell}^{[2]} + \left(1+\frac{\ell}{n+1}\right)\frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}},$$

which after some simple algebraic manipulations yields in

$$a_{n,\ell}^{[2]} = \frac{n+1}{\ell}\left(\frac{\binom{n+1-\ell}{\ell-1}}{\binom{n+1}{\ell}} - \frac{n+1-\ell}{n+1}\frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}}\right) = \frac{(\ell-1)(n-\ell+1)}{(n+2-2\ell)\ \ell}\ \frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}} = \frac{\binom{n-\ell}{\ell-2}}{\binom{n}{\ell-1}}. \tag{4.17}$$

Plugging all the results above together delivers the following theorem.

Theorem 4.1.4 (Helmi and Panholzer, [15, Theorem 1]). The exact distribution of the number of hired candidates h_n when applying hiring above the median for a sequence of length n is given by

$$\mathbb{P}(h_n = k) = \begin{cases} \frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}}, & k = 2\ell - 1, \\ \frac{\binom{n-\ell}{\ell-2}}{\binom{n}{\ell-1}}, & k = 2\ell - 2. \end{cases}$$

Proof. The claim is a direct consequence of Formula (4.1) together with (4.16) and (4.17). □

With these exact results for the probabilities of h_n we can also determine its limiting distribution which we state in the corollary below.

Corollary 4.1.5. The limiting distribution of the random variable h_n for $n \to \infty$ is, after normalization, asymptotically Rayleigh distributed, with parameter $\sigma = \sqrt{2}$, i.e.

$$\frac{h_n}{\sqrt{n}} \xrightarrow{(d)} \widehat{R},$$

where $\widehat{R}$ has the density function

$$\widehat{f}(x) = \frac{x}{2} e^{-\frac{x^2}{4}}, \quad x > 0.$$

Furthermore, the expectation of h_n satisfies $\mathbb{E}(h_n) = \sqrt{\pi n} + \mathcal{O}(1)$.

Proof. Using the notation of Corollary 1.3.5 shows the identities

$$\begin{aligned} a_{n,\ell}^{[1]} &= \frac{\ell}{n+1-2\ell}\, c(n,\ell), \\ a_{n,\ell}^{[2]} &= \frac{(\ell-1)(n-\ell+1)}{(n+1-2\ell)(n+2-2\ell)}\, c(n,\ell), \end{aligned}$$

where the numbers $c(n,\ell)$ are defined by

$$c(n,\ell) = \frac{\binom{n-\ell}{\ell}}{\binom{n}{\ell}}.$$

This yields in the asymptotic representations

$$a_{n,\ell}^{[1]} = \frac{\ell}{n} e^{-\frac{\ell^2}{n}} \cdot \left(1 + \mathcal{O}\left(\frac{\ell}{n}\right) + \mathcal{O}\left(\frac{\ell^3}{n^2}\right)\right), \tag{4.18}$$

$$a_{n,\ell}^{[2]} = \frac{\ell}{n} e^{-\frac{\ell^2}{n}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\frac{\ell}{n}\right) + \mathcal{O}\left(\frac{\ell^3}{n^2}\right)\right), \tag{4.19}$$

which are an immediate consequence of Corollary 1.3.5 and hold for $1 \leq \ell \leq n^{\frac{1}{2}+\varepsilon}$, where ε is an arbitrary chosen constant in $0 < \varepsilon < \frac{1}{6}$. For $\ell \geq n^{\frac{1}{2}+\varepsilon}$ these numbers are sub-exponentially small. The considerations above together with Theorem 4.1.4 lead to the asymptotic expansion for $\mathbb{P}(h_n = k)$ given by

$$\mathbb{P}(h_n = k) \sim \frac{k}{2n} e^{-\frac{k^2}{4n}},$$

which holds for $k \in [n^{\frac{1}{2}-\varepsilon}, n^{\frac{1}{2}+\varepsilon}]$ when n tends to infinity. Substituting $k = x\sqrt{n}$ shows that $\frac{h_n}{\sqrt{n}}$ converges for $n \to \infty$ in distribution to a random variable $\widehat{R}$ with density

$$\widehat{f}(x) = \frac{x}{2} e^{-\frac{x^2}{4}}, \quad x > 0,$$

which is the density function of a Rayleigh distribution with parameter $\sigma = \sqrt{2}$.

For the expectation we get

$$\mathbb{E}(h_n) = \sum_{k=1}^{n} k \; \mathbb{P}(h_n = k) = \sum_{\ell=1}^{n} \left((2\ell - 1) a_{n,\ell}^{[1]} + (2\ell - 2) a_{n,\ell}^{[2]} \right). \tag{4.20}$$

The asymptotic expansions for $a_{n,\ell}^{[1]}$ and $a_{n,\ell}^{[2]}$ given in the Formulae (4.18)-(4.19) then imply that

$$\begin{aligned} \mathbb{E}(h_n) &= \sum_{\ell=1}^{n^{\frac{1}{2}+\varepsilon}} \frac{4\ell^2}{n} e^{-\frac{\ell^2}{n}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\frac{\ell}{n}\right) + \mathcal{O}\left(\frac{\ell^3}{n^2}\right) \right) \\ &= 4\sqrt{n} \int_0^\infty x^2 e^{-x^2} dx \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right) = \sqrt{\pi n} + \mathcal{O}(1), \end{aligned}$$

where we used the substitution $x = \frac{\ell}{\sqrt{n}}$. For computing the integral one may consider the fact that

$$\frac{\sqrt{\pi}}{2} = \int_0^\infty e^{-x^2} dx = x e^{-x^2} \Big|_{x=0}^{\infty} - \int_0^\infty x e^{-x^2} (-2x) dx = 2 \int_0^\infty x^2 e^{-x^2} dx,$$

which, for instance, can be found in [7, p.744].

□

When putting the exact results of our auxiliary variables into (4.20) one comes to the exact representation

$$\mathbb{E}(h_n) = \sum_{\ell=1}^{n} \frac{\ell(2\ell - 1)(n + 2 - 2\ell) + 2(\ell - 1)^2(n + 1 - \ell)}{\ell(n + 1 - \ell)} \; \frac{\binom{n-\ell+1}{\ell-1}}{\binom{n}{\ell}}.$$

As the first candidate gets hired for sure it hypothesizes that there is a strong dependency of the number of hired candidates in a sequence of n applicants on the relative score of candidate X_1. A low score will lead to a high number of hirings and vice versa. Thus, we consider the random variable $h_{n,r} := [h_n \mid \sigma_n(1) = r]$ which denotes the number of hirings after n interviews conditioned by the event that the relative score of the first candidate, given by $\sigma_n(1)$, is r. Note that in contrast to hiring above the m-th best we have no additional parameter m in order to tighten or relax the hiring criteria. Thus, the dependency on the score of the first candidate is even more important.

Theorem 4.1.6. The exact distribution of the random variable $h_{n,r}$ defined above is, for $1 \leq k, r \leq n$, given as follows:

$$\mathbb{P}(h_{n,r} = k) = \begin{cases} (\ell-1)\, \frac{\binom{n-r-\ell}{\ell-2}}{\binom{n-r}{\ell}}, & k = 2\ell - 1, \\ \frac{(\ell-1)(\ell-2)}{\ell}\, \frac{\binom{n-r-\ell+1}{\ell-2}}{\binom{n-r}{\ell}}, & k = 2\ell - 2, \end{cases}$$

where the threshold candidate has the ℓ-th highest score. Asymptotically, for $n \to \infty$ and provided that $n - r \to \infty$, $h_{n,r}$ converges in distribution to a continuous limiting distribution S, i.e. $\frac{h_{n,r}}{\sqrt{n-r}} \xrightarrow{(d)} S$, which is characterized by its density function

$$h(x) = \frac{x^3}{8} e^{-\frac{x^2}{4}}, \quad x > 0.$$

Proof. The claim can be shown by introducing the sequence of auxiliary variables

$$\overline{a}_{n,k,r} := \mathbb{P}(h_n = k \ \wedge \ \sigma_n(1) = r),$$

where $\sigma_n \in \mathcal{S}_n$ denotes the permutation of the relative ranks. For determining the values of $\overline{a}_{n,k,r}$ we first show that all these values are fully determined by $\overline{a}_{n,k,1}$. Thus, we consider the set of all permutations $\sigma \in \mathcal{S}_n$ satisfying $\sigma(1) = r$ and $h(\sigma) = k$. This set is given by $\mathcal{S}_{n,k,r} := \{\sigma \in \mathcal{S}_n \mid \sigma(1) = r \ \wedge \ h(\sigma) = k\}$. If the rank of the first candidate equals r then none of the candidates holding a lower rank than r get hired because the first candidate is the first threshold and the rank of the threshold does not decrease during the hiring process. This allows us to reduce the sequence by eliminating all of these $r-1$ candidates. When performing the hiring process on this reduced sequence of candidates we will get the same number of hired candidates as we did before by eliminating the low quality candidates. Formally this means that we eliminate $r-1$ candidates from a permutation $\sigma \in \mathcal{S}_{n,k,r}$ and receive - after relabelling $r, r+1, \ldots n$ to $1, 2, \ldots, n-r+1$ - a permutation $\widehat{\sigma} \in \mathcal{S}_{n-r+1,k,1}$. We can also reverse this construction to get to the original permutation again. Therefore we have to insert the other $r-1$ candidates, which can be inserted at $r-1$ out of $n-1$ positions. Together with the possible permutations of the $r-1$ candidates we get the following recursion for $\mathcal{S}_{n,k,r}$:

$$|\mathcal{S}_{n,k,r}| = |\mathcal{S}_{n-r+1,k,1}| \binom{n-1}{r-1} (r-1)!$$

As a direct consequence we get a recursion for the numbers $\overline{a}_{n,k,r}$ given by

$$\overline{a}_{n,k,r} = \frac{|\mathcal{S}_{n,k,r}|}{n!} = \frac{|\mathcal{S}_{n-r+1,k,1}| \binom{n-1}{r-1} (r-1)!}{n!} = \frac{n-r+1}{n}\, \overline{a}_{n-r+1,k,1}. \tag{4.21}$$

The only thing we still have to do is determining the starting values $\overline{a}_{n,k,1}$. Therefore we use the fact that $\mathbb{P}(h_n = k) = \sum_{r=1}^{n} \overline{a}_{n,k,r}$, which, when plugging (4.21) into it leads to

$$\mathbb{P}(h_n = k) = \frac{1}{n} \sum_{r=1}^{n} (n-r+1) \overline{a}_{n-r+1,k,1} = \frac{1}{n} \sum_{r=1}^{n} r\ \overline{a}_{r,k,1},$$

where the last equality results from flipping the index $r \mapsto n-r+1$. Multiplying by n and building differences gives us the starting values

$$\overline{a}_{n,k,1} = \frac{n\mathbb{P}(h_n = k) - (n-1)\mathbb{P}(h_{n-1} = k)}{n},$$

which then shows the general solution for $\overline{a}_{n,k,r}$:

$$\overline{a}_{n,k,r} = \frac{(n-r+1)\mathbb{P}(h_{n-r+1} = k) - (n-r)\mathbb{P}(h_{n-r} = k)}{n}$$

The probabilities for the random variable $h_{n,r}$ is then given by

$$\mathbb{P}(h_{n,r} = k) = \frac{\overline{a}_{n,k,r}}{\mathbb{P}(\sigma_n(1) = r)} = n\ \overline{a}_{n,k,r} = (n-r+1)\mathbb{P}(h_{n-r+1} = k) - (n-q)\mathbb{P}(h_{n-r} = k). \tag{4.22}$$

Plugging the results from Theorem 4.1.4 into (4.22) finally shows the claim.

The result for the limit distribution is a direct result of Corollary 1.3.5 applied to the explicit result for the probabilities. Thus, it is true that

$$\mathbb{P}(h_{n,r} = k) \sim \frac{k^3}{8(n-r)^2} e^{-\frac{k^2}{4(n-q)}},$$

uniformly for $k = \mathcal{O}\left((n-r)^{\frac{1}{2}+\varepsilon}\right)$ and $0 \le \varepsilon < \frac{1}{6}$. Moreover, the probabilities are sub-exponentially small for k larger.

□

In the next step we observe the probability for any candidate to get hired, i.e. $\mathbb{P}(X_n = 1)$. In contrast to hiring above the m-th best this quantity is strongly dependent on the random variables $X_1, \dots, X_{n-1}$, as for the possible ranks for a hiring we have $\mathcal{R}(\sigma) = k$, where $\sigma \in \mathcal{S}_{n-1}$ denotes the permutation of the ranks amongst the first $n-1$ candidates and the threshold holds rank $n-k$ (i.e. the k-th highest rank among *all* already interviewed applicants). On the other hand hiring above the m-th best satisfies $\mathcal{R}(\sigma) = m$ for all permutations σ.

Remark 4.1.7. According to Theorem 1.1.9 each pragmatic hiring strategy recruits the $\mathcal{R}(\sigma)$ very best candidates until step $|\sigma| = n-1$. Hence, assuming the threshold candidate holds the ℓ-th highest relative rank among all hired candidates, it follows that he even holds the ℓ-th highest relative rank among all applicants seen so far. Thus, the conditioned probability for a hiring of candidate X_n is given by $\mathbb{P}(X_n = 1 \mid \tau = n-\ell) = \frac{\ell}{n}$, where τ denotes the rank of the threshold candidate.

■

Proposition 4.1.8. Consider the random variable X_n which indicates a hiring of the n-th candidate. Then for the probability of recruiting the n-th candidate the following holds:

$$\mathbb{P}(X_n = 1) = \begin{cases} 1, & n = 1, \\ \sum\limits_{\ell=1}^{n-1} \frac{(2\ell-1)n + (-3\ell+2)\ell}{(n-\ell)^2} \frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}}, & \text{else.} \end{cases}$$

Furthermore, the asymptotic expansion, for $n \to \infty$, is given by

$$\mathbb{P}(X_n = 1) = \frac{2\sqrt{\pi}}{\sqrt{n}} + \mathcal{O}\left(\frac{1}{n}\right).$$

Proof. The case for $n = 1$ follows trivially by the definition of our strategy. For $n \geq 2$ the law of complete probability shows that

$$\mathbb{P}(X_n = 1) = \sum_{\ell=1}^{n} \mathbb{P}(X_n = 1 \mid \tau = \ell)\ \mathbb{P}(\tau = \ell) = \sum_{\ell=1}^{n} \frac{\ell}{n} \left(a_{n-1,\ell}^{[1]} + a_{n-1,\ell}^{[2]}\right),$$

where τ denotes the threshold candidate after $n-1$ interviews. Applying (4.16) as well as (4.17) to the formula above, together with some careful algebraic manipulations yield in

$$\mathbb{P}(X_n = 1) = \sum_{\ell=1}^{n-1} \frac{\ell}{n} \left(\frac{\binom{n-1-\ell}{\ell-1}}{\binom{n-1}{\ell}} + \frac{\binom{n-1-\ell}{\ell-2}}{\binom{n-1}{\ell-1}} \right) = \sum_{\ell=1}^{n-1} \frac{(2\ell-1)n + (-3\ell+2)\ell}{(n-\ell)^2} \frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}}.$$

When taking the asymptotic expansion of the numbers $a_{n,\ell}^{[1]}$ and $a_{n,\ell}^{[2]}$ into account we get, similarly to the computations for the expectation of h_n, the following formula:

$$\begin{aligned}\mathbb{P}(X_n = 1) &= \sum_{\ell=1}^{n^{\frac{1}{2}}} \frac{2\ell^2}{n^2} e^{-\frac{\ell^2}{n}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\frac{\ell}{n}\right) + \mathcal{O}\left(\frac{\ell^3}{n^2}\right)\right) \\ &= \frac{2}{\sqrt{n}} \int_0^\infty x^2 e^{-x^2} dx \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) = \frac{2\sqrt{\pi}}{\sqrt{n}} + \mathcal{O}\left(\frac{1}{n}\right).\end{aligned}$$

□

Unfortunately we were not able to find a closed form for the exact probabilities of X_n either, which is surprising as the probabilities of h_n have a very simple representation.

We have already computed the probabilities for the (unconditioned and conditioned) size of the hiring set and the probability for hiring a single candidate until now. So the question after the probability for hiring a candidate holding a specific score r comes up. Therefore we define the numbers $a_{n,\ell,r}^{[i]}$, $i = 1, 2$ which give the probabilities that in a sequence of n candidates the threshold candidate has the ℓ-th highest rank, an odd or even number, respectively, has been recruited and that the candidate with the $(\ell + r)$-th largest relative rank amongst all n candidates is getting hired. Note that all candidates with a score equal or better than the threshold candidate get hired for sure. As an immediate consequence the probability for hiring a candidate with relative rank r, which we denote by $p_{n,r}$ is given by the following recurrence:

$$p_{n,r} = \sum_{\ell=1}^{n-r} \left(a_{n,\ell,n-\ell-r+1}^{[1]} + a_{n,\ell,n-\ell-r+1}^{[2]}\right) + \sum_{\ell=n-r+1}^{n} \left(a_{n,\ell}^{[1]} + a_{n,\ell}^{[2]}\right). \tag{4.23}$$

We only need to compute the values for $a_{n,\ell,r}^{[i]}$, $i = 1, 2$ and as a consequence we have a representation for $p_{n,r}$. We will only sketch these considerations as the approach is very similar to that one for computing the numbers $a^{[i]}_{n,\ell}$, $i = 1,2$ which we considered in the beginning of this section. A more detailed computations can be found in [15, p.23ff]. First of all it is suitable to modify the automaton given in Figure 4.1 which then leads to the automaton

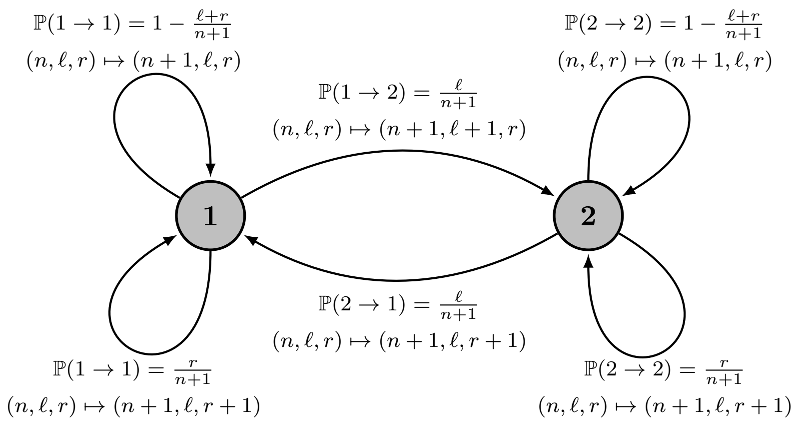

Figure 4.2: The modified automaton describing the relation and transition between the variables $a^{[1]}_{n,\ell,r}$ and $a^{[2]}_{n,\ell,r}$.

Considering Figure 4.2 leads us to the system of recursions

$$\begin{aligned} a^{[1]}_{n,\ell,r} &= \frac{\ell}{n}\, a^{[2]}_{n-1,\ell,r-1} + \frac{r-1}{n}\, a^{[1]}_{n-1,\ell,r-1} + \left(1 - \frac{\ell+r}{n}\right) a^{[1]}_{n-1,\ell,r}, \\ a^{[2]}_{n,\ell,r} &= \frac{\ell-1}{n}\, a^{[1]}_{n-1,\ell-1,r} + \frac{r-1}{n}\, a^{[2]}_{n-1,\ell,r-1} + \left(1 - \frac{\ell+r}{n}\right) a^{[2]}_{n-1,\ell,r}, \end{aligned}$$

with the initial conditions $a^{[i]}_{n,\ell,0} = a^{[i]}_{n,\ell}$, $i = 1,2$. Similar to the computation of $a^{[i]}_{n,\ell}$ the normalization

$$b^{[i]}_{n,\ell,r} = \frac{n!}{(\ell-1)!\,(r-1)!\,(n-\ell-r)!}\, a^{[i]}_{n,\ell,r}, \quad i = 1,2,$$

and uncoupling the recursions leads to the PDE

$$v(1-z-zv)B^{[1]}_v(z,u,v) - \frac{z^2u^2v}{1-z-zv} B^{[1]}_u(z,u,v) - \left(1 - z + \frac{z^2uv}{1-z-zv}\right) B^{[1]}(z,u,v) = 0,$$

where $B^{[1]}(z,u,v) = \sum_{n=1}^{\infty}\sum_{\ell=1}^{n}\sum_{r=1}^{n-\ell} b^{[1]}_{n,\ell,r} z^n u^\ell v^r$ denotes the multivariate generating function of the numbers $b^{[1]}_{n,\ell,r}$. Introducing $\widehat{B}(z,u,w) := B^{[1]}\left(z,u,\frac{(1-z)w}{1+zw}\right)$ and $\widehat{b}_\ell(z,w) := [u^\ell]\widehat{B}(z,u,w)$ yields in the recursion

$$w\,\frac{\partial \widehat{b}_\ell}{\partial w}(z,w) - \widehat{b}_\ell(z,w) = \frac{\ell z^2 w}{1-z}\,\widehat{b}_{\ell-1}(z,w).$$

For solving this differential equation we introduce the numbers

$$\overline{b}_\ell(z,w) := \frac{\widehat{b}_\ell(z,w)}{\ell!\,\left(\frac{z^2}{1-z}\right)^\ell},$$

and the corresponding generating function $\overline{B}(z,u,w) := \sum_{\ell=1}^{\infty} \overline{b}_\ell(z,w)u^\ell$, which leads to the differential equation

$$w\overline{B}_w(z,u,w) - (1+uw)\overline{B}(z,u,w) = 0.$$

The solution for this differential equation is given by

$$\overline{B}(z,u,w) = \left[\sum_{\ell=2}^{\infty}\left(\frac{1}{(\ell-2)!}(1-z)^{\ell+1}u^\ell \sum_{n=\ell}^{\infty}\binom{n-\ell+1}{\ell} z^{n-2\ell}\right)\right] we^{uw}.$$

Transforming this solution backwards gives us a representation for the numbers $b^{[1]}_{n,\ell,r}$ which is given by

$$\frac{b^{[1]}_{n,\ell,r}}{\ell!} = [z^n u^\ell v^r]\overline{B}\left(z, \frac{uz^2}{1-z}, \frac{v}{1-z-zv}\right) = [z^n u^\ell v^r] \frac{u^2 z^3 v}{(1-z)^2(1-z-zv)}\, e^{\frac{uz^2(1-v)}{1-z-zv}}. \tag{4.24}$$

The definition of the numbers $b^{[1]}_{n,\ell,r}$ together with extracting coefficients from (4.24) gives us

$$a^{[1]}_{n,\ell,r} = \frac{\ell-1}{n\binom{n-1}{\ell}\binom{n-\ell-1}{r-1}} \sum_{k=0}^{\ell-2}\binom{\ell-2}{k}\binom{r+k-1}{\ell-2}\binom{n-\ell+1}{r+k+1}, \tag{4.25}$$

$$a^{[2]}_{n,\ell,r} = \frac{\ell-1}{n\binom{n-1}{\ell-1}\binom{n-\ell}{r}} \sum_{k=0}^{\ell-2}\binom{\ell-2}{k}\binom{r+k-1}{\ell-3}\binom{n-\ell+1}{r+k+1}. \tag{4.26}$$

With this representation for the numbers $a^{[1]}_{n,\ell,r}$ and $a^{[2]}_{n,\ell,r}$ we get the following theorem.

Theorem 4.1.9. Let $p_{n,r}$ denote the probability that the candidate holding the relative rank r in a sequence of n candidates is getting hired. Then for the exact values of $p_{n,r}$, having $1 \le r \le n$, it holds that

$$\begin{aligned} p_{n,r} = \sum_{\ell=1}^{n-r}\Bigg[&\frac{\ell-1}{n\binom{n-1}{\ell}\binom{n-\ell-1}{n-\ell-r}} \sum_{k=0}^{\ell-2}\binom{\ell-2}{k}\binom{n-\ell-r+k}{\ell-2}\binom{n-\ell+1}{n-\ell-r+k+2} \\ &+ \frac{\ell-1}{n\binom{n-1}{\ell-1}\binom{n-\ell}{n-\ell-r+1}} \sum_{k=0}^{\ell-2}\binom{\ell-2}{k}\binom{n-\ell-r+k}{\ell-3}\binom{n-\ell+1}{n-\ell-r+k+2}\Bigg] \\ &+ \sum_{\ell=n-r+1}^{n}\left[\frac{\binom{n-\ell}{\ell-1}}{\binom{n}{\ell}} + \frac{\binom{n-\ell}{\ell-2}}{\binom{n}{\ell-1}}\right]. \end{aligned}$$

Proof. The claim is a direct consequence of (4.23) together with (4.25), (4.26), (4.16) and (4.17).

□

As we already have some results about the behavior of the random variables X_n, and the required quality an applicant needs for a hiring, we now observe the other hiring parameters to get a deeper insight into this strategy. We start with the waiting time until N candidates get recruited. As we do not have the parameter m, like we had with hiring above the m-th best, we cannot use the results of the waiting time for this aim, but the results give a outline about how many interviews probably have to be made.

Theorem 4.1.10 (Helmi and Panholzer, [15, Theorem 2]). For the exact probability of the waiting time W_N until N candidates get hired it holds that

$$\mathbb{P}(W_N = k) = \begin{cases} 0, & k < N, \\ \delta_{k,1}, & N = 1, \\ \frac{\ell}{N} \frac{\binom{k-1-\ell}{\ell-2}}{\binom{k-1}{\ell-1}}, & N = 2\ell - 1 \geq 3,\ k \geq N, \\ \frac{\ell}{N} \frac{\binom{k-1-\ell}{\ell-1}}{\binom{k-1}{\ell}}, & N = 2\ell \geq 2,\ k \geq N. \end{cases}$$

Asymptotically, for $N \to \infty$, W_N is distributed as follows: $\frac{W_N}{N^2} \xrightarrow{(d)} W$, where W has the density function

$$g(x) = \frac{1}{4x^2} e^{-\frac{1}{4x}}, \quad x > 0.$$

Proof. Again the cases of $N = 1$ as well as for $k < N$ are obvious by definition of the strategy. For the other cases let us consider Remark 2.1.6 together with Remark 4.1.7 which directly leads to

$$\begin{aligned} \mathbb{P}(W_N = k) &= \mathbb{P}(h_{k-1} = N-1)\ \mathbb{P}(X_k = 1 \mid h_{k-1} = N-1) \\ &= \begin{cases} a_{k-1,\ell}^{[2]}\ \frac{\ell}{k}, & N = 2\ell - 1 \geq 3, \\ a_{k-1,\ell}^{[1]}\ \frac{\ell}{k}, & N = 2\ell \geq 2. \end{cases} \end{aligned} \tag{4.27}$$

Plugging the exact results for $a_{n,\ell}^{[1]}$ and $a_{n,\ell}^{[2]}$ into (4.27) shows the claim. When considering the asymptotic results for the two auxiliary variables given in (4.18) - (4.19) we obtain the following result:

$$\mathbb{P}(W_N = k) = \frac{N^2}{4k^2} e^{-\frac{N^2}{4k}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{N}\right) + \mathcal{O}\left(\frac{N}{k}\right) + \mathcal{O}\left(\frac{N^3}{k^2}\right)\right),$$

when $N \to \infty$ and $k \gg n^{\frac{3}{2}}$. Substituting $k = xN^2,\ x > 0$ finally leads to

$$\mathbb{P}(W_N = k) = \frac{1}{N^2} \frac{1}{4x^2} e^{-\frac{1}{4x}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{N}\right)\right),$$

which shows the claim.

□

The next statements give information about how the hiring process will perform at the end of the sequence and how the requested ranks for a hiring will behave in time.

Theorem 4.1.11 (Helmi and Panholzer, [15, Theorem 3]). Consider the index of the last hired candidate L_n in a sequence of length n. Then the exact distribution of L_n when applying hiring above the median is given as follows:

$$\mathbb{P}(L_n = i) = \begin{cases} \frac{1}{n}, & i = 1, \\ \sum_{\ell=1}^{i-1} \frac{\binom{i-1-\ell}{\ell-2}}{\binom{n}{\ell}} + \sum_{\ell=1}^{i} \frac{\ell-1}{\ell} \frac{\binom{i-\ell}{\ell-2}}{\binom{n}{\ell}}, & 2 \leq i \leq n, \\ 0, & \text{else.} \end{cases}$$

Suitably normalized, L_n converges in distribution to a continuous random variable L, i.e. $\frac{n-L_n}{\sqrt{n}} \xrightarrow{(d)} L$, which is characterized by its density function $f(x)$ given by

$$f(x) = 2\int_0^\infty t^2 e^{-t(x+t)} dt, \quad x > 0.$$

Proof. The case of $L_n = 1$ appears if and only if exactly 1 candidate is hired, i.e. the first candidate has rank n. Thus, it is true that $\mathbb{P}(L_n = 1) = \frac{1}{n}$. The case of $L_n = i$, where it holds that $i > n$ can not appear as we hire no future candidates, which shows the claim. The non-trivial case of $2 \leq i \leq n$ simply states that $X_i = 1$ and $X_j = 0$ for all $i < j \leq n$. As an immediate consequence we have

$$\mathbb{P}(L_n = i) = \mathbb{P}(X_i = 1) \prod_{j=i+1}^{n} \mathbb{P}(X_j = 0 \mid X_i = 1, X_{i+1} = 0, X_{i+2} = 0, \ldots, X_{j-1} = 0). \quad (4.28)$$

For computing (4.28) we have to do a careful distinction whether we already recruited an odd number of candidates before X_i or an even number. If we recruited an even number of candidates then ℓ will increase with the hiring of X_i, which does not happen in the other case. So we must not simply plug the result from Proposition 4.1.8 into (4.28) but we have to do this separately for the even and the odd case. Thus, considering Remark 4.1.7 and the fact mentioned above, we have

$$\begin{aligned} \mathbb{P}(L_n = i) &= \left(\sum_{\ell=1}^{i-1} \frac{\binom{i-1-\ell}{\ell-2}}{\binom{i-1}{\ell-1}} \frac{\ell}{i} + \sum_{\ell=1}^{i} \frac{\binom{i-\ell}{\ell-2}}{\binom{i-1}{\ell-1}} \frac{\ell-1}{i} \right) \cdot \prod_{j=i+1}^{n} \left(1 - \frac{\ell}{i}\right) \\ &= \sum_{\ell=1}^{i-1} \frac{\binom{i-1-\ell}{\ell-2}}{\binom{n}{\ell}} + \sum_{\ell=1}^{i} \frac{\ell-1}{\ell} \frac{\binom{i-\ell}{\ell-2}}{\binom{n}{\ell}}. \end{aligned}$$

In the next step we determine the limit distribution of L_n. Therefore we use the asymptotic expansion

$$\frac{\binom{n-k-1-\ell}{\ell-2}}{\binom{n}{\ell}} = \frac{\ell^2}{n^2} e^{-\frac{\ell(k+\ell)}{n}} \left(\mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\frac{k+\ell}{n}\right) + \mathcal{O}\left(\frac{(k+\ell)^3}{n^2}\right) \right), \quad (4.29)$$

which holds uniformly for $k, \ell = \mathcal{O}(n^{\frac{1}{2}+\varepsilon})$ and can be shown similarly to Lemma 1.3.5. Moreover, this expression is sub-exponentially small for $\ell \geq n^{\frac{1}{2}+\varepsilon}$ and arbitrary $k \in \mathbb{N}$. Setting

$i = n - k$ and plugging (4.29) into the exact values for $\mathbb{P}(L_n = i)$ yields the asymptotic result for $k = \mathcal{O}(n^{\frac{1}{2}+\varepsilon})$ given by

$$\mathbb{P}(L_n = n-k) = 2\sum_{\ell=1}^{\infty} \frac{\ell^2}{n^2} e^{-\frac{\ell(k+\ell)}{n}} \left(\mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\frac{k+\ell}{n}\right) + \mathcal{O}\left(\frac{(k+\ell)^3}{n^2}\right)\right). \tag{4.30}$$

The claim follows from interpreting the sum in (4.30) as a Riemann sum together with substituting $k = x\sqrt{n}$ which results in the integral representation

$$\mathbb{P}(L_n = n-k) \sim \frac{2}{\sqrt{n}} \int_0^\infty t^2 e^{-t(x+t)} dt, \tag{4.31}$$

having $x > 0$. Thus, it is true that $\frac{n-L_n}{\sqrt{n}} \xrightarrow{(d)} L$, where L has the density function given in (4.31).

□

We also state the expectation of L_n which is probably a better indicator about the last hiring. Especially the asymptotic expansion shows that there will appear hirings throughout the end of the sequence of candidates.

Corollary 4.1.12. The expectation of the index of the last hiring L_n is given by

$$\mathbb{E}(L_n) = \sum_{\ell=1}^{n} \frac{(2\ell^2 - 3\ell + 1)n^2 - (3\ell^3 - 11\ell^2 + 10\ell - 3)n - 6\ell^3 + 12\ell^2 - 8\ell + 2}{\ell^2(n-\ell+1)} \frac{\binom{n-\ell+1}{\ell-1}}{\binom{n}{\ell}}.$$

Asymptotically for $n \to \infty$ we have the expansion

$$\mathbb{E}(L_n) = n - \sqrt{\pi n} + \mathcal{O}(\log n).$$

Proof. The explicit formula can be computed straightforwardly by

$$\begin{aligned}
\mathbb{E}(L_n) &= \sum_{i=1}^{n} i\, \mathbb{P}(L_n = i) = \frac{1}{n} + \sum_{i=2}^{n} i \left(\sum_{\ell=1}^{i-1} \frac{\binom{i-1-\ell}{\ell-2}}{\binom{n}{\ell}} + \sum_{\ell=1}^{i} \frac{\ell-1}{\ell} \frac{\binom{i-\ell}{\ell-2}}{\binom{n}{\ell}} \right) \\
&= \sum_{\ell=1}^{n-1} \frac{1}{\binom{n}{\ell}} \sum_{i=\ell+1}^{n} i \binom{i-1-\ell}{\ell-2} + \sum_{\ell=1}^{n} \frac{\ell-1}{\ell\binom{n}{\ell}} \sum_{i=\ell}^{n} i \binom{i-\ell}{\ell-2} \\
&= \sum_{\ell=1}^{n-1} \frac{1}{\binom{n}{\ell}} \underbrace{\sum_{i=0}^{n-\ell-1} (i+\ell+1) \binom{i}{\ell-2}}_{=:A} + \sum_{\ell=1}^{n} \frac{\ell-1}{\ell\binom{n}{\ell}} \underbrace{\sum_{i=0}^{n-\ell} (i+\ell) \binom{i}{\ell-2}}_{=:B}.
\end{aligned} \tag{4.32}$$

Using some simple algebraic manipulations and using the fact that

$$\sum_{i=0}^{n} \binom{i}{\ell} = \binom{n+1}{\ell+1},$$

we can compute the inner sums by

$$\begin{aligned} A &= \sum_{i=0}^{n-\ell-1} (i+\ell+1)\binom{i}{\ell-2} = (\ell-1)\sum_{i=0}^{n-\ell-1}\binom{i+1}{\ell-1} + \ell\sum_{i=0}^{n-\ell-1}\binom{i}{\ell-2} \\ &= (\ell-1)\left[\binom{n-\ell+1}{\ell} - \binom{0}{\ell-1}\right] + \ell\binom{n-\ell}{\ell-1} = (\ell-1)\binom{n-\ell+1}{\ell} + \ell\binom{n-\ell}{\ell-1}, \\ B &= (\ell-1)\binom{n-\ell-2}{\ell} + (\ell-1)\binom{n-\ell+1}{\ell-1}, \end{aligned}$$

where the representation for B can be received in an analog way as for A. Plugging the results for A and B into (4.32) together with some careful algebraic manipulation finally yields

$$\mathbb{E}(L_n) = \sum_{\ell=1}^{n} \frac{(2\ell^2-3\ell+1)n^2 - (3\ell^3-11\ell^2+10\ell-3)n - 6\ell^3+12\ell^2-8\ell+2}{\ell^2(n-\ell+1)} \frac{\binom{n-\ell+1}{\ell-1}}{\binom{n}{\ell}}. \tag{4.33}$$

Note that this representation results in 1 for the trivial case of $n=1$.

In order to get the asymptotic expansion we apply some modifications to the sum occurring in (4.33) and get

$$\mathbb{E}(L_n) = \sum_{\ell=1}^{n} \frac{(2\ell^2-3\ell+1)n^2 - (3\ell^3-11\ell^2+10\ell-3)n - 6\ell^3+12\ell^2-8\ell+2}{\ell(n-2\ell+2)(n-2\ell+2)} c(n,\ell),$$

where $c(n,\ell) := \frac{\binom{n-\ell}{\ell}}{\binom{n}{\ell}}$ is defined like in Corollary 1.3.5. Applying polynomial division to the fraction occurring in the sum together with the asymptotic expansion for $c(n,\ell)$ stated in 1.3.5 then shows that

$$\mathbb{E}(L_n) = \sum_{\ell=1}^{n} e^{-\frac{\ell^2}{n}}\left(2\ell-3+\frac{1}{\ell}+\frac{5\ell^2}{n}-\frac{2\ell^4}{n^2}\right)\cdot\left(1+\mathcal{O}\left(\frac{1}{\ell^2}\right)+\mathcal{O}\left(\frac{\ell^6}{n^4}\right)\right).$$

Applying Mellin transformation (see Section 1.5) leads us to the integral representation

$$\begin{aligned} \sum_{\ell=1}^{\infty} \ell^k e^{-\frac{\ell^2}{n}} &= \sum_{\ell=1}^{\infty}\left(\frac{\ell^k}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(s)\frac{n^s}{\ell^{2s}}ds\right) = \int_{c-i\infty}^{c+i\infty}\Gamma(s)n^s\cdot\left(\sum_{\ell=1}^{\infty}\ell^{k-2s}\right)ds \\ &= \int_{c-i\infty}^{c+i\infty}\Gamma(s)n^s\zeta(2s-k)ds, \end{aligned}$$

for fixed $k \geq -1$ and $c > \frac{k+1}{2}$ due to the fundamental strip of $\zeta(s)$. Note that we may change the order of integration and summation as both, the integral and the sum, converge in the fundamental strip. The statement follows from shifting the integral to the left and collecting the residues of the integrand. The significant singularity lies at $s = \frac{k+1}{2}$. This pole of the zeta function generates a residuum for the whole integrand, where we have to compute this separately for $k=-1$, which implies $s=0$, as $\Gamma(s)$ has a pole in this case as well.

- *Case 1:* $k=-1$ We estimate the sum directly in this case and get

$$\sum_{\ell}^{\infty}\frac{1}{\ell}e^{1-\frac{\ell^2}{n}} \sim \sum_{\ell=1}^{n^{\frac{1}{2}+\varepsilon}}\frac{1}{\ell}e^{-\frac{\ell^2}{n}} \leq \sum_{\ell=1}^{n^{\frac{1}{2}+\varepsilon}}\frac{1}{\ell} = \mathcal{O}(\log n),$$

  where the summands are sub-exponential small for $\ell \geq n^{\frac{1}{2}+\varepsilon}$.

- *Case 2:* $k \geq 0$ For the residuum at $s = \frac{k+1}{2}$ we have

$$\begin{aligned}\operatorname{Res}\left(\Gamma(s)n^s\zeta(2s-k), s = \frac{k+1}{2}\right) &= \operatorname{Res}\left(\Gamma(s)n^s\zeta\left(1+2\left(s-\frac{k+1}{2}\right)\right), s = \frac{k+1}{2}\right)\\ &= \operatorname{Res}\left(\frac{\Gamma(s)n^s}{1+2\left(s-\frac{k+1}{2}\right)}, s = \frac{k+1}{2}\right)\\ &= \Gamma\left(\frac{k+1}{2}\right)\frac{n^{\frac{k+1}{2}}}{2}.\end{aligned}$$

Thus, for the asymptotic expansion of the sum it holds that

$$\sum_{\ell=1}^{\infty} \ell^k e^{-\frac{\ell^2}{n}} = \Gamma\left(\frac{k+1}{2}\right)\frac{n^{\frac{k+1}{2}}}{2} + \mathcal{O}(1),$$

where the $\mathcal{O}$-term results from the next singularity at $s = 0$. Detailed information about this technique may be found in [6].

Merging all these results finally leads to

$$\begin{aligned}\mathbb{E}(L_n) &= n - \sqrt{n}\left(\frac{3}{2}\Gamma\left(\frac{1}{2}\right) - \frac{5}{2}\Gamma\left(\frac{3}{2}\right) + \Gamma\left(\frac{5}{2}\right)\right) + \mathcal{O}(\log n) + \mathcal{O}(1)\\ &= n - \sqrt{\pi n} + \mathcal{O}(\log n),\end{aligned}$$

where the computation of the values of $\Gamma(s)$ can be found in [7, p.744].

□

Theorem 4.1.13 (Helmi and Panholzer, [15, Theorem 6]). Consider the rank R_n of the last recruited candidate in a sequence of n candidates when using the hiring above the median strategy. Then the following holds for the exact distribution of R_n:

$$\mathbb{P}(R_n = r) = \sum_{\ell=n+1-r}^{n-1} \frac{\binom{n-\ell}{\ell}}{\binom{n}{\ell}} \frac{2n+3\ell+1}{(n-\ell)(n-2\ell+1)}.$$

Suitably normalized, R_n converges in distribution to a continuous random variable R, i.e. $\frac{n-R_n}{\sqrt{n}} \xrightarrow{(d)} R$, which is characterized by its density function $g(x)$ given by

$$g(x) = 2\int_0^{\infty} e^{-(x+t)^2} dt, \quad x > 0.$$

Proof. For convenience we consider the random variable $\widehat{R_n} := n+1-R_n$, i.e. $\widehat{R}_n = r$ denotes the r-th largest rank of the last recruited candidate. When considering the automaton given in Figure 4.1 and taking the index $L_n = m$ of the last hiring into account gives us the following formula

$$\begin{aligned}\mathbb{P}(\widehat{R}_n = r) = &\sum_{m=1}^{n-1}\sum_{\ell=r}^{m}\left[a_{m,\ell}^{[1]}\frac{1}{m+1}\prod_{j=m+1}^{n-1}\left(1-\frac{\ell+1}{j+1}\right) + a_{m,\ell}^{[2]}\frac{1}{m+1}\prod_{j=m+1}^{n-1}\left(1-\frac{\ell}{j+1}\right)\right]\\ &+ a_{1,1}^{[1]}\prod_{j=1}^{n-1}\left(1-\frac{1}{j+1}\right)\delta_{r,1}.\end{aligned}$$

The formula can be explained easily. First we again have to distinguish between the even and the odd case for h_n. When considering the position m of the last hiring, the possible ranks for the last hired candidate are all ranks higher than the rank ℓ of the threshold candidate at step m. Additionally we have to consider the case consisting of only one hiring, which only carries weight for $r = 1$. When plugging (4.16) and (4.17) into the formula above we get

$$\begin{aligned}
\mathbb{P}(\widehat{R}_n = r) &= \sum_{m=1}^{n-1} \sum_{\ell=r}^{m} \left[\frac{\binom{m-\ell}{\ell-1}}{\binom{m}{\ell}} \frac{(n-1-\ell)!\, m!}{(m-\ell)!\, n!} + \frac{\binom{m-\ell}{\ell-2}}{\binom{m}{\ell-1}} \frac{(n-\ell)!\, m!}{(m+1-\ell)!\, n!} \right] + \frac{\delta_{r,1}}{n} \\
&= \sum_{m=1}^{n-1} \sum_{\ell=r}^{m} \left[\frac{\binom{m-\ell}{\ell-1}}{\binom{n}{\ell+1}} \frac{1}{(\ell+1)!} + \frac{\binom{m-\ell}{\ell-2}}{\binom{n}{\ell}} \frac{1}{\ell} \right] + \frac{\delta_{r,1}}{n} \\
&= \sum_{\ell=r}^{n-1} \left[\frac{1}{(\ell+1)\binom{n}{\ell+1}} \sum_{m=\ell}^{n-1} \binom{m-\ell}{\ell-1} + \frac{1}{\ell\binom{n}{\ell}} \sum_{m=\ell}^{n-1} \binom{m-\ell}{\ell-2} \right] + \frac{\delta_{r,1}}{n} \\
&= \sum_{\ell=r}^{n-1} \left[\frac{\binom{n-\ell}{\ell}}{\binom{n}{\ell+1}} \frac{1}{\ell+1} + \frac{\binom{n-\ell}{\ell-1} - \delta_{\ell,1}}{\binom{n}{\ell}} \frac{1}{\ell} \right] + \frac{\delta_{r,1}}{n} = \sum_{\ell=r}^{n-1} \frac{\binom{n-\ell}{\ell}}{\binom{n}{\ell}} \frac{2n+3\ell+1}{(n-\ell)(n-2\ell+1)},
\end{aligned}$$

which shows the claim.

The asymptotic behavior can be determined in a same manner as for L_n. Therefore we again consider Corollary 1.3.5 and plug the expansion for the binomial coefficients into the exact distribution yielding

$$\mathbb{P}(\widehat{R}_n = r) = \sum_{\ell=r}^{n^{\frac{1}{2}+\varepsilon}} \frac{2}{n} e^{-\frac{\ell^2}{n}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\frac{\ell}{n}\right) + \mathcal{O}\left(\frac{\ell^3}{n^2}\right)\right).$$

Substituting $r = x\sqrt{n}$ and $t = \frac{\ell}{\sqrt{n}}$ as well as interpreting the sum as a Riemann sum finally shows that $\frac{n-R_n}{\sqrt{n}} \xrightarrow{(d)} R$, where R has the density function

$$g(x) = 2\int_x^{\infty} e^{-t^2} dt = 2\int_0^{\infty} e^{-(x+t)^2} dt, \quad x > 0.$$

□

Corollary 4.1.14. The moments $\mathbb{E}(R^s)$ of the limit distribution R are for $s > -1$ given by

$$\mathbb{E}(R^s) = \frac{2}{s+1}\Gamma\left(\frac{s}{2}+1\right).$$

Proof.

$$\begin{aligned}
\mathbb{E}(R^s) &= \int_0^{\infty} x^s g(x) dx = 2\int_0^{\infty} x^s \int_0^{\infty} e^{-t^2} dt dx = 2\int_0^{\infty} e^{-t^2} \int_0^{t} x^s dx dt \\
&= 2\int_0^{\infty} e^{-t^2} \frac{t^{s+1}}{s+1} dt = \frac{2}{s+1}\int_0^{\infty} e^{-u} u^{\frac{s+1}{2}} \frac{du}{2\sqrt{u}} = \frac{2}{s+1}\Gamma\left(\frac{s}{2}+1\right),
\end{aligned}$$

where we substituted $u = t^2$.

□

For studying the parameter M_n which denotes the rank of the best discarded candidate in a sequence with hiring above the median we make the same approach as we already did for this quantity when analyzing hiring above the m-th best. The only difference is that now we have to take the parity of h_n into account again. Therefore we introduce auxiliary variables $\widehat{a}^{[1]}_{n,\ell,q}$ and $\widehat{a}^{[2]}_{n,\ell,q}$ for $0 \le q \le n-\ell$ that denote the probabilities that for n interviewed candidates the threshold candidate holds the ℓ-th largest score and at least all of the $\ell+q$ best ranked candidates are hired. Like for hiring above the m-th best we then get

$$\mathbb{P}(M_n = r) = \sum_{\ell=1}^{n-r} \left(\widehat{a}^{[1]}_{n,\ell,n-\ell-r} - \widehat{a}^{[1]}_{n,\ell,n-\ell-r+1} \right) + \sum_{\ell=1}^{n-r} \left(\widehat{a}^{[2]}_{n,\ell,n-\ell-r} - \widehat{a}^{[2]}_{n,\ell,n-\ell-r+1} \right). \tag{4.34}$$

The automaton from Figure 4.1 helps us to find a recursion for the new auxiliary variables but we have to be careful about the transition probabilities when discarding a candidate. Therefore assume that at step $n-1$ the best $\ell+q$ candidates are hired. If we want this property to hold in step n too when discarding candidate X_n, he must not belong to the best $\ell+q$ candidates. Note that for solely discarding X_n (without respecting the investigated property) he must not belong to the best ℓ candidates. Thus, the recursions are given by

$$\begin{aligned} \widehat{a}^{[1]}_{n,\ell,q} &= \frac{\ell}{n}\, \widehat{a}^{[2]}_{n-1,\ell,q-1} + \left(1 - \frac{\ell+q}{n}\right) \widehat{a}^{[1]}_{n-1,\ell,q}, \\ \widehat{a}^{[2]}_{n,\ell,q} &= \frac{\ell-1}{n}\, \widehat{a}^{[1]}_{n-1,\ell-1,q} + \left(1 - \frac{\ell+q}{n}\right) \widehat{a}^{[2]}_{n-1,\ell,q}. \end{aligned}$$

For the starting values Remark 4.1.7 implies that $\widehat{a}^{[1]}_{n,\ell,0} = a^{[1]}_{n,\ell}$ and $\widehat{a}^{[2]}_{n,\ell,0} = a^{[2]}_{n,\ell}$. Similarly to the computations for h_n introducing

$$\widehat{b}^{[i]}_{n,\ell,q} := \frac{n!}{\ell!\;(n-q-\ell)!}\, a^{[i]}_{n,\ell,q}, \quad i = 1, 2$$

together with a generating function approach and uncoupling the recursions lead to the PDE

$$\widehat{B}^{[1]}_u(z,u,v) - \frac{(1-z)^2}{z^2u^2v}\widehat{B}^{[1]}(z,u,v) + \frac{z(1-z)}{(1-z-z^2u)^2} = 0.$$

As we did for hiring above the m-th best in Lemma 3.2.13 extracting the coefficients from this PDE leads to the recurrence

$$b^{[1]}_\ell(z,v) = \frac{z^2v}{(1-z)^2}(\ell-1)b^{[1]}_{\ell-1}(z,v) + \frac{z^3v}{(1-z)^3}(\ell-1)\left(\frac{z^2}{1-z}\right)^{\ell-2},$$

where we defined $b^{[1]}_\ell(z,v) := [u^\ell]B^{[1]}(z,u,v)$. Substituting

$$c_\ell(z,v) := \frac{b^{[1]}_\ell(z,v)(1-z)^{2\ell}}{(\ell-1)!\; z^{2\ell}v^\ell}$$

and solving the resulting recursion leads to

$$c_\ell(z,v) = \frac{1}{z}\sum_{j=2}^{\ell} \frac{1}{(j-2)!}\left(\frac{1-z}{v}\right)^{j-1}, \quad \ell \ge 0.$$

Finally resubstituting and extracting the coefficients from the solution for $b_\ell^{[1]}(z,v)$ yields in

$$\widehat{a}_{n,\ell,q}^{[1]} = \frac{\binom{\ell-1}{q}\binom{n-\ell+q}{\ell+q-1}}{\binom{n}{q}\binom{n-q}{\ell}}, \tag{4.35}$$

$$\widehat{a}_{n,\ell,q}^{[2]} = \frac{\binom{\ell-1}{q+1}\binom{n-\ell+q+1}{\ell+q-1}}{\binom{n}{q+1}\binom{n-q-1}{\ell-1}}, \tag{4.36}$$

where the range of q is given by $0 \le q \le n-\ell$. A more detailed execution for computing $\widehat{a}_{n,\ell,q}^{[1]}$ and $\widehat{a}_{n,\ell,q}^{[2]}$ can be found in [15, p.20ff]. With this result we are able to formulate the following theorem.

Theorem 4.1.15 (Helmi and Panholzer, [15, Theorem 7]). For the exact probabilities of the rank of the best discarded candidate M_n it holds that

$$\mathbb{P}(M_n = r) = \begin{cases} \frac{1}{\binom{n}{\lceil \frac{n}{2} \rceil}}, & r = 0, \\ \sum_{\ell=1}^{n-r} \frac{\binom{\ell-1}{n-\ell-r}\binom{2n-2\ell-r}{n-r-1}}{\binom{n-r}{\ell}\binom{n}{r}} \left(1 + \frac{(n-2\ell-r+1)(2n-2\ell-r+1)}{r(n-r)}\right) & \\ + \sum_{\ell=1}^{n-r} \frac{\binom{\ell-1}{n-\ell-r+1}\binom{2n-2\ell-r+1}{n-r-1}}{\binom{n-r}{\ell-1}\binom{n}{r}} \left(1 + \frac{(n-2\ell-r+2)(2n-2\ell-r+2)}{r(n-r)}\right), & \text{else.} \end{cases}$$

The limiting distribution of M_n for $n \to \infty$ is, after normalization, a Rayleigh distribution with parameter $\sigma = \frac{1}{\sqrt{2}}$, i.e.
$\frac{n-M_n}{\sqrt{n}} \xrightarrow{(d)} \widetilde{R}$, where $\widetilde{R}$ has the density function

$$\widetilde{f}(x) = 2xe^{-x^2}, \quad x > 0.$$

Proof. By definition we have $M_n = 0$ if and only if no candidate is discarded. Thus, it holds that $\mathbb{P}(M_n = 0) = \mathbb{P}(h_n = n) = \frac{1}{\binom{n}{\lceil n \rceil}}$. The result for $1 \le r \le n$ is a direct consequence of (4.34) together with (4.35) and (4.36).

For the asymptotic result we consider the random variable $\widehat{M_n} := n - M_n$. Applying the asymptotic expansion from Corollary 1.3.5 to the exact probabilities for $\widehat{M_n}$ leads to

$$\mathbb{P}(\widehat{M_n} = r) = \sum_{\ell=0}^{r^{\frac{1}{2}+\varepsilon}} \frac{4\ell}{n} e^{-\frac{\ell^2}{r} - \frac{r^2}{n}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\ell}\right) + \mathcal{O}\left(\frac{\ell}{r}\right) + \mathcal{O}\left(\frac{\ell r}{n}\right) + \mathcal{O}\left(\frac{r^3}{n^2}\right)\right).$$

Considering the sum as a Riemann sum and substituting $r = x\sqrt{n}$ and $\ell = t\sqrt{r}$ yields

$$\mathbb{P}(\widehat{M_n} = x) \sim \frac{4}{n} e^{-\frac{r^2}{n}} \sum_{\ell=0}^{r^{\frac{1}{2}+\varepsilon}} \ell e^{-\frac{\ell^2}{r}} \sim \frac{4x}{\sqrt{n}} e^{-x^2} \int_0^\infty t e^{-t^2} dt = \frac{2xe^{-x^2}}{\sqrt{n}}, \quad x > 0,$$

which shows the claim.

□

The last parameter we observe for hiring above the median without the replacement mechanism is the distance between the last two hirings Δ_n, which we do not expect to grow in a too fast way, due to the results for the index of the last hired candidate.

Theorem 4.1.16 (Helmi and Panholzer, [15, Theorem 4]). The exact probability of the distance between the last two hirings in a sequence of n candidates Δ_n when using hiring above the median is given by

$$\mathbb{P}(\Delta_n = d) = \begin{cases} \frac{1}{n}, & d = 0, \\ \sum\limits_{m=1}^{n-d-1} \sum\limits_{\ell=1}^{m} \frac{\binom{m-\ell}{\ell-1}}{\binom{n}{\ell+1}} \left(\frac{\ell}{m+d-\ell} + \frac{\ell(\ell-1)}{(\ell+1)(n-2\ell+2)} \right) + \frac{1}{n(n-1)}, & 1 \leq d \leq n-1. \end{cases}$$

Suitably normalized, Δ_n converges in distribution to a continuous random variable Δ, i.e. $\frac{\Delta_n}{\sqrt{n}} \xrightarrow{(d)} \Delta$, which is characterized by its density function $f(x)$ given by

$$f(x) = 2 \int_0^\infty t^2 e^{-t(x+t)} dt, \quad x > 0.$$

Proof. By definition we have $\Delta_n = 0$ if and only if only one candidate gets recruited. Thus, it follows directly $\mathbb{P}(\Delta_n = 0) = \frac{1}{n}$. For computing the exact probabilities when $d \geq 1$ we have to consider the positions of the last two hirings and $d-1$ discarded candidates between them. Like for L_n we have to respect the parity of the hiring set, i.e. a possible change of the threshold candidate. Therefore let m denote the position of the second last hiring. Due to the minimum number of d candidates after the second last hiring, we get $1 \leq m \leq n-d-1$ for the range of m. Thus, we have

$$\begin{aligned} \mathbb{P}(\Delta_n = d) = & \sum_{m=1}^{n-d-1} \sum_{\ell=1}^{m} a_{m,\ell}^{[1]} \frac{\ell}{m+1} \prod_{j=m+1}^{m+d-1} \left(1 - \frac{\ell+1}{j+1}\right) \frac{\ell+1}{m+d+1} \prod_{j=m+d+1}^{n-1} \left(1 - \frac{\ell+1}{j+1}\right) \\ & + \sum_{m=1}^{n-d-1} \sum_{\ell=1}^{m} a_{m,\ell}^{[2]} \frac{\ell}{m+1} \prod_{j=m+1}^{m+d-1} \left(1 - \frac{\ell}{j+1}\right) \frac{\ell}{m+d+1} \prod_{j=m+d+1}^{n-1} \left(1 - \frac{\ell+1}{j+1}\right) \\ & + a_{1,1}^{[1]} \prod_{j=1}^{d-1} \left(1 - \frac{1}{j+1}\right) \frac{1}{d+1} \prod_{j=d+1}^{n-1} \left(1 - \frac{2}{j+1}\right), \end{aligned} \tag{4.37}$$

where the last line represents the case that the second last hired candidate is the first candidate X_1. Plugging the results of (4.16) and (4.17) into (4.37) together with simplifying the products yields

$$\begin{aligned} \mathbb{P}(\Delta_n = d) = & \sum_{m=1}^{n-d-1} \sum_{\ell=1}^{m} \frac{\binom{m-\ell}{\ell-1}}{\binom{m}{\ell}} \frac{m!\,(n-1-\ell)!\,(\ell+1)\ell}{n!\,(m-\ell)!\,(m+d-\ell)} \\ & + \sum_{m=1}^{n-d-1} \sum_{\ell=1}^{m} \frac{\binom{m-\ell}{\ell-2}}{\binom{m}{\ell-1}} \frac{m!\,(n-\ell-1)!\,\ell^2}{n!\,(m+1-\ell)!} + \frac{1}{n(n-1)}. \end{aligned}$$

Finally merging the sums and applying some algebraic manipulations show the claim for the exact probabilities.

When determining the asymptotic behavior of Δ_n it is applicable to start with another representation for the exact probabilities given by

$$\mathbb{P}(\Delta_n = d) = \sum_{\ell=2}^{n-d-1} \frac{\ell}{\ell+1} \frac{\binom{n-d-\ell}{\ell-1}}{\binom{n}{\ell+1}} + \sum_{\ell=2}^{n-d-2} \frac{\ell}{\ell-1} \frac{\binom{n-d-\ell-1}{\ell-1}}{\binom{n}{\ell+1}} + R(n,\ell),$$

having

$$R(n,\ell) = \frac{2}{n(n-1)}(H_{n-2} - H_{d-1}) + \frac{1}{n(n-1)} - \sum_{m=3}^{n-d-1} \sum_{\ell=2}^{m-1} \frac{\ell d}{(m+d-\ell)(m-\ell)} \frac{\binom{m-\ell}{\ell-1}}{\binom{n}{\ell+1}},$$

which holds for $1 \leq d \leq n-2$ and $n \geq 3$ and can be obtained by an easy but a bit lengthy computation when changing the order of summation and applying some combinatorial identities. Let us first consider an asymptotic expansion of $R(n,\ell)$. When decomposing $R(n,\ell)$ into its summands it firstly holds that

$$\frac{2}{n(n-1)}(H_{n-2} - H_{d-1}) = \mathcal{O}\left(\frac{\log n}{n^2}\right), \qquad \frac{1}{n(n-1)} = \mathcal{O}\left(\frac{1}{n^2}\right),$$

and secondly

$$\sum_{m=3}^{n-d-1} \sum_{\ell=2}^{m-1} \frac{\ell d}{(m+d-\ell)(m-\ell)} \frac{\binom{m-\ell}{\ell-1}}{\binom{n}{\ell+1}} = \underbrace{\sum_{m=5}^{n-d-1} \sum_{\ell=3}^{m-2} \frac{\ell d}{(m+d-\ell)(m-\ell)} \frac{\binom{m-\ell}{\ell-1}}{\binom{n}{\ell+1}}}_{=:S} + \mathcal{O}\left(\frac{\log n}{n^2}\right),$$

where for the sum it is true that

$$\begin{aligned} S &\leq \sum_{m=5}^{n-d-1} \sum_{\ell=3}^{m-2} \frac{\ell d}{(\ell-1)(\ell-2)} \frac{\binom{m-\ell-2}{\ell-3}}{\binom{n}{\ell+1}} = \sum_{\ell=3}^{n-d-3} \frac{\ell d}{(\ell-1)(\ell-2)} \frac{\binom{n-d-\ell-2}{\ell-2}}{\binom{n}{\ell+1}} \\ &= \mathcal{O}\left(\sum_{\ell=3}^{n-d-3} \frac{\binom{n-d-\ell-1}{\ell-1}}{\binom{n}{\ell+1}} \frac{d}{n}\right) = \mathcal{O}\left(\frac{d}{n} \sum_{\ell=3}^{\infty} \frac{\ell^2}{n^2} e^{-\frac{\ell(d+\ell)}{n}}\right) = \mathcal{O}\left(\frac{d}{n^{\frac{3}{2}}}\right). \end{aligned}$$

Thus, the term $R(n,\ell)$ is asymptotically negligible and the main contribution comes from the other two sums. Using Corollary 1.3.5 and applying computations similar to the proof of Theorem 4.1.11 yield the asymptotic expansion

$$\mathbb{P}(\Delta_n = d) \sim \frac{2}{\sqrt{n}} \int_0^\infty t^2 e^{-t(x+t)} dt, \quad d = x\sqrt{n}, \quad x > 0.$$

Hence, for $n \to \infty$ we get $\frac{\Delta_n}{\sqrt{n}} \xrightarrow{(d)} \Delta$, where Δ has the density function stated in the claim. □

Last but not least let us investigate the behavior of hiring above the median together with the replacement mechanism. Therefore we make the same approach as we did for hiring above the

m-th best. We introduce the indicator variables Y_i that denotes the event of a replacement in step i. The following lemma is a statement about the probability of a replacement at some certain step i.

Lemma 4.1.17. When adding the replacement mechanism to the hiring above the median strategy the exact probability that a replacement appears when interviewing candidate X_i satisfies

$$\mathbb{P}(Y_i = 1) = \sum_{\ell=1}^{j-1} \left(\frac{\ell-1}{i} \frac{\binom{j-1-\ell}{\ell-1}}{\binom{j-1}{\ell}} + \frac{\ell-2}{i} \frac{\binom{j-1-\ell}{\ell-2}}{\binom{j-1}{\ell-1}} \right).$$

Proof. Let us first consider the conditioned probability of Y_i under $h_n^{[f]} = k$. One may easily prove the fact that

$$\mathbb{P}(Y_i = 1 \mid \tau_i = \ell) = \begin{cases} \frac{\ell-1}{i}, & k \equiv 1(2), \\ \frac{\ell-2}{i}, & k \equiv 0(2), \end{cases}$$

where τ_i denotes the event that the threshold candidate at step i has the ℓ-th highest rank. The stated probabilities are a direct consequence of the possible ranks for a replacement. The law of complete probability implies that

$$\mathbb{P}(Y_i = 1) = \sum_{\ell=1}^{i-1} \mathbb{P}(Y_i = 1 \mid \tau_i = \ell)\mathbb{P}(\tau_i = \ell) = \sum_{\ell=1}^{i-1} \left(a_{i-1,\ell}^{[1]} \frac{\ell-1}{i} + a_{j-1,\ell}^{[2]} \frac{\ell-2}{i} \right). \tag{4.38}$$

Plugging (4.16) and (4.17) into (4.38) shows the claim.

□

Theorem 4.1.18. Consider the number of replacements f_n that appear in a sequence of n candidates when applying the replacement mechanism together with hiring above the median. Then the expectation is asymptotically, for $n \to \infty$ given by

$$\mathbb{E}(f_n) = \sqrt{\pi n} + \mathcal{O}(\log n).$$

Proof. The expectation of f_n can be computed by using the linearity of the expectation and considering the results from Lemma 4.1.17. Thus, we have

$$\mathbb{E}(f_n) = \sum_{i=1}^{n} \mathbb{P}(Y_i = 1) = \sum_{i=1}^{n} \sum_{\ell=1}^{i-1} \left(\frac{\ell-1}{i} \frac{\binom{i-1-\ell}{\ell-1}}{\binom{i-1}{\ell}} + \frac{\ell-2}{i} \frac{\binom{i-1-\ell}{\ell-2}}{\binom{i-1}{\ell-1}} \right)$$

Considering the inner sum yields

$$\sum_{\ell=2}^{i-1} \left(\frac{\ell-1}{i} \frac{\binom{i-1-\ell}{\ell-1}}{\binom{i-1}{\ell}} + \frac{\ell-2}{i} \frac{\binom{i-1-\ell}{\ell-2}}{\binom{i-1}{\ell-1}} \right) = \frac{\sqrt{\pi}}{2\sqrt{i}} + \mathcal{O}\left(\frac{1}{i}\right),$$

where the result is a consequence of Corollary 1.3.5 together with interpreting the sum as a Riemann sum and computing the resulting integral. The claim follows by summation over i.

□

Other possible hiring strategies that are similar to hiring above the median can be obtained when taking another position of the threshold candidate instead of the median.

Remark 4.1.19. According to [15, 20] using the lower median as threshold candidate instead of the upper median. They called this strategy the $\frac{1}{2}$-percentile rule in [20]. This strategiy leads to significantly different results as for hiring above the median. The results for the lower median have even more complicated representations.

■

Remark 4.1.20. Another approach is hiring above the α-quantile, which was introduced by [1]. Here the threshold candidate has the $\lceil \alpha k \rceil$-th highest rank of the hiring set, where k denotes the number of recruited candidates. Hiring above the α-quantile is closely related to the p-percentile rule introduced in [20]. Helmi and Panholzer [14] were first able to determine the limit distribution for $p = \frac{1}{2}$. Later Gaither and Ward [9] were first to determine asymptotics of the expectation for $p \in (0, 1)$.

■

Chapter 5

Probabilistic hiring

A very important characteristic of all so far investigated strategies was the score of a threshold candidate whom we used for our decision about a hiring. In particular we observed how the threshold changes during the hiring process. Whilst hiring above the minimum never changed the threshold candidate, hiring above the maximum did this every time a hiring happened and for hiring above the median the threshold candidate changed after every second hiring.

In this chapter we will consider *probabilistic hiring strategies*. Probabilistic, referring to the change of the threshold candidate. Thus, we do not know for sure if the threshold candidate changes after a hiring or not. For all the other strategies we have always known what would happen after recruiting a candidate. Thus, we may centralize all the strategies we have considered until yet as *deterministic hiring strategies*. In fact we could imagine changing the threshold candidate as a function that acts in some way after a hiring (we could also interpret it as a Turing machine, or like the automaton in Figure 4.1).

We can not do this for the probabilistic strategies as we have no fixed rule for changing the threshold. One may say that we have a double uncertainty, namely if a hiring happens on the one hand and on the other hand if the threshold candidate changes. In Section 5.2 we will build up a connection between the deterministic strategies and the probabilistic approach.

5.1 Probabilistic change of the threshold

The strategy we will concretely observe acts like that, that the first candidate gets hired for sure and becomes the first threshold candidate. Further candidates get hired if and only if their score is better than the score of the candidate currently holding the threshold position. If a new candidate is hired there may appear two cases. Firstly, the threshold candidate does not change, or secondly the very next better ranked and already hired candidate becomes the new threshold candidate, which happens with a fixed probability $p \in [0, 1)$. We will consider the case $p = 1$ in Section 5.2 as the following results do not hold in that case. As usual the threshold candidate does not change if we discard a candidate in order to receive a pragmatic strategy. Let us describe this formally like we did for all the other strategies.

Strategy 5.1.1 (Probabilistic change of the threshold). For a sequence $X_1, X_2, \ldots, X_n$ of candidates and the permutation of their relative ranks $\sigma \in \mathcal{S}_n$ the probabilistic hiring strategy works in the following way:

- In our strategy the first candidate gets hired for sure and becomes the first threshold candidate. Thus, we have

$$X_1 = 1, \qquad \tau = \sigma_1(1).$$

- When interviewing candidate $X_i,\ i \geq 2$, we consider the hiring set of step $i-1$, which is given by $\mathcal{H}_{i-1}(\sigma) = \{c_1, c_2, \ldots, c_k\}$. Furthermore, assume that the hired candidates are ordered by their ranking, i.e. $\sigma_i(c_1) < \sigma_i(c_2) < \ldots < \sigma_i(c_k)$ and let $T_i := c_\ell,\ 1 \leq \ell \leq k$, denote the current threshold candidate. The hiring process acts in the following way, where $\tau_i = \sigma_i(c_\ell)$ denotes the threshold value:

$$X_i = \begin{cases} 1, & \sigma_i(i) > \tau_i, \\ 0, & \sigma_i(i) < \tau_i. \end{cases}$$

Moreover, the threshold changes applying the following (probabilistic) rule:

$$T_{i+1} = \begin{cases} c_\ell, & X_i = 0, \\ c_\ell, & X_i = 1 \wedge \mathcal{T} = 0, \\ \widetilde{c}_{\ell+1}, & X_i = 1 \wedge \mathcal{T} = 1, \end{cases}$$

where $\mathcal{T}$ denotes the random variable that indicates a change of the threshold, i.e. $\mathbb{P}(\mathcal{T} = 1) = p$ and $\mathbb{P}(\mathcal{T} = 0) = 1-p$. The candidate $\widetilde{c}_{\ell+1}$ is, analogous to step $i-1$, the candidate with the $(\ell+1)$-th smallest relative rank of all hired candidates after step i. Thus, it holds that

$$\widetilde{c}_{\ell+1} = \begin{cases} i, & \sigma_i(c_\ell) < \sigma_i(i) < \sigma_i(c_{\ell+1}), \\ c_{\ell+1}, & \sigma_i(i) > \sigma_i(c_{\ell+1}). \end{cases}$$

Note that τ always denotes the relative rank of the threshold candidate in the current step. If we refer to a specific step i we write τ_i. Obviously it holds that $\tau_i = \tau_{i-1} + 1$ if a candidate is discarded.

■

We can also define the strategy by building up the hiring set from step to step like we did for the other strategies. This would result in the following rules, where we use the notation from the definition of the strategy above.

- $\mathcal{H}_1(\sigma) = \{1\}$.
- $\mathcal{H}_i(\sigma) = \begin{cases} H_{i-1}, & \sigma_i(i) < \tau_i, \\ H_{i-1} \cup \{i\}, & \sigma_i(i) > \tau_i, \end{cases} \qquad i \geq 2.$

When analyzing this strategy it turns out that the approach is, by reason of the parameter p quite involved and we were not even able to determine the exact probabilities of the parameter h_n. However, we were able to determine its distribution by the integer moments.

For our approach we introduce the auxiliary variables $a_{n,k,d}$ that describe the probability that after n interviewed candidates, the threshold candidate has the k-th highest relative rank and the hiring has the size d, i.e. $a_{n,k,d} := \mathbb{P}(h_n = d \wedge \tau_n = k)$. Obviously it holds that $k \leq d$ as the threshold candidate may only change after a hiring. Furthermore, it holds that $a_{1,1,1} = 1$. We can easily gain a recursion for the numbers $a_{n,k,d}$ given by

$$a_{n,k,d} = \frac{n-k}{n} a_{n-1,k,d} + p\frac{k}{n} a_{n-1,k,d-1} + (1-p)\frac{k-1}{n} a_{n-1,k-1,d-1}, \tag{5.1}$$

for $n \geq 2$, $1 \leq d \leq n$, as well as $1 \leq k \leq d$. The recursion reflects the three cases we have had in Strategy 5.1.1 for the change of the threshold. Either the candidate is discarded or the candidate is recruited and the threshold changes or the candidate is recruited without a change of the threshold. Note that k does not change if a candidate is recruited together with an increase of the threshold.

For our purposes it is good to consider the sequence $a_{n,k}(v) := \sum_{d=1}^{n} a_{n,k,d} v^d$. With that definition (5.1) translates to

$$a_{n,k}(v) = \frac{n-k}{n}\, a_{n-1,k}(v) + pv\, \frac{k}{n}\, a_{n-1,k}(v) + (1-p)v\, \frac{k-1}{n}\, a_{n-1,k-1}(v),$$

for $n \geq 2$ and $1 \leq k \leq n$. As a starting value for this functional recursion we have $a_{1,1}(v) = v$. Multiplying the recursion by n leads to

$$n a_{n,k}(v) = (n-k) a_{n-1,k}(v) + pvk a_{n-1,k}(v) + (1-p)v(k-1) a_{n-1,k-1}(v),$$

and introducing the multivariate generating function $F(z,u,v) = \sum_{n=1}^{\infty} \sum_{k=1}^{n} a_{n,k}(v) z^n u^k$ yields the PDE

$$\begin{aligned} F_z(z,u,v) - vu = {} & zF_z(z,u,v) + F(z,u,v) - uF_u(z,u,v) + pvuF_u(z,u,v) \\ & + (1-p)vu^2 F_u(z,u,v), \end{aligned}$$

which we can write in the canonical form

$$(1-z)F_z(z,u,v) + u(1 - v(p + (1-p)u))F_u(z,u,v) - F(z,u,v) - vu = 0. \tag{5.2}$$

For convenience we will omit the arguments in the following, e.g. $F(z,u,v) \equiv F$. The PDE (5.2) can be solved using the method of characteristics. For a detailed description of this method see [22]. The method of characteristics gives us the form $f(z,u,F) = 0$ having

$$(1-u)f_z + u(1 - v(p + (1-p)u))f_u + (F + vu)f_F = 0$$

Thus, we receive a system of ODE given by

$$\begin{aligned} \dot{z} &= 1 - z, \\ \dot{u} &= u(1 - v(p + (1-p)u)), \\ \dot{F} &= F + vu. \end{aligned}$$

Considering the derivation

$$\frac{dz}{du} = \frac{1-u}{u(1 - v(p + (1-p)u))},$$

and separation of the variables followed by partial fraction decomposition yields in

$$\frac{dz}{1-z} = \left(\frac{1}{(1-vp)u} + \frac{v(1-p)}{(1-vp)(1-v(p+(1-p)u))}\right) du.$$

Integrating the equation by using the substitution $x = v(p+(1-p)u)$ on the right hand side results in

$$\begin{aligned}\ln\left(\frac{1}{1-z}\right) &= \frac{1}{1-vp}\ln(u) - \frac{(1-p)v}{(1-vp)(1-p)v}\ln(1-v(p+(1-p)u)) + \widetilde{C}_1\\ &= \frac{1}{1-vp}\ln\left(\frac{u}{1-v(p+(1-p)u)}\right) + \widetilde{C}_1\end{aligned}$$

and some simple algebraic manipulations show that

$$C_1 = \frac{\left(\frac{1-v(p+(1-p)u)}{u}\right)^{\frac{1}{1-vp}}}{1-z}, \tag{5.3}$$

where we define $C_1 := e^{\widetilde{C}_1}$. For convenience it is useful to define

$$R(u,v) := \frac{u}{1-v(p+(1-p)u}.$$

Secondly we consider

$$\frac{dF}{du} = \frac{F+vu}{u(1-v(p+(1-p)u))} = \frac{F}{u(1-v(p+(1-p)u))} + \frac{v}{1-v(p+(1-p)u)}. \tag{5.4}$$

Solving the homogenous differential equation

$$\frac{dF}{du} = \frac{F}{u(1+v(p+(1-p)u))},$$

by separation of the variables leads to

$$\begin{aligned}\frac{dF}{F} &= \frac{du}{u(1-v(p+(1-p)u))} = \left(\frac{1}{(1-vp)u} + \frac{v(1-p)}{(1-vp)(1-v(p+(1-p)u))}\right) du\\ \ln(F) &= \frac{1}{1-vp}\ln(R(u,v)) + \widetilde{C}_2\\ F &= C_2 \cdot R(u,v)^{\frac{1}{1-vp}},\end{aligned}$$

with $C_2 := e^{\widetilde{C}_2}$. Applying variation of the constant leads to the following identities:

$$F = C(u) \cdot R(u,v)^{\frac{1}{1-vp}}, \tag{5.5}$$

$$\frac{dF}{du} = C'(u) \cdot R(u,v)^{\frac{1}{1-vp}} + \frac{C(u)R(u,v)^{\frac{1}{1-vp}-1}}{1-vp} \cdot \frac{1-v(p+(1-p)u)+uv(1-p)}{(1-v(p+(1-p)u))^2}, \tag{5.6}$$

which when plugging (5.5) and (5.6) into (5.4) and applying some careful algebraic manipulations yields

$$C'(u) = \frac{v}{u}\ R(u,v)^{-\frac{vp}{1-vp}}.$$

Hence, when using the fundamental theorem of calculus, which can be found in [16], we receive the solution

$$C(u) = \int \frac{v}{u} \; R(u,v)^{-\frac{vp}{1-vp}} du = v \int_{u_0}^{u} \frac{1}{t} \; R(t,v)^{-\frac{vp}{1-vp}} dt, \tag{5.7}$$

with an arbitrary $u_0 \in (0,\infty)$. When adding (5.7) to the solution of the homogenous ODE yields

$$F = R(u,v)^{\frac{1}{1-vp}} \cdot \int_{u_0}^{u} \frac{1}{t} R(t,v)^{-\frac{vp}{1-vp}} dt + C_2 \cdot R(u,v)^{-\frac{vp}{1-vp}},$$

or equivalently,

$$C_2 = F \cdot R(u,v)^{\frac{vp}{1-vp}} - \int_{u_0}^{u} \frac{1}{t} \; R(t,v)^{-\frac{vp}{1-vp}} dt. \tag{5.8}$$

Thus, we may write f in the form $f(z,u,F) = G(C_1,C_2)$ with an arbitrary differentiable function G. As by our approach it holds $G(C_1,C_2) = 0$ we may apply the fundamental theorem of implicit functions, which can be found in [16], to achieve the representation $C_2 = g(C_1)$ with an arbitrary differentiable function g and plugging (5.3) and (5.8) into it yields

$$F \cdot R(u,v)^{\frac{vp}{1-vp}} - \int_{u_0}^{u} \frac{1}{t} \; R(t,v)^{-\frac{vp}{1-vp}} dt = g\left(\frac{R(u,v)^{-\frac{1}{1-vp}}}{1-z}\right).$$

Choosing $u_0 = 1$ finally leads to a representation for F by

$$F(z,u,v) = R(u,v)^{\frac{1}{1-vp}} \cdot \left[\int_{1}^{u} \frac{1}{t} \; R(t,v)^{-\frac{vp}{1-vp}} dt + g\left(\frac{R(u,v)^{-\frac{1}{1-vp}}}{1-z}\right)\right].$$

To determine the function g we use the initial condition $F(0,u,v) = 0$ which follows trivially from the combinatorial structure of the hiring problem. Thus, we get the representation

$$-v \int_{1}^{u} \frac{1}{t} \; R(t,v)^{-\frac{vp}{1-vp}} dt = g\left(R(u,v)^{-\frac{1}{1-vp}}\right).$$

When applying the substitution $x = R(u,v)^{-\frac{1}{1-vp}} = \left(\frac{1-v(p(1-p)u)}{u}\right)^{\frac{1}{1-vp}}$ together with some careful algebraic manipulations we achieve the representation

$$g(x) = -v \int_{1}^{\frac{1-vp}{x^{1-vp}+v(1-p)}} \frac{1}{t} \; R(t,v)^{-\frac{vp}{1-vp}} dt,$$

which then, after substituting $y = \frac{x}{1-z}$, implies that

$$
\begin{aligned}
F(z,u,v) &= vR(u,v)^{\frac{1}{1-vp}} \left[\int_1^u \frac{1}{t} \, R(t,v)^{-\frac{vp}{1-vp}} dt - \int_1^{\frac{1-vp}{y^{1-vp}+v(1-p)}} \frac{1}{t} \, R(t,v)^{-\frac{vp}{1-vp}} dt \right] \\
&= vR(u,v)^{\frac{1}{1-vp}} \int_{\frac{1-vp}{y^{1-vp}+v(1-p)}}^{u} \frac{1}{t} \, R(t,v)^{-\frac{vp}{1-vp}} dt \\
&= vR(u,v)^{\frac{1}{1-vp}} \int_{\frac{(1-vp)u(1-z)^{1-vp}}{1-v(p+(1-p)u)+(1-p)uv(1-z)^{1-vp}}}^{u} \frac{1}{t} \, R(t,v)^{-\frac{vp}{1-vp}} dt
\end{aligned}
$$

As we are mainly interested in the size of the hiring set independent of the score of the threshold we consider the function

$$
\begin{aligned}
F(z,v) := F(z,1,v) = vR(1,v)^{-\frac{1}{1-vp}} & \int_{\frac{(1-vp)(1-z)^{1-vp}}{1-v+(1-p)v(1-z)^{1-pv}}}^{1} \frac{1}{t} \, R(t,v)^{-\frac{vp}{1-vp}} dt \\
= \frac{v}{(1-v)^{\frac{1}{1-vp}}} & \int_{\frac{(1-vp)(1-z)^{1-vp}}{1-v+(1-p)v(1-z)^{1-pv}}}^{1} \frac{1}{t} \, R(t,v)^{-\frac{vp}{1-vp}} dt.
\end{aligned}
$$

Substituting $s = \frac{1}{t}$ then leads to

$$
\begin{aligned}
F(z,v) &= \frac{v}{(1-v)^{\frac{1}{1-vp}}} \int_1^{\frac{1-v+(1-p)v(1-z)^{1-pv}}{(1-vp)(1-z)^{1-vp}}} \frac{1}{s} \left((1-vp)s - v(1-p)\right)^{\frac{vp}{1-vp}} ds \\
&= \frac{v}{1-v} \int_1^{\frac{1-v}{(1-vp)(1-z)^{1-vp}} + \frac{(1-p)v}{1-vp}} \frac{1}{s} \left(\frac{(1-vp)s-(1-p)v}{1-v} \right)^{\frac{vp}{1-vp}} ds,
\end{aligned}
$$

and another substitution by $r = \left(\frac{(1-vp)s-(1-p)v}{1-v} \right)^{\frac{1}{1-vp}}$ yields in the representation

$$
\begin{aligned}
F(z,v) &= \frac{v}{1-v} \int_1^{\frac{1}{1-z}} \frac{(1-vp)(1-v)}{(1-v)r^{1-vp} + (1-p)v} dr \\
&= v(1-vp) \int_1^{\frac{1}{1-z}} \frac{dr}{(1-v)r^{1-vp} + (1-p)v}. \qquad (5.9)
\end{aligned}
$$

With this representation we are now able to determine the expectation and the moments of the parameter h_n.

Theorem 5.1.2. For the expectation of the size of the hiring set h_n when applying the probabilistic hiring strategy it holds that

$$\mathbb{E}(h_n) = -\frac{p}{1-p} + \frac{1}{(1-p)(2-p)}\binom{n+1-p}{n} = \frac{n^{1-p}}{(1-p)\,\Gamma(3-p)} + \mathcal{O}(1),$$

where the asymptotic expansion holds uniformly for $n \to \infty$.

Proof. Following Theorem 1.4.7, the PGF of h_n is given by

$$P(v) = \frac{[z^n]F(z,v)}{[z^n]F(z,1)}.$$

Considering (5.9) implies that

$$F(z,1) = (1-p)\int_1^{\frac{1}{1-z}} \frac{dr}{1-p} = \frac{z}{1-z},$$

and as a direct consequence we have $[z^n]F(z,1) = 1$, for $n \geq 1$. This result is consistent with the definition of $a_{n,k,d}$, as for $u = v = 1$ we only sum up the probabilities for fixed n. Hence, the PGF is given by

$$P(v) = [z^n]v(1-pv)\int_1^{\frac{1}{1-z}} \frac{dr}{(1-v)r^{1-vp} + (1-p)v},$$

which means for the expectation that

$$\mathbb{E}(h_n) = P'(v)\big|_{v=1} = [z^n]\frac{\partial}{\partial v}F(z,v)\big|_{v=1} \tag{5.10}$$

For the derivative of F it is an easy task to prove that we may change the order of integration and differentiation which immediately shows that

$$\begin{aligned}\frac{\partial}{\partial v}F(z,v) &= (1-2pv)\int_1^{\frac{1}{1-z}} \frac{dr}{(1-v)r^{1-vp} + (1-p)v} \\ &\quad - v(1-vp)\int_1^{\frac{1}{1-z}} \frac{-r^{1-vp} + (1-v)r^{1-vp}(-p)\ln(r) + (1-p)}{(1-v)r^{1-vp} + (1-p)v)^2}.\end{aligned} \tag{5.11}$$

When plugging (5.11) into (5.10) and evaluating at $v = 1$ it leads to

$$\begin{aligned}\mathbb{E}(h_n) &= [z^n]\left((1-2p)\int_1^{\frac{1}{1-z}} \frac{dr}{1-p} - (1-p)\int_1^{\frac{1}{1-z}} \frac{-r^{1-p} + (1-p)}{(1-p)^2}dr\right) \\ &= [z^n]\left(\frac{1-2p}{1-p}\frac{z}{1-z} + \frac{1}{(1-p)(2-p)}\left(\frac{1}{(1-z)^{2-p}} - 1\right) - \frac{z}{1-z}\right) \\ &= -\frac{p}{1-p} + \frac{1}{(1-p)(2-p)}\binom{n+1-p}{n},\end{aligned}$$

for $n \geq 1$. Asymptotically, for $n \to \infty$, applying the identity

$$\binom{n+1-p}{n} = \frac{\Gamma(n+2-p)}{\Gamma(n+1)\Gamma(2-p)},$$

together with Stirling's formula for the gamma function, Theorem 1.3.8 shows that

$$\begin{aligned}\mathbb{E}(h_n) &= \frac{1}{(1-p)(2-p)}\binom{n+1-p}{n} - \frac{p}{1-p}\\ &= \frac{1}{(1-p)(2-p)\Gamma(2-p)}\frac{(n+2-p)^{n+2-p-\frac{1}{2}}\ e^{n+1}}{(n+1)^{n+1-\frac{1}{2}}\ e^{n+2-p}}\cdot\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right)+\mathcal{O}(1)\\ &= \frac{n^{1-p}}{(1-p)\Gamma(3-p)}+\mathcal{O}(1).\end{aligned}$$

□

We were also able to determine the other moments $\mathbb{E}(X^r)$, for arbitrary $r > 0$, of h_n. As the computations are a bit lengthy we will only sketch a proof. It is useful to consider the function

$$\widehat{F}(z,w) = F(z,w+1) = \left(1+\frac{1-2p}{1-p}w - \frac{p}{1-p}w^2\right)\cdot\int_1^{\frac{1}{1-z}}\frac{dx}{1-w\left(\frac{x^{1-p}}{1-p}e^{-pw\ln(x)}-1\right)}.$$

When defining $S(w,x) := \frac{x^{1-p}}{1-p}e^{-pw\ln(x)} - 1$ and applying the identity $\sum_{k=0}^{\infty} y^k = \frac{1}{1-y}$ to the integrand shows that

$$\begin{aligned}[w^r]\widehat{F}(z,w) = &\sum_{j=0}^{r}\int_1^{\frac{1}{1-z}} S(w,x)^j dx + \frac{1-2p}{1-p}\sum_{j=0}^{r-1}[w^{r-1-j}]\int_1^{\frac{1}{1-u}} S(w,x)^j dx\\ &- \frac{p}{1-p}\sum_{j=0}^{r-2}[w^{r-2-j}]\int_1^{\frac{1}{1-z}} S(w,x)^j dx.\end{aligned}$$

When changing the order of integration and the coefficient extraction operator and defining

$$T(w,x,j) := \sum_{\ell=0}^{j}\binom{j}{\ell}(-1)^{j-\ell}\frac{x^{(1-p)\ell}e^{-\ell pw\ln(x)}}{(1-p)^\ell}$$

then shows that

$$\begin{aligned}[w^r]\widehat{F}(z,w) = &\sum_{j=0}^{r}\int_1^{\frac{1}{1-z}} T(w,x,j)dx + \frac{1-2p}{1-p}\sum_{j=0}^{r-1}\int_1^{\frac{1}{1-z}} T(w,x,j)dx\\ &- \frac{p}{1-p}\sum_{j=0}^{r-2}\int_1^{\frac{1}{1-z}} T(w,x,j)dx.\end{aligned}$$

Hence, we get the asymptotic representation

$$\begin{aligned}[w^r]\widehat{F}(z,w) &= \int_1^{\frac{1}{1-z}}\frac{x^{(1-p)r}}{(1-p)^r}dx + \mathcal{O}\left(\int_1^{\frac{1}{1-z}} x^{(1-p)(r-1)}\ln(x)dx\right)\\ &= \frac{1}{(1-p)^r((1-p)r+1)(1-z)^{(1-p)r+1}} + \mathcal{O}\left(\frac{1}{(1-z)^{(1-p)(r-1)+1}}\ln\left(\frac{1}{1-z}\right)\right).\end{aligned}$$

One may now easily prove that

$$[z^n w^r]\widehat{F}(z,w) = \frac{n^{(1-p)r}}{(1-p)^r\Gamma((1-p)r+1)} \cdot \left(1 + \mathcal{O}\left(\frac{\ln(n)}{n^{1-p}}\right)\right).$$

Now we are able to determine all the moments $\mathbb{E}(h_n^r)$, for $r > 0$.

Theorem 5.1.3. For a number $r \in \mathbb{N}$, the r-th moment of the random variable h_n, that denotes the number of hired candidates when applying the probabilistic strategy, is asymptotically, for $n \to \infty$, given by

$$\mathbb{E}(h_n^r) = \frac{r!\ n^{(1-p)r}}{(1-p)^r\Gamma((1-p)r+2)} \cdot \left(1 + \mathcal{O}\left(\frac{\ln(n)}{n^{1-p}}\right)\right).$$

Moreover, suitably normalized, h_n converges, for $n \to \infty$, in distribution to a random variable H, i.e. $\frac{h_n}{n^{1-p}} \xrightarrow{(d)} H$, where H is fully determined by its integer moments

$$\mathbb{E}(H^r) = \frac{r!}{(1-p)^r\Gamma((1-p)r+2)}.$$

Proof. We first consider the r-th factorial moment of h_n, which is, following Lemma 1.4.8, given by

$$\mathbb{E}(h_n^{\underline{r}}) = r!\ [z^n w^r]\widehat{F}(z,w) = \frac{r! n^{(1-p)r}}{(1-p)^r\Gamma((1-p)r+2)} \cdot \left(1 + \mathcal{O}\left(\frac{\ln(n)}{n^{1-p}}\right)\right).$$

It is a simple task to prove that the same asymptotic expansion holds for the r-th moments as well. Furthermore, it is easy to see that the moments of $\frac{h_n}{n^{1-p}}$ satisfy the conditions of the Fréchet-Shohat Theorem 1.6.4, hence, we directly get

$$\frac{h_n}{n^{1-p}} \xrightarrow{(d)} H,$$

where H is fully determined by its integer moments.

□

5.2 Connection between the probabilistic and deterministic strategies

In this section we will compare the probabilistic strategy from Section 5.1 to the other strategies we considered in the Chapters 2, 3 and 4.

5.2.A Hiring above the maximum

First of all let us observe the not yet investigated case $p = 1$. Here the threshold candidate obviously changes after each hiring. It is an easy task to prove that the threshold candidate is the best already hired candidate at each step n. In particular we have the following induction:

- $n = 1$: The strategy hires the first candidate. Thus, the first threshold candidate is the best already hired candidate, since he is the only one.
- $n > 1$: By inductive hypothesis the threshold candidate after step $n-1$ is the best so far seen candidate. A new candidate gets hired if and only if he is better than the threshold, i.e. if he is the best candidate. As the threshold changes after each hiring the new candidate becomes the new threshold.

As a direct consequence we can see that for $p = 1$ we have hiring above the maximum. We can also see this when plugging $p = 1$ into (5.1):

$$a_{n,k,d} = \frac{n-k}{n} a_{n-1,k,d} + \frac{k}{n} a_{n-1,k,d-1}, \quad n \geq 2.$$

The initial value is $a_{1,1,1} = 1$. As we can see the parameter k does not change in the recursion, i.e. $k = 1$ for sure. This leads to the recursion

$$\widehat{a}_{n,d} = \frac{n-1}{n} \widehat{a}_{n-1,d} + \frac{1}{n} \widehat{a}_{n-1,d-1}, \quad n \geq 2, \tag{5.12}$$

with the initial value $\widehat{a}_{1,1} = 1$ and $\widehat{a}_{n,d} = a_{n,1,d}$. Obviously (5.12) is a recursion that describes hiring above the maximum.

Remark 5.2.1. A generalization for this probabilistic strategy to hiring above the m-th best seems rather difficult as we have to change the behavior of the threshold transition after m candidates. However, asymptotically for $n \to \infty$ and m fixed, $p = 1$ represents hiring above the m-th best, too.

■

5.2.B Hiring above the minimum

Secondly we consider $p = 0$ which is the contrary case to $p = 1$ considered before. Again this is a deterministic approach, as the threshold does not change. So the hiring process recruits the first candidate who becomes the first threshold. Thus, any further candidate gets hired if and only if he is better than the first candidate. Hence, we have hiring above the minimum for $p = 0$. This is also confirmed by the results in Section 5.1:

$$\mathbb{E}(h_n) = -\frac{p}{1-p} + \frac{1}{(1-p)(2-p)} \binom{n+1-p}{n} = \frac{1}{2} \binom{n+1}{n} = \frac{n+1}{2}.$$

5.2.C Hiring above the median

Finally we compare the probabilistic strategy with $p = \frac{1}{2}$ to hiring above the median. Let us recap our considerations about hiring above the median. By using the automaton from Figure 4.1 we could observe that the threshold candidate changes after every second hiring, depending on the parity of the current hiring set. Thus, if we do not know the parity of the hiring set, the probability for changing the threshold lies at $p = \frac{1}{2}$. On the other hand, hiring above the median is a deterministic strategy. This shows that hiring above the median is an instance of the probabilistic strategy with $p = \frac{1}{2}$. Thus, one would expect that the two strategies lead, asymptotically, to the same results, but with other constants. We can observe this behavior for the expectation $\mathbb{E}(h_n)$. For hiring above the median we saw that

$$\mathbb{E}(h_n) = \sqrt{\pi n} + \mathcal{O}(1),$$

and when plugging $p = \frac{1}{2}$ into the results for the probabilistic strategy we have

$$\mathbb{E}(h_n) = \frac{8}{3\sqrt{\pi}}\sqrt{n} + \mathcal{O}(1).$$

We were also able to determine the limit distribution of h_n for $n \to \infty$ when applying the probabilistic strategy with parameter $p = \frac{1}{2}$.

Theorem 5.2.2. When applying the probabilistic strategy with parameter $p = \frac{1}{2}$, the number of recruited candidates h_n converges, asymptotically, for $n \to \infty$, in distribution to a limiting distribution H that is characterized by its density function, i.e.

$$\frac{h_n}{\sqrt{n}} \xrightarrow{(d)} H,$$

where H has the density function

$$h(x) = \frac{x}{\sqrt{\pi}} \int_x^\infty \frac{e^{-\frac{t^2}{16}}}{t^2} dt, \quad x > 0.$$

Proof. By Theorem 5.1.3 $\frac{h_n}{\sqrt{n}}$ converges in distribution to a unique random variable H whose moments are given by

$$\mathbb{E}(H^r) = \frac{2^{2r}\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\left(\frac{r}{2}+1\right)}.$$

The stated density function $h(x)$ satisfies

$$\begin{aligned}\mathbb{E}(X_h^r) &= \int_0^\infty x^r h(x)dx = \frac{1}{\sqrt{\pi}} \int_0^\infty x^{r+1} \int_x^\infty \frac{e^{-\frac{t^2}{16}}}{t^2} dt\ dx = \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-\frac{t^2}{16}}}{t^2} \int_0^t x^{r+1} dx\ dt \\ &= \frac{1}{\sqrt{\pi}(r+2)} \int_0^\infty e^{-\frac{t^2}{16}} t^r dt,\end{aligned}$$

where X_h denotes the corresponding random variable that belongs to $h(x)$. Applying the substitution $\frac{t^2}{16} = u$ leads to

$$\mathbb{E}(X_h^r) = \frac{2 \cdot 4^r}{\sqrt{\pi}(r+2)} \int_0^\infty u^{\frac{r-1}{2}} e^{-u} du = \frac{2^{2r}\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\left(\frac{r}{2}+1\right)}.$$

By applying Legendre's duplication formula, Lemma 1.3.9, one can easily show that the random variables H and X_h have the same moments and as H is uniquely determined by its moments we have $H = X_h$.

□

Bibliography

[1] Margaret Archibald and Conrado Martínez. The hiring problem and permutations. *DMTCS Proceedings*, (01):63–76, 2009.

[2] Barry C Arnold, Narayanaswamy Balakrishnan, and Haikady Navada Nagaraja. *Records*, volume 768. John Wiley & Sons, 2011.

[3] Andrei Z Broder, Adam Kirsch, Ravi Kumar, Michael Mitzenmacher, Eli Upfal, and Sergei Vassilvitskii. The hiring problem and lake wobegon strategies. *SIAM Journal on Computing*, 39(4):1233–1255, 2009.

[4] John H Curtiss. A note on the theory of moment generating functions. *The Annals of Mathematical Statistics*, 13(4):430–433, 1942.

[5] Volker Diekert, Manfred Kufleitner, and Gerhard Rosenberger. *Elemente der diskreten Mathematik: Zahlen und Zählen, Graphen und Verbände*. Walter de Gruyter, 2013.

[6] Philippe Flajolet, Xavier Gourdon, and Philippe Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical computer science*, 144(1):3–58, 1995.

[7] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, New York, NY, USA, 1 edition, 2009.

[8] M Fréchet and J Shohat. A proof of the generalized second-limit theorem in the theory of probability. *Transactions of the American Mathematical Society*, 33(2):533–543, 1931.

[9] J Gaither and MD Ward. Analytic methods for select sets. *Probability in the Engineering and Informational Sciences*, 26(04):561–568, 2012.

[10] Martin Gardner. Mathematical games. *Scientific American*, 243(6):18–28, 1980.

[11] John P Gilbert and Frederick Mosteller. Recognizing the maximum of a sequence. In *Selected Papers of Frederick Mosteller*, pages 355–398. Springer, 2006.

[12] Bernhard Gittenberger. Diskrete Methoden. Course Notes, Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien. Version Summer term 2010. `http://www.dmg.tuwien.ac.at/gittenberger/dmeth.pdf`. Online; accessed 12/04/2013.

[13] DMBTL Griffiths and MIJJB Tenenbaum. Hierarchical topic models and the nested chinese restaurant process. *Advances in neural information processing systems*, 16:17, 2004.

[14] Ahmed Helmi, Conrado Martínez, and Alois Panholzer. Hiring above the m-th best candidate: a generalization of records in permutations. In *LATIN 2012: Theoretical Informatics*, pages 470–481. Springer, 2012.

[15] Ahmed Helmi and Alois Panholzer. Analysis of hiring above the median: A Lake Wobegon strategy for the hiring problem. In *ANALCO*, pages 76–83. SIAM, 2012.

[16] Harro Heuser. *Lehrbuch der Analysis*. Springer, 2009.

[17] Michael Kaltenbäck and Harald Woracek. Komplexe Analysis für Technische Mathematik. Course Notes, Institut ASC, Technische Universität Wien. Version 25.6.2013. `http://asc.tuwien.ac.at/~funkana/kaltenbaeck/mykana/mykana.pdf`. Online; accessed 07/05/2014.

[18] Donald E. Knuth. *The Art of Computer Programming, Volume 1 (3rd Ed.): Fundamental Algorithms*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1997.

[19] Donald Erwin Knuth, Ronald Graham, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison, 2 edition, 1994.

[20] Abba M Krieger, Moshe Pollak, Ester Samuel-Cahn, et al. Select sets: rank and file. *The Annals of Applied Probability*, 17(1):360–385, 2007.

[21] Markus Kuba and Helmut Prodinger. A note on stirling series. *Integers*, 10(4):393–406, 2010.

[22] Meyberg Kurt and Vachenauer Peter. *Höhere Mathematik 2 - Differentialgleichungen, Funktionentheorie, Fourier-Analysis, Variationsrechnung*. Springer-Verlag Berlin Heidelberg New York, 2006.

[23] N. Kusolitsch. *Maß- und Wahrscheinlichkeitstheorie: Eine Einführung*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2014.

[24] Michel Loeve. Probability theory. *Graduate Texts in Mathematics*, 45, 1963.

[25] Peter Norvig. Hiring: The lake wobegon strategy. *Official Google Research*, 2006.

[26] Jim Pitman et al. Combinatorial stochastic processes. Technical report, Springer, 2002.

[27] Walter Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 1987.