

The hiring problem with rank-based strategies*

Svante Janson[†]

Dedicated to the memory of Philippe Flajolet (1948–2011)

Abstract

The hiring problem is studied for general strategies based only on the relative ranking of the candidates; this includes some well known strategies studied before such as hiring above the median. We give general limit theorems for the number of hired candidates and some other properties, extending previous results. The results exhibit a dichotomy between two classes of rank-based strategies: either the asymptotics of the process are determined by the early events, with a.s. convergence of suitably normalized random variables, or there is a mixing behaviour without long-term memory and with asymptotic normality.

Keywords: hiring problem; multiple secretary problem; hiring above median.

AMS MSC 2010: Primary 60C05; 68W40, Secondary 62L10; 68W27.

Submitted to EJP on September 11, 2018, final version accepted on October 27, 2019.

Supersedes arXiv:1809.03800v1.

1 Introduction

The *hiring problem* is a variant of the well-known secretary problem, in which we want to hire many good applicants and not just the best. An informal formulation is that a large number of candidates are examined (interviewed) one by one; immediately after each interview we decide whether to hire the candidate or not, based on the value of the candidate (which is assumed to be revealed during the interview) and of the values of the candidates seen earlier. This is thus an on-line type of decision problem. The mathematical model assumes that the values of the candidates are i.i.d. random variables, with some continuous distribution (which prevents ties). See below and Section 3 for formal details.

There are two conflicting aims in the hiring problem: we want to hire (rather) many candidates but we also want them to be good. Thus there is no single goal in the hiring problem, and thus we cannot talk about an optimal solution. Instead, the mathematical

*Partly supported by the Knut and Alice Wallenberg Foundation.

[†]Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.

E-mail: svante.janson@math.uu.se

problem is to analyse properties of various proposed strategies. The property that has been most studied is the number of accepted candidates among the first n examined, here denoted M_n . We will also study the inverse function N_m , the number of candidates examined until m are accepted. Some other properties, such as the distribution of the value of the accepted candidates, are discussed in Sections 9–11.

The hiring problem seems to have been studied first by Preater [18], and later (independently) by Broder et al [4], who also invented the name ‘hiring problem’. Further papers studying the hiring problem under various strategies are Krieger, Pollak and Samuel-Cahn [14, 15, 16], Archibald and Martínez [2], Gaither and Ward [6], Helmi and Panholzer [10, 11], Helmi, Martínez and Panholzer [8, 9]; see [9] for a more detailed history.

There are two main groups of strategies. In the present paper we study only *rank-based strategies*, i.e., strategies that depend only on the rank of the candidate among the ones seen so far; in other words, on the relative order of the values of the candidates. A typical example is ‘hiring above the median’, see below; see also [14, 4, 2, 6, 10, 8, 11, 9]. In statistical terms, the values of the candidates are regarded as on an ordinal scale. Thus, the distribution of the value of a candidate does not matter, and can freely be chosen as e.g. uniform (an obvious standard choice used by some previous authors) or exponential (used in the analysis in the present paper). Furthermore, for rank-based strategies, it is equivalent (for a fixed n) to assume that the values of the first n candidates form a uniformly random permutation of $\{1, \dots, n\}$ [14, 2, 6, 10, 8, 11, 9].

The alternative is to use a strategy depending on the actual values; a typical example is ‘hiring above the mean’ [18, 4, 15, 16]. For such strategies, the results depend on the given distribution of the value of a candidate; several different distributions have been investigated in the papers just mention. Such strategies will not be considered in the present paper.

In the present paper we thus study rank-based strategies. More precisely, following Krieger, Pollak and Samuel-Cahn [14], we consider strategies of the following form (which seems to include all reasonable rank-based strategies). We assume throughout the paper that we are given a sequence of integers $r(m)$, $m \geq 0$, such that

$$r(0) = 1 \quad \text{and} \quad r(m) \leq r(m + 1) \leq r(m) + 1, \quad m \geq 0. \quad (1.1)$$

Note that this implies $1 \leq r(m) \leq m + 1$ for every $m \geq 0$. Then the strategy is that if so far $m \geq 0$ candidates have been accepted, then (if $r(m) \leq m$) the next candidate is accepted if her value is above the $r(m)$ -th best value of the ones already accepted. If $r(m) = m + 1$, we always accept the next candidate.

Remark 1.1. Some authors have chosen to define smaller values as better. This is obviously equivalent, but some care has to be taken when comparing definitions and results. (In the discussion here, we have when necessary translated to our setting.)

One basic example is ‘*hiring above the median*’, where we hire every candidate that is better than the median of the ones seen so far. For m odd, this means taking $r(m) = (m + 1)/2$, but for m even, there are two possible values of the median. Tradition established by previous authors [4, 2, 11] says that we choose the smaller value as the threshold. (This is thus the less restrictive policy of the two possibilities.) This means taking $r(m) = m/2 + 1$ when m is even, so we can summarize the strategy ‘hiring above the median’ by

$$r(m) := \lfloor m/2 \rfloor + 1 = \lceil (m + 1)/2 \rceil, \quad m \geq 0. \quad (1.2)$$

The sequence (starting with $r(0)$) thus is 1, 1, 2, 2, 3, 3, . . .

The alternative interpretation of ‘median’ when m is even is instead a special case of the strategy known as ‘*the α -percentile rule*’, which, again by tradition [14, 11], is

defined by

$$r(m) := \lceil \alpha m \rceil, \quad m \geq 1. \tag{1.3}$$

In the case $\alpha = \frac{1}{2}$, we thus take $r(m) = m/2$ when $m \geq 2$ is even, meaning a smaller $r(m)$ and thus a larger threshold than in ‘hiring above the median’; the sequence (starting with $r(0)$) is 1, 1, 1, 2, 2, 3, . . .

A third simple example is ‘hiring above the r -th best’, where $r \geq 1$ is a fixed number [2, 8, 9]. This means $r(m) = r$ when $m \geq r$; we always hire the r first candidates, so the complete definition in accordance with (1.1) is

$$r(m) := \min\{r, m + 1\}, \quad m \geq 0. \tag{1.4}$$

Note that the case $r = 1$ gives the strategy of hiring only the candidates that are better than everyone seen earlier, i.e., the records.

The present paper gives a general analysis of strategies of the type above, with an arbitrary sequence $r(m)$ satisfying (1.1). Our main results give the asymptotic distribution of N_m (in general) and M_n (under weak regularity assumptions on $r(m)$), see in particular Theorems 4.6, 7.5, 1.2 and 1.3. In particular, this gives new proofs of known results (and some new) for the examples above.

It turns out that there is a dichotomy: the general results are of two different types, depending on whether $\sum_m r(m)^{-2} < \infty$ or $\sum_m r(m)^{-2} = \infty$; we will call these two cases *large* $r(m)$ and *small* $r(m)$, respectively. Note that the strategies ‘hiring above the median’ and ‘the α -percentile rule’ have large $r(m)$, while ‘hiring above the r -th best’ has small $r(m)$; indeed, the limit theorems found by previous researchers for these cases are of different types, compare e.g. [11] and [9].

The main differences between the two cases can be summarized as follows, at least assuming some further regularity of $r(m)$. For simplicity, we consider here only M_n ; the same types of results hold for N_m .

large $r(m)$, $\sum_m r(m)^{-2} < \infty$ $M_n/\mathbb{E}M_n$ converges to a non-degenerate distribution on $[0, \infty)$. (Thus, the limit is not normal.) Furthermore, $M_n/\mathbb{E}M_n$ converges a.s. Hence, the limit and the asymptotic behaviour are essentially determined by what happens very early, i.e., by the values of the first few candidates. This also means a strong long-range dependence in the sequence M_n .

small $r(m)$, $\sum_m r(m)^{-2} = \infty$ Asymptotic normality of M_n . There is no long-range dependency; instead M_n is asymptotically independent of what happened with the first n_0 candidates, for any fixed n_0 . In particular, there is no a.s. convergence.

Intuitively, the reason for the difference between the two cases is that when $r(m)$ is small, each accepted candidate has (typically) a rather large influence on the threshold, and thus on the future of the process, and these influences add up and eventually dominate over the influences of the values of the first candidates, while if $r(m)$ is large, then the influences of later candidates are small, and the effects of the first few candidates dominate.

We state here two theorems that exemplify our main results, one for large $r(m)$ and one for small. In both cases, we assume a regularity condition on $r(m)$ yielding simpler results; proofs and more general results are given in Sections 4–5 and 7. See also the examples in Section 8, including the counterexample Example 8.9.

First, consider the case $r(m) = \alpha m + O(1)$, where $0 < \alpha \leq 1$; note that this includes ‘hiring above the median’ (with $\alpha = \frac{1}{2}$) and ‘the α -percentile rule’, and that such $r(m)$ are large.

Theorem 1.2. Suppose that $r(m) = \alpha m + O(1)$, where $0 < \alpha \leq 1$. Then,

$$M_n/n^\alpha \xrightarrow{\text{a.s.}} W, \tag{1.5}$$

for a random variable W , which can be represented as in (5.8) below. Furthermore, all moments converge in (1.5), and, for every $s \geq 0$,

$$\mathbb{E} M_n^s/n^{\alpha s} \rightarrow \mathbb{E} W^s = \frac{\Gamma(s+1)}{\Gamma(s\alpha+2)} \prod_{k=1}^{\infty} \frac{1+s/k}{1+s\alpha/r(k)} < \infty. \tag{1.6}$$

The moments in (1.6) can be explicitly computed in many cases, see [6], Theorem 6.1 and Examples 8.1–8.2.

For our example result in the case of small $r(m)$, we assume that the sequence $r(m)$ is regularly varying. (See e.g. [3, p. 52] for definition, and [1] for definition of a mixing limit theorem.)

Theorem 1.3. Assume that $r(m)$ is a regularly varying sequence such that $\sum_m r(m)^{-2} = \infty$. Let $\mu(n)$, $n \geq 1$, be a sequence of real numbers such that

$$\sum_{k=1}^{\mu(n)} \frac{1}{r(k)} = \log n + O(1), \tag{1.7}$$

and let

$$\beta(m)^2 := r(m)^2 \sum_{k=1}^m \frac{1}{r(k)^2} = \sum_{k=1}^m \frac{r(m)^2}{r(k)^2}, \tag{1.8}$$

$$\gamma(n) := \beta(\lfloor \mu(n) \rfloor). \tag{1.9}$$

Then, as $n \rightarrow \infty$,

$$\frac{M_n - \mu(n)}{\gamma(n)} \xrightarrow{d} N(0, 1). \tag{1.10}$$

Furthermore, (1.10) is mixing.

Section 3 contain some basic general results. Then Theorem 1.2 and related results for large $r(m)$ are proved in Sections 4–6, while Theorem 1.3 and related results for small $r(m)$ are proved in Section 7. Section 8 contains some examples. The remaining sections consider some additional properties that have been considered by previous authors. Section 9 consider results conditioned on the value of the first candidate. Section 10 treat the probability of accepting the next candidate, and also the number of unsuccessful candidates since the last accepted one. Section 11 studies the distribution of the accepted values.

Remark 1.4. Some previous papers, in particular [11, 9], contain also interesting exact results on the exact distribution of M_n for finite n . We do not consider such results here.

2 Notation

$\text{Exp}(a)$ denotes the exponential distribution with rate a . In other words, if $X \in \text{Exp}(a)$, then $\mathbb{P}(X > x) = e^{-ax}$ for $x \geq 0$; equivalently, $aX \in \text{Exp}(1)$. Hence, $\mathbb{E} X = 1/a$.

$\text{Ge}(p)$ denotes the Geometric distribution started at 1 (also called First Success distribution), with $\mathbb{P}(X = n) = p(1-p)^{n-1}$, $n \geq 1$.

E, E_i, E'_j and so on will always denote $\text{Exp}(1)$ variables, independent of each other.

γ is Euler’s constant.

$\xrightarrow{\text{a.s.}}$, \xrightarrow{p} and \xrightarrow{d} denote convergence almost surely (a.s.), in probability and in distribution, respectively, for random variables.

$a_m \sim b_m$, where a_m and b_m are real numbers, means $a_m/b_m \rightarrow 1$ as $m \rightarrow \infty$. Furthermore, assuming $b_m > 0$, $a_m = O(b_m)$ means $\sup_m |a_m|/b_m < \infty$ and $a_m = o(b_m)$ means $a_m/b_m \rightarrow 0$ as $m \rightarrow \infty$. Moreover, if X_m are random variables and $b_m > 0$ are real numbers, then $X_m = O_p(b_m)$ means that the sequence $|X_m|/b_m$ is stochastically bounded (tight), i.e., $\sup_m \mathbb{P}(|X_m|/b_m > K) \rightarrow 0$ as $K \rightarrow \infty$, and $X_m = o_p(b_m)$ means $X_m/b_m \xrightarrow{P} 0$ as $m \rightarrow \infty$. We sometimes omit ‘as $m \rightarrow \infty$ ’ when clear from the context.

With high probability (w.h.p.) means with probability $1 - o(1)$ (as, e.g., $n \rightarrow \infty$).

C and c are used for unspecified constants, which may vary from one occurrence to another. For constants that depend on some parameter (but not on other variables such as m or n), we similarly use e.g. C_K and $c(\delta)$.

If $x \in \mathbb{R}$, then $\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}$ and $\lceil x \rceil := \min\{n \in \mathbb{Z} : n \geq x\}$.

3 General limit theorems

We begin by formalising the hiring strategy discussed in Section 1, at the same time introducing some further notation. Recall that for a rank-based strategy, the result does not depend on the (continuous) distribution of the values of the candidates. We choose this distribution to be exponential.

Thus, let X_1, X_2, \dots be i.i.d. random variables with $X_i \in \text{Exp}(1)$, representing the values of the candidates. We assume without further mention that these values are distinct (which happens a.s.), so we ignore the possibility of ties below. For convenience we identify a candidate and her value; we will thus say both ‘candidate n is accepted’ and ‘value X_n is accepted’.

Let N_m be the index of the m -th accepted candidate, and denote the m -th accepted value by $X_m^* := X_{N_m}$. Conversely, let M_n be the number of candidates accepted among $1, \dots, n$. Thus,

$$M_n \geq m \iff N_m \leq n. \tag{3.1}$$

The hiring strategy is defined by a given function $r : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{> 0}$ satisfying (1.1), and thus, in particular, $1 \leq r(m) \leq m + 1$.

The basic rule of the strategy is that if $m \geq 0$ values have been accepted so far, then the next value is accepted if and only if it exceeds a threshold Y_m , which is the $r(m)$ -th largest value among the m values already accepted, interpreted as $Y_m := 0$ when $r(m) = m + 1$. (In particular, $Y_0 = 0$.)

Remark 3.1. It is easy to see [14] that this threshold Y_m is the same as the $r(m)$ -th best value of all candidates seen so far, since all previous candidates with values at or above Y_m were accepted. If $r(m) \leq m$, then the strategy is thus to accept a candidate if her value is among the $r(m)$ best of all values seen so far (including her own). It follows by symmetry that conditioned on $M_n = m$, and on everything else that has happened earlier, the probability that candidate $n + 1$ is accepted equals $r(m)/(n + 1)$, see [14].

The threshold Y_m is thus updated when a new value is accepted. This is described by the following lemma which is simple but basic for our analysis. In particular, note that Y_m never decreases.

Lemma 3.2. (i) If $r(m + 1) = r(m) + 1$, then $Y_{m+1} = Y_m$.

(ii) If $r(m + 1) = r(m)$, then Y_{m+1} is the smallest of the $r(m)$ values that are larger than Y_m among the selected values X_1^*, \dots, X_{m+1}^* .

Proof. Let m be fixed and order the accepted values X_1^*, \dots, X_m^* in decreasing order as $X_{(1)}^* > X_{(2)}^* > \dots > X_{(m)}^*$; define further $X_{(m+1)}^* := 0$. Then $Y_m = X_{(r(m))}^*$. By assumption, $X_{(m+1)}^* > Y_m$. Thus there are in the set $\{X_1^*, \dots, X_{m+1}^*\}$ exactly $r(m)$ values that are larger than $Y_m = X_{(r(m))}^*$, viz. $\{X_{(i)}^* : 1 \leq i < r(m)\} \cup \{X_{(m+1)}^*\}$. Hence, if $r(m + 1) = r(m)$, then

Y_{m+1} is the smallest of these values, while if $r(m + 1) = r(m)$, then Y_{m+1} is the next smaller accepted value (or 0), which is $X_{(r(m))}^* = Y_m$. \square

So far, the argument has been deterministic. We now use our assumption that the values X_i are i.i.d. random variables as above; this is where the choice of exponential distribution is convenient and greatly simplifies the argument.

Lemma 3.3. *Assume as above that X_1, X_2, \dots are i.i.d. and $\text{Exp}(1)$. Then, the increments $Y_{m+1} - Y_m$, $m \geq 0$, are independent random variables with*

$$Y_{m+1} - Y_m \in \begin{cases} \text{Exp}(r(m)), & r(m + 1) = r(m), \\ 0, & r(m + 1) = r(m) + 1. \end{cases} \quad (3.2)$$

Proof. Run the hiring process as above, but keep the values X_n secret as long as possible, revealing only enough to determine whether to accept X_n or not, and to determine the next threshold Y_m . To be precise, when a new candidate n is examined, reveal first only whether her value X_n is larger than the current threshold Y_m or not. If not, we forget this candidate and move on to the next. Suppose instead that $X_n > Y_m$, so that we accept n . Then we also have to update Y_m . By Lemma 3.2, this is trivial if $r(m + 1) = r(m) + 1$. However, if $r(m + 1) = r(m)$, then there are $r(m)$ accepted candidates (including the latest recruit, n) that have values $> Y_m$. We now reveal the minimum of these values, giving Y_{m+1} , but we do not reveal the remaining $r(m) - 1$ of them.

Claim 3.4. *Conditioned on $Y_m = y$ and on everything else that has been revealed so far, the $r(m) - 1$ (still hidden) accepted values that are larger than Y_m have the distribution of $r(m) - 1$ i.i.d. random variables with the distribution $\mathcal{L}(X \mid X > y)$.*

To show the claim, we use induction on m . We condition on $Y_m = y$ and everything else that has been revealed so far, and note that when we accept the next X_n , we know just that $X_n > y$, so X_n too has the distribution $\mathcal{L}(X \mid X > y)$. Hence, by the induction hypothesis, the $r(m)$ accepted values that are larger than $Y_m = y$ are $r(m)$ (conditionally) independent random variables with this distribution. By Lemma 3.2, this completes the induction step when $r(m + 1) = r(m) + 1$; otherwise we reveal the minimum Y_{m+1} of them, and note that conditioned on $Y_{m+1} = y' > y$, the remaining $r(m) - 1 = r(m + 1) - 1$ variables are i.i.d. with the distribution $\mathcal{L}(X \mid X > y')$.

This proves the claim. Furthermore, since $X \in \text{Exp}(1)$, this distribution $\mathcal{L}(X \mid X > y)$ is the same as the distribution of $X + y$. (The standard lack-of-memory property of exponential distributions.) Hence, if $r(m + 1) = r(m)$, then the claim and its proof yield that, conditioned on $Y_m = y$ and everything else revealed so far,

$$Y_{m+1} = \min_{1 \leq j \leq r(m)} (E_j + y) = Y_m + \min_{1 \leq j \leq r(m)} E_j, \quad (3.3)$$

where E_1, E_2, \dots are i.i.d. and $\text{Exp}(1)$. In particular, $Y_{m+1} - Y_m$ is independent of Y_1, \dots, Y_m . Furthermore, (3.2) holds, since $\min_{1 \leq j \leq r(m)} E_j \in \text{Exp}(r(m))$. \square

Let

$$\delta_m := \mathbf{1}\{r(m) = r(m - 1)\} = 1 + r(m - 1) - r(m). \quad (3.4)$$

Lemma 3.3 yields the following representation of Y_m .

Lemma 3.5. *There exists a sequence E_1, E_2, \dots of i.i.d. $\text{Exp}(1)$ random variables such that*

$$Y_m = \sum_{k=1}^m \frac{\delta_k}{r(k)} E_k, \quad m \geq 0. \quad (3.5)$$

Proof. An immediate consequence of Lemma 3.3, with

$$E_k := r(k)(Y_k - Y_{k-1}) = r(k-1)(Y_k - Y_{k-1}) \in \text{Exp}(1) \tag{3.6}$$

when $\delta_k = 1$, and $E_k \in \text{Exp}(1)$ arbitrary but independent of everything else when $\delta_k = 0$. \square

We are now prepared for a theorem giving an exact representation of the sequence N_m .

Theorem 3.6. *The sequence $N_m, m \geq 0$, is given by*

$$N_m = \sum_{k=0}^{m-1} V_k, \tag{3.7}$$

where, conditioned on the sequence $(Y_m)_1^\infty$ given by Lemma 3.5, the random variables V_k are independent with $V_k \in \text{Ge}(e^{-Y_k})$.

Proof. Fix m and condition on Y_1, \dots, Y_m and N_1, \dots, N_m . Each new candidate after N_m has probability $\mathbb{P}(X_n > Y_m \mid Y_m) = e^{-Y_m}$ of exceeding the threshold Y_m , and these events are independent. Hence, still conditioned on the past,

$$N_{m+1} - N_m \in \text{Ge}(e^{-Y_m}). \tag{3.8}$$

Furthermore, still conditioned on the past, this waiting time $N_{m+1} - N_m$ is independent of the value of the next accepted candidate X_{m+1}^* . Hence, the argument in the proof of Lemma 3.3 extends and shows that conditioned on Y_1, \dots, Y_m and N_1, \dots, N_m , the increments $Y_{m+1} - Y_m$ and $N_{m+1} - N_m$ are independent, with the (conditional) distributions given by (3.2) and (3.8).

This implies that conditioned on $(Y_m)_1^\infty$, the increments $V_k := N_{k+1} - N_k$ are independent, with (conditionally) $V_k \in \text{Ge}(e^{-Y_k})$. \square

Remark 3.7. As said above, our choice $X_n \in \text{Exp}(1)$ simplifies the argument, but it is not really essential. An equivalent argument has been used by e.g. [4] with values $X_n \in U(0, 1)$; then one considers the gap $1 - Y_m$ and shows that these gaps can be written as products of independent random variables. Taking (minus) the logarithm of the gap yields a sum of independent random variables (which is more convenient than a product for limit theorems), and that is equivalent to our version with exponentially distributed values X_n .

So far we have given exact formulas, but now we start to approximate in order to obtain simpler formulas. First we approximate the geometric distributions in (3.7) by exponential distributions.

Theorem 3.8. *As $m \rightarrow \infty$, a.s.*

$$N_m \sim \sum_{k=0}^{m-1} e^{Y_k} E'_k, \tag{3.9}$$

where Y_k are given by (3.5) and $E'_k \in \text{Exp}(1)$ are independent of each other and of $(Y_k)_1^\infty$.

Proof. We use continuous time, and assume that candidate n is examined at time τ_n , where the waiting times $\tau_n - \tau_{n-1}$ (with $\tau_0 = 0$) are i.i.d. $\text{Exp}(1)$. In other words, the candidates arrive according to a Poisson process on $[0, \infty)$ with intensity 1. Note that by the law of large numbers, $\tau_n/n \xrightarrow{\text{a.s.}} 1$ as $n \rightarrow \infty$. We assume that the times $(\tau_n)_n$ are independent of the values $(X_n)_n$.

Let $T_m := \tau_{N_m}$ be the time the m -th candidate is accepted. Then, as $m \rightarrow \infty$ and thus $N_m \rightarrow \infty$,

$$\frac{T_m}{N_m} = \frac{\tau_{N_m}}{N_m} \xrightarrow{\text{a.s.}} 1. \tag{3.10}$$

We argue as in the proof of Theorem 3.6. Condition on Y_1, \dots, Y_m and T_1, \dots, T_m for some m . Then, after T_m , candidates arrive according to a Poisson process with intensity 1, and thus candidates with a value $> Y_m$ arrive as a Poisson process with intensity e^{-Y_m} . Consequently, conditioned on the past, the waiting time $T_{m+1} - T_m$ has an exponential waiting time

$$T_{m+1} - T_m \in \text{Exp}(e^{-Y_m}). \tag{3.11}$$

By the same argument as in the proof of Theorem 3.6, this implies that conditioned on $(Y_k)_1^\infty$, the increments $\widehat{V}_k := T_{k+1} - T_k$ are independent with $\widehat{V}_k \in \text{Exp}(e^{-Y_k})$. Define $E'_k := e^{-Y_k} \widehat{V}_k \in \text{Exp}(1)$. Then

$$T_m = \sum_{k=0}^{m-1} \widehat{V}_k = \sum_{k=0}^{m-1} e^{Y_k} E'_k. \tag{3.12}$$

Furthermore, conditioned on $(Y_k)_1^\infty$, the variables E'_k are i.i.d. $\text{Exp}(1)$; hence, $(Y_k)_1^\infty$ and $(E'_k)_0^\infty$ are independent.

Finally, the exact continuous-time representation (3.12) implies the approximation (3.9) by (3.10). □

The results above are valid for any sequence $r(m)$ fulfilling the conditions (1.1). For further approximations, we treat the cases of large and small $r(m)$ separately, in Sections 4–6 and Section 7, respectively.

We define, recalling (3.5),

$$y_m := \mathbb{E} Y_m = \sum_{k=1}^m \frac{\delta_k}{r(k)} = \sum_{\substack{1 \leq k \leq m: \\ r(k)=r(k-1)}} \frac{\delta_k}{r(k)} = \sum_{k=1}^m \frac{1}{r(k)} - \sum_{\ell=2}^{r(m)} \frac{1}{\ell}, \tag{3.13}$$

where the final equality follows because each $\ell \in \{2, \dots, r(m)\}$ equals $r(k)$ for exactly one $k \in \{1, \dots, m\}$ with $r(k) > r(k-1)$.

4 Large $r(m)$

In this section we assume that $r(m)$ is large, i.e.,

$$\sum_{m=1}^{\infty} \frac{1}{r(m)^2} < \infty. \tag{4.1}$$

Lemma 4.1. *Suppose that $\sum_{m=1}^{\infty} r(m)^{-2} < \infty$. Then $Y_m - y_m \xrightarrow{\text{a.s.}} Z$ as $m \rightarrow \infty$, where*

$$Z := \sum_{k=1}^{\infty} \frac{\delta_k}{r(k)} (E_k - 1), \tag{4.2}$$

with, as always, $(E_k)_1^\infty$ are i.i.d. $\text{Exp}(1)$.

Proof. By (3.5) and (3.13),

$$Y_m - y_m = \sum_{k=1}^m \frac{\delta_k}{r(k)} (E_k - 1), \tag{4.3}$$

so $Y_m - y_m$ converges to Z given by (4.2) whenever the latter sum converges. Furthermore, this occurs a.s., since the summands in (4.2) are independent random variables with mean 0 and sum of variances

$$\sum_{k=1}^{\infty} \mathbb{E} \left(\frac{\delta_k}{r(k)} (E_k - 1) \right)^2 \leq \sum_{k=1}^{\infty} \frac{1}{r(k)^2} < \infty. \tag{4.4}$$

□

We let Z denote the sum in (4.2) whenever (4.1) holds.

Remark 4.2. Since the $\text{Exp}(1)$ distribution is infinitely divisible with Lévy measure $x^{-1}e^{-x} dx$, it follows from (4.2) that Z is infinitely divisible with Lévy measure, arguing as in (3.13),

$$\sum_{k=1}^{\infty} \delta_k e^{-r(k)x} \frac{dx}{x} = \left(\sum_{k=1}^{\infty} e^{-r(k)x} - \frac{1}{e^{2x} - e^x} \right) \frac{dx}{x}, \quad x > 0. \tag{4.5}$$

Theorem 4.3. Suppose that $\sum_{m=1}^{\infty} r(m)^{-2} < \infty$. Then, as $m \rightarrow \infty$, a.s.

$$N_m \sim e^Z \sum_{k=0}^{m-1} e^{y_k} E'_k, \tag{4.6}$$

where Z and y_k are given by (4.2) and (3.13), and $E'_k \in \text{Exp}(1)$ are independent of each other and of Z .

Proof. Let $\varepsilon_k := Y_k - y_k - Z$; then a.s. $\varepsilon_k \rightarrow 0$ as $k \rightarrow \infty$ by Lemma 4.1. Furthermore, $Y_k = y_k + Z + \varepsilon_k$, and thus

$$\sum_{k=0}^{m-1} e^{Y_k} E'_k = e^Z \sum_{k=0}^{m-1} e^{\varepsilon_k} e^{y_k} E'_k. \tag{4.7}$$

Hence, the result (4.6) follows from (3.9) and the simple deterministic Lemma 4.4 below, noting that $\sum_{k=0}^{\infty} e^{y_k} E'_k \geq \sum_{k=0}^{\infty} E'_k = \infty$ a.s. □

Lemma 4.4. Suppose that $a_k > 0$, $\sum_{k=0}^{\infty} a_k = \infty$ and $\varepsilon_k \rightarrow 0$ as $k \rightarrow \infty$. Then, as $m \rightarrow \infty$,

$$\frac{\sum_{k=0}^m (1 + \varepsilon_k) a_k}{\sum_{k=0}^m a_k} \rightarrow 1. \tag{4.8}$$

Proof. Let $\eta > 0$. Then there exists K such that if $k > K$, then $|\varepsilon_k| < \eta$. Consequently, for $m > K$,

$$\left| \sum_{k=0}^m \varepsilon_k a_k \right| \leq \sum_{k=0}^K |\varepsilon_k| a_k + \eta \sum_{k=K+1}^m a_k, \tag{4.9}$$

which is less than $2\eta \sum_{k=0}^m a_k$ if m is large enough. This implies (4.8). □

Lemma 4.5. Suppose that

$$\sum_{m=1}^{\infty} \frac{1}{r(m)^2} < \infty. \tag{4.10}$$

Then

$$\sum_{m=0}^{\infty} \frac{e^{2y_m}}{\left(\sum_{k=0}^m e^{y_k} \right)^2} < \infty. \tag{4.11}$$

Proof. Let $m \geq 0$ and let $a(m) := r(\lfloor m/2 \rfloor) \leq \lfloor m/2 \rfloor + 1$. If $0 \leq i < a(m)$, then $i \leq a(m) - 1 \leq m/2$, and thus $m - i \geq m/2$. Hence,

$$y_m - y_{m-i} = \sum_{j=m-i+1}^m \frac{\delta_j}{r(j)} \leq \frac{i}{r(\lceil m/2 \rceil)} \leq \frac{i}{a(m)} \leq 1. \tag{4.12}$$

Consequently,

$$\sum_{k=0}^m e^{y_k} \geq \sum_{i=0}^{a(m)-1} e^{y_{m-i}} \geq \sum_{i=0}^{a(m)-1} e^{y_{m-1}} = a(m)e^{y_{m-1}} \tag{4.13}$$

and thus

$$\sum_{m=0}^{\infty} \frac{e^{2y_m}}{\left(\sum_{k=0}^m e^{y_k}\right)^2} \leq \sum_{m=0}^{\infty} \frac{e^2}{a(m)^2} = 2 \sum_{\ell=0}^{\infty} \frac{e^2}{r(\ell)^2} < \infty. \tag{4.14}$$

□

Define, with y_k given by (3.13),

$$\lambda_m := \sum_{k=0}^{m-1} e^{y_k}, \quad m \geq 0. \tag{4.15}$$

Theorem 4.6. *Suppose that $\sum_{m=1}^{\infty} r(m)^{-2} < \infty$. Then, as $m \rightarrow \infty$,*

$$N_m / \lambda_m \xrightarrow{\text{a.s.}} e^Z, \tag{4.16}$$

where Z and λ_m are given by (4.2) and (4.15).

Proof. Let $W_m := e^{y_m}(E'_m - 1)$ and $b_m := \lambda_{m+1}$. Then $\mathbb{E} W_m = 0$ and Lemma 4.5 shows that $\sum_{m=0}^{\infty} \text{Var}(W_m) / b_m^2 < \infty$. Consequently, see [5, Theorem VII.8.3], $b_m^{-1} \sum_{k=0}^m W_k \xrightarrow{\text{a.s.}} 0$. Hence,

$$\frac{\sum_{k=0}^{m-1} e^{y_k} E'_k}{\lambda_m} = 1 + \frac{\sum_{k=0}^{m-1} e^{y_k} (E'_k - 1)}{\lambda_m} = 1 + \frac{\sum_{k=0}^{m-1} W_k}{b_{m-1}} \xrightarrow{\text{a.s.}} 1, \tag{4.17}$$

and the result follows from (4.6). □

Equivalently, a.s. $N_m \sim \lambda_m e^Z$ as $m \rightarrow \infty$. Hence, N_m grows as the deterministic sequence λ_m , with a random factor (asymptotically independent of m) given by e^Z .

Theorem 4.6 gives the asymptotics (and in particular the asymptotic distribution) of N_m , the number of candidates examined until m have been accepted. By inversion, we obtain corresponding asymptotic results for M_n , the number of accepted candidates when n have been examined. We state one general result as the next theorem. More explicit results require inversion of the function $m \mapsto \lambda_m$, which easily is done under further assumptions; we study an important case in Section 5 below.

Note first that $M_n \xrightarrow{\text{a.s.}} \infty$ as $n \rightarrow \infty$, since for every m , a.s. some future candidate $n > N_m$ will satisfy $X_n > Y_m$ and thus be accepted.

Theorem 4.7. *Suppose that $\sum_{m=1}^{\infty} r(m)^{-2} < \infty$. Then, as $n \rightarrow \infty$,*

$$\lambda_{M_n} / n \xrightarrow{\text{a.s.}} e^{-Z}, \tag{4.18}$$

where Z and λ_m are given by (4.2) and (4.15).

Proof. Since $M_n \rightarrow \infty$ a.s. as $n \rightarrow \infty$, (4.16) and Lemma 4.8 below imply that a.s. $N_{M_n} / \lambda_{M_n} \xrightarrow{\text{a.s.}} e^Z$ and $N_{M_n+1} / \lambda_{M_n} \xrightarrow{\text{a.s.}} e^Z$. Furthermore, the definitions imply $N_{M_n} \leq n < N_{M_n+1}$. Hence, $n / \lambda_{M_n} \xrightarrow{\text{a.s.}} e^Z$. □

Lemma 4.8. *If $r(m) \rightarrow \infty$ as $m \rightarrow \infty$, then $\lambda_{m+1}/\lambda_m \rightarrow 1$.*

Proof. By (4.15) and (4.13),

$$\lambda_{m+1} - \lambda_m = e^{y_m} \leq \frac{e}{a(m)} \sum_{k=0}^m e^{y_k} = \frac{e}{a(m)} \lambda_{m+1}. \tag{4.19}$$

Since $a(m) := r(\lfloor m/2 \rfloor) \rightarrow \infty$ as $m \rightarrow \infty$ by assumption, it follows that $(\lambda_{m+1} - \lambda_m)/\lambda_{m+1} \rightarrow 0$, and thus $\lambda_m/\lambda_{m+1} \rightarrow 1$. \square

5 Roughly linear rank thresholds

As said in the introduction, the strategies ‘hiring above the median’ and ‘the α -percentile rule’ satisfy $r(m) = \alpha m + O(1)$ for some constant $\alpha > 0$, and we stated Theorem 1.2 for this case. Note that $r(m) = \alpha m + O(1)$ implies $r(m)^{-1} = (\alpha m)^{-1} + O(1/m^2)$, and thus

$$\sum_{m=1}^{\infty} |r(m)^{-1} - (\alpha m)^{-1}| < \infty. \tag{5.1}$$

In fact, by the proof below, Theorem 1.2 holds under the weaker assumption (5.1). (For example, if $r(m) = \alpha m + O(m^{1-\eta})$ for some $\eta > 0$.)

We assume throughout this section that (5.1) holds, with $0 < \alpha \leq 1$. We then define

$$\rho := \sum_{m=1}^{\infty} \left(\frac{1}{r(m)} - \frac{1}{\alpha m} \right) \in \mathbb{R}. \tag{5.2}$$

Note that (5.1) implies

$$r(m) \sim \alpha m \quad \text{as } m \rightarrow \infty. \tag{5.3}$$

This is an easy exercise, but for the reader’s convenience we give a proof in Appendix A. Note also that (5.3) implies $\sum_{m=1}^{\infty} r(m)^{-2} < \infty$, so $r(m)$ is large and we can use the results of Section 4; our goal in this section is to use the assumption (5.1) to make the results more explicit.

Lemma 5.1. *Suppose that (5.1) holds for some $\alpha \in (0, 1]$. Then, as $m \rightarrow \infty$,*

$$y_m = (\alpha^{-1} - 1) \log m + (\alpha^{-1} - 1)\gamma - \log \alpha + \rho + 1 + o(1). \tag{5.4}$$

Proof. Let $H_m := \sum_{k=1}^m 1/k$, the m -th harmonic number, and recall that $H_m = \log m + \gamma + o(1)$. Hence, by (3.13) and (5.2),

$$\begin{aligned} y_m &= \sum_{k=1}^m \left(\frac{1}{r(k)} - \frac{1}{\alpha k} \right) + \frac{1}{\alpha} H_m - (H_{r(m)} - 1) \\ &= \rho + \alpha^{-1} (\log m + \gamma) - \log r(m) - \gamma + 1 + o(1). \end{aligned} \tag{5.5}$$

The result (5.4) follows since $\log(r(m)) = \log(\alpha m) + o(1) = \log m + \log \alpha + o(1)$ by (5.3). \square

Lemma 5.2. *Suppose that (5.1) holds for some $\alpha \in (0, 1]$. Then, as $m \rightarrow \infty$,*

$$\lambda_m \sim e^{(\alpha^{-1}-1)\gamma+\rho+1} m^{1/\alpha}. \tag{5.6}$$

Proof. Lemma 5.1 implies, where $o(1) \rightarrow 0$ as $k \rightarrow \infty$,

$$\lambda_m := \sum_{k=0}^{m-1} e^{y_k} = e^{(\alpha^{-1}-1)\gamma+\rho+1} \alpha^{-1} \sum_{k=0}^{m-1} k^{\alpha^{-1}-1} (1 + o(1)), \tag{5.7}$$

and the result (5.6) follows. \square

We are prepared to prove the convergence (1.5) in Theorem 1.2.

Lemma 5.3. *Suppose that (5.1) holds for some $\alpha \in (0, 1]$. Then, as $n \rightarrow \infty$,*

$$M_n/n^\alpha \xrightarrow{\text{a.s.}} W := e^{(\alpha-1)\gamma-\alpha\rho-\alpha} e^{-\alpha Z}, \tag{5.8}$$

with ρ and Z given by (5.2) and (4.2).

Proof. Theorem 4.7 and Lemma 5.2 imply that a.s.

$$M_n^{1/\alpha}/n \rightarrow e^{-(\alpha^{-1}-1)\gamma-\rho-1} e^{-Z}, \tag{5.9}$$

and thus (5.8) follows. □

We proceed to compute moments of W .

Lemma 5.4. *Suppose that (5.1) holds for some $\alpha \in (0, 1]$. If $-\infty < u < \alpha$, then*

$$\mathbb{E} e^{uZ} = e^{-u\rho-u+u(1-\alpha^{-1})\gamma} \frac{\Gamma(1-u/\alpha)}{\Gamma(2-u)} \prod_{k=1}^{\infty} \frac{1-u/\alpha k}{1-u/r(k)} < \infty. \tag{5.10}$$

Hence, at least for $-1 < s < \infty$,

$$\mathbb{E} W^s = \frac{\Gamma(s+1)}{\Gamma(s\alpha+2)} \prod_{k=1}^{\infty} \frac{1+s/k}{1+s\alpha/r(k)}. \tag{5.11}$$

Proof. It is elementary that $\mathbb{E} e^{uE_k} = 1/(1-u)$ for $u < 1$. Hence, (4.2) implies

$$\mathbb{E} e^{uZ} = \prod_{k \geq 1: \delta_k=1} \frac{e^{-u/r(k)}}{1-u/r(k)}. \tag{5.12}$$

Arguing as in (3.13), this can be written

$$\mathbb{E} e^{uZ} = \prod_{k=1}^{\infty} \frac{e^{-u/r(k)}}{1-u/r(k)} \prod_{\ell=2}^{\infty} \frac{1-u/\ell}{e^{-u/\ell}} \tag{5.13}$$

where the products are absolutely convergent as a consequence of (5.3) and simple estimates. Using (5.2), we rewrite this as

$$\mathbb{E} e^{uZ} = e^{-u\rho} \prod_{k=1}^{\infty} \frac{e^{-u/\alpha k}}{1-u/\alpha k} \prod_{k=1}^{\infty} \frac{1-u/\alpha k}{1-u/r(k)} \prod_{\ell=2}^{\infty} \frac{1-u/\ell}{e^{-u/\ell}}. \tag{5.14}$$

The standard product formula for the Gamma function [17, (5.8.2)] can be written $\prod_{k=1}^{\infty} e^{z/k}/(1+z/k) = ze^{\gamma z}\Gamma(z) = e^{\gamma z}\Gamma(z+1)$, and thus (5.14) yields

$$\mathbb{E} e^{uZ} = e^{-u\rho} e^{-u\alpha^{-1}\gamma} \Gamma(1-u/\alpha) \frac{e^{-u}}{1-u} e^{u\gamma} \Gamma(1-u)^{-1} \prod_{k=1}^{\infty} \frac{1-u/\alpha k}{1-u/r(k)},$$

which simplifies to (5.10).

Finally, (5.11) follows from (5.8) and (5.10), with $u = -\alpha s$. □

Remark 5.5. Let $r_* := \inf\{r(k) : \delta_k = 1\} \geq 1$ be the first repeated value in the sequence $r(0), r(1), \dots$. Then the proof shows that (5.10) holds for $\alpha \leq u < r_*$ too, with the right hand side interpreted by continuity when necessary, and that $\mathbb{E} e^{uZ} = \infty$ for $u \geq r_*$. Consequently, (5.11) holds for $s \geq -r_*/\alpha$. Furthermore, by analytic continuation, or by the proof above, (5.11) extends to complex s with $\text{Re } s > -r_*/\alpha$. However, $\mathbb{E} W^{-r_*/\alpha} = \infty$.

Next we bound moments of M_n , using a series of lemmas. Recall $T_m := \tau_{N_m}$ from the proof of Theorem 3.8. We tacitly continue to assume (5.1).

Lemma 5.6. *For every integer $K \geq 1$ there exists a constant C_K such that if $m \geq 2K + 2$, then*

$$\mathbb{E}(T_m/m^{1/\alpha})^{-K} \leq C_K. \tag{5.15}$$

Proof. For convenience, we first consider T_{2m} . By (3.12),

$$T_{2m} \geq \sum_{k=m}^{2m-1} e^{Y_k} E'_k \geq e^{Y_m} \sum_{k=m}^{2m-1} E'_k =: e^{Y_m} U, \tag{5.16}$$

where $U := \sum_{k=m}^{2m-1} E'_k$ has a Gamma distribution $\Gamma(m)$ and is independent of Y_m . Consequently,

$$\mathbb{E} T_{2m}^{-K} \leq \mathbb{E} e^{-K(Y_m - y_m)} e^{-Ky_m} \mathbb{E} U^{-K}. \tag{5.17}$$

We estimate the three factors on the right-hand side separately.

For the first factor, by (4.3) and (5.12), for any real u ,

$$\mathbb{E} e^{u(Y_m - y_m)} = \prod_{k=m: \delta_k=1} \frac{e^{-u/r(k)}}{1 - u/r(k)} \leq \mathbb{E} e^{uZ}, \tag{5.18}$$

since each factor in (5.12) is ≥ 1 (by the explicit form or by Jensen's inequality). In particular,

$$\mathbb{E} e^{-K(Y_m - y_m)} \leq \mathbb{E} e^{-KZ} = C_K, \tag{5.19}$$

using (5.10) to see that $\mathbb{E} e^{-KZ} < \infty$.

For the second factor in (5.17), Lemma 5.1 yields

$$e^{-Ky_m} = m^{-K(\alpha^{-1}-1)} C_K e^{o(1)} \leq m^{-K(\alpha^{-1}-1)} C_K. \tag{5.20}$$

Finally, for the third factor, we use the fact that $U \in \Gamma(m)$ and compute

$$\mathbb{E} U^{-K} = \frac{1}{\Gamma(m)} \int_0^\infty x^{-K} x^{m-1} e^{-x} dx = \frac{\Gamma(m-K)}{\Gamma(m)} \leq C_K m^{-K}, \tag{5.21}$$

for $m \geq K + 1$.

Combining (5.19), (5.20) and (5.21) with (5.17), we find that if $m \geq K + 1$, then

$$\mathbb{E} T_{2m}^{-K} \leq C_K m^{-K/\alpha}. \tag{5.22}$$

This proves (5.15) for even $m \geq 2K + 2$, and the case of odd m follows because $T_{2m+1} \geq T_{2m}$. \square

Lemma 5.7. *For every $K > 0$ there exists a constant C_K such that*

$$\mathbb{E}(N_m/m^{1/\alpha})^{-K} \leq C_K, \quad m \geq 1. \tag{5.23}$$

Proof. We may assume that K is an integer (by Lyapunov's inequality). Furthermore, $N_m \geq m \geq 1$, and thus $\mathbb{E} N_m^{-K} \leq 1$. It follows that (5.23) holds trivially for $m < 4K + 2$, so we may assume $m \geq 4K + 2$. In this case, we use the Cauchy-Schwarz inequality and obtain by Lemma 5.6

$$\left(\mathbb{E} \left(\frac{N_m}{m^{1/\alpha}} \right)^{-K} \right)^2 \leq \mathbb{E} \left(\frac{T_m}{m^{1/\alpha}} \right)^{-2K} \mathbb{E} \left(\frac{N_m}{T_m} \right)^{-2K} \leq C_K \mathbb{E} \left(\frac{T_m}{N_m} \right)^{2K}. \tag{5.24}$$

Furthermore, conditioned on $N_m = n$, $T_m = \tau_n$ is the sum of n i.i.d. waiting times $E_i'' \in \text{Exp}(1)$. Since E_i'' have moments of all orders, the law of large numbers holds with moment convergence [7, Theorem 6.10.2], and thus $\mathbb{E} \tau_n^{2K} / n^{2K} \rightarrow 1$ as $n \rightarrow \infty$, and therefore $\mathbb{E} \tau_n^{2K} / n^{2K} \leq C_K$ for all $n \geq 1$. Consequently,

$$\mathbb{E} \left(\left(\frac{T_m}{N_m} \right)^{2K} \mid N_m = n \right) = \frac{\mathbb{E} \tau_n^{2K}}{n^{2K}} \leq C_K \tag{5.25}$$

and thus

$$\mathbb{E} \left(\frac{T_m}{N_m} \right)^{2K} = \mathbb{E} \left(\left(\frac{T_m}{N_m} \right)^{2K} \mid N_m \right) \leq C_K. \tag{5.26}$$

The result follows by (5.24) and (5.26). □

Lemma 5.8. *For every $K > 0$ there exists a constant C_K such that*

$$\mathbb{E}(M_n/n^\alpha)^K \leq C_K, \quad n \geq 1. \tag{5.27}$$

Proof. Let $x > 0$ and $n \geq 1$, and let $m := \lceil xn^\alpha \rceil$. If $M_n \geq xn^\alpha$, then $M_n \geq m$, and thus, by (3.1), $N_m \leq n \leq x^{-1/\alpha} m^{1/\alpha}$. Consequently, for any $L > 0$, by Markov's inequality and Lemma 5.7,

$$\mathbb{P}(M_n \geq xn^\alpha) \leq \mathbb{P}(N_m \leq x^{-1/\alpha} m^{1/\alpha}) \leq x^{-L/\alpha} \mathbb{E}(N_m/m^{1/\alpha})^{-L} \leq C_L x^{-L/\alpha}. \tag{5.28}$$

Choosing $L := (K + 1)\alpha$, we obtain $\mathbb{P}(M_n \geq xn^\alpha) \leq C_K x^{-(K+1)}$, which implies (5.27). □

Proof of Theorem 1.2. We have shown the a.s. convergence (1.5) in Lemma 5.3, and moment convergence follows from this and the uniform estimate in Lemma 5.8. Finally, the moments of W are computed in Lemma 5.4. □

Although perhaps of less interest, we show further that the moment convergence in Theorem 1.2 holds also for some, but not all, $s < 0$. Let $r_* \geq 1$ be as in Remark 5.5.

Lemma 5.9. *For every real $u < r_*$, there exists a constant C_u such that*

$$\mathbb{E}(N_m/m^{1/\alpha})^u \leq C_u, \quad m \geq 1. \tag{5.29}$$

Proof. The case $u < 0$ is Lemma 5.7, and $u = 0$ is trivial, so we may assume $u > 0$. We consider again first $T_m = \tau_{N_m}$.

Assume first that $r(1) = 2$, so that $r_* \geq 2$. It then suffices to prove (5.29) for $1 \leq u < r_*$, so we assume this. Recall from Remark 5.5 that then $\mathbb{E} e^{uZ} < \infty$. Hence, (3.12), Minkowski's inequality, (5.18), (4.15) and (5.6) yield, with $\|X\|_u := (\mathbb{E} |X|^u)^{1/u}$,

$$\|T_m\|_u \leq \sum_{k=0}^{m-1} \|e^{Y_k} E'_k\|_u = \sum_{k=0}^{m-1} e^{y_k} (\mathbb{E} e^{u(Y_k - y_k)})^{1/u} \|E'_k\|_u \tag{5.30}$$

$$\leq C_u (\mathbb{E} e^{uZ})^{1/u} \sum_{k=0}^{m-1} e^{y_k} = C_u \lambda_m \leq C_u m^{1/\alpha}. \tag{5.31}$$

In other words,

$$\mathbb{E} T_m^u \leq C_u m^{u/\alpha}, \quad m \geq 1. \tag{5.32}$$

The argument above fails if $r(1) = 1$, since then $r_* = 1$ and $\mathbb{E} e^{uZ} = \infty$ for every $u \geq 1$, see again Remark 5.5. (And we need $u \geq 1$ in order to use Minkowski's inequality.) In this case, let $k_0 := \min\{k : r(k) = 2\}$ and consider

$$\tilde{T}_m := \sum_{k=k_0}^{m-1} e^{Y_k - Y_{k_0}} E'_k, \quad m \geq 1. \tag{5.33}$$

We have, cf. (4.3) and (5.12),

$$\mathbb{E} e^{Y_k - Y_{k_0} - (y_k - y_{k_0})} = \prod_{k_0 < j \leq m: \delta_j = 1} \frac{e^{-1/r(j)}}{1 - 1/r(j)} \leq \prod_{j=k_0}^{\infty} \frac{e^{-1/r(j)}}{1 - 1/r(j)} < \infty, \quad (5.34)$$

and thus, using (4.15) and (5.6) as above,

$$\mathbb{E} \tilde{T}_m := \sum_{k=k_0}^{m-1} \mathbb{E} e^{Y_k - Y_{k_0}} \leq C \sum_{k=k_0}^{m-1} e^{y_k - y_{k_0}} \leq C \lambda_m \leq C m^{1/\alpha}. \quad (5.35)$$

Consequently, for $0 < u < 1 = r_*$, using $T_m = T_{k_0} + e^{Y_{k_0}} \tilde{T}_m$ and the subadditivity of $x \mapsto x^u$,

$$\begin{aligned} \mathbb{E} T_m^u &\leq \mathbb{E} T_{k_0}^u + \mathbb{E} e^{uY_{k_0}} \mathbb{E} \tilde{T}_m^u \leq \sum_{k=0}^{k_0-1} \mathbb{E} e^{uY_k} \mathbb{E} (E'_k)^u + \mathbb{E} e^{uY_{k_0}} (\mathbb{E} \tilde{T}_m)^u \\ &\leq C_u + C_u m^u. \end{aligned} \quad (5.36)$$

Hence, (5.32) holds in this case too.

Finally, by the law of large numbers, $\mathbb{P}(\tau_n > n/2) \rightarrow 1$ as $n \rightarrow \infty$, and thus $\mathbb{P}(\tau_n > n/2) \geq c$ for every $n \geq 1$. Hence,

$$\mathbb{E}(T_m^u \mid N_m) \geq c(N_m/2)^u = c_u N_m^u \quad (5.37)$$

and thus

$$\mathbb{E} T_m^u \geq c_u \mathbb{E} N_m^u. \quad (5.38)$$

Consequently, (5.29) follows from (5.32). \square

Lemma 5.10. For every $s > -r_*/\alpha$,

$$\mathbb{E}(M_n/n^\alpha)^s \leq C_s. \quad (5.39)$$

Proof. By Lemma 5.8, it suffices to consider $s < 0$. Then, let $-r_*/\alpha < t < s$ and $u := -\alpha t \in (0, r_*)$. We argue as in the proof of Lemma 5.8 with minor modifications. Let $x > 0$ and let $m := \lfloor xn^\alpha \rfloor$. Then, by Lemma 5.9,

$$\begin{aligned} \mathbb{P}(M_n + 1 < xn^\alpha) &\leq \mathbb{P}(M_n < m) = \mathbb{P}(N_m > n) \leq n^{-u} \mathbb{E} N_m^u \\ &\leq C_u m^{u/\alpha} n^{-u} \leq C_u x^{u/\alpha} = C_u x^{-t}. \end{aligned} \quad (5.40)$$

It follows that, since trivially $1 + M_n \leq 2M_n$,

$$\mathbb{E} M_n^s \leq 2^{-s} \mathbb{E}(M_n + 1)^s \leq C_s n^{s\alpha}. \quad (5.41)$$

\square

Theorem 5.11. The moment convergence (1.6) holds for every real $s > -r_*/\alpha$, and more generally for every complex s with $\operatorname{Re} s > -r_*/\alpha$, with a finite limit. On the other hand, if $s \leq -r_*/\alpha$, then $\mathbb{E} M_n^s \rightarrow \infty$.

Proof. Note that the right-hand side of (1.6) is finite for $\operatorname{Re} s > -r_*/\alpha$ and an analytic function of s in that half-plane, see Remark 5.5.

Lemma 5.10 implies uniform integrability of $(M_n/n^\alpha)^s$ for every $s \in (-r_*/\alpha, 0)$, and thus by Theorem 1.2 for every real $s > -r_*/\alpha$, which implies (1.6) for $\operatorname{Re} s > -r_*/\alpha$. On the other hand, if $s \leq -r_*/\alpha$, then $\mathbb{E} W^s = \infty$, see Remark 5.5, and thus $\mathbb{E} M^s/n^{\alpha s} \rightarrow \infty$ by (1.5) and Fatou's lemma. \square

6 Linear-periodic rank thresholds

In this section we restrict $r(m)$ further to an important case where we can evaluate the moments in (5.11) explicitly. We assume that

$$r(m + q) = r(m) + \nu, \quad m \geq 1, \tag{6.1}$$

for some integers ν and q with $1 \leq \nu \leq q$. Hence, with $\alpha := \nu/q \in (0, 1]$, $r(m) - \alpha m$ is periodic. In particular, $r(m) = \alpha m + O(1)$, and thus the results of Sections 4–5 apply. Moreover, it is obvious from (1.2)–(1.3) that ‘hiring above the median’ satisfies (6.1) with $\nu = 1, q = 2$, and that ‘the α -percentile rule’ satisfies (6.1) whenever $\alpha = \nu/q$ is rational.

When (6.1) holds, the asymptotic moments can be evaluated explicitly; for convenience in applications, we give several equivalent formulas.

Theorem 6.1. *Suppose that (6.1) holds, with $1 \leq \nu \leq q$, and let $\alpha := \nu/q$. Then, for $0 \leq s < \infty$, as $n \rightarrow \infty$,*

$$\mathbb{E} M_n^s / n^{\alpha s} \rightarrow \mathbb{E} W^s = \frac{\Gamma(s + 1)}{\Gamma(s\alpha + 2)} \prod_{i=1}^q \frac{\Gamma(\frac{i}{q}) \Gamma(\frac{s}{q} + \frac{r(i)}{\nu})}{\Gamma(\frac{r(i)}{\nu}) \Gamma(\frac{s}{q} + \frac{i}{q})} \tag{6.2}$$

$$= \frac{q^s}{\Gamma(s\alpha + 2)} \prod_{i=1}^q \frac{\Gamma(\frac{s}{q} + \frac{r(i)}{\nu})}{\Gamma(\frac{r(i)}{\nu})} \tag{6.3}$$

$$= \frac{q^s}{\nu^{\alpha s}} \prod_{i=1}^q \frac{\Gamma(\frac{s}{q} + \frac{r(i)}{\nu})}{\Gamma(\frac{r(i)}{\nu})} \prod_{j=2}^{\nu+1} \frac{\Gamma(\frac{j}{\nu})}{\Gamma(\frac{s}{q} + \frac{j}{\nu})} \tag{6.4}$$

$$= \frac{q^{s+1}}{\nu^{\alpha s}(\nu s + q)} \prod_{i=1}^q \frac{\Gamma(\frac{s}{q} + \frac{r(i)}{\nu})}{\Gamma(\frac{r(i)}{\nu})} \prod_{j=1}^{\nu} \frac{\Gamma(\frac{j}{\nu})}{\Gamma(\frac{s}{q} + \frac{j}{\nu})}. \tag{6.5}$$

Proof. Theorem 1.2 applies and shows (1.6), and it remains only to compute the infinite product there. Standard manipulations yield, for every $s \geq 0$, writing $k = jq + i$ and noting that $r(jq + i) = j\nu + r(i)$,

$$\begin{aligned} \prod_{k=1}^{\infty} \frac{1 + s/k}{1 + s\alpha/r(k)} &= \prod_{k=1}^{\infty} \frac{(k + s)r(k)}{k(r(k) + s\alpha)} = \prod_{i=1}^q \prod_{j=0}^{\infty} \frac{(jq + i + s)r(jq + i)}{(jq + i)(r(jq + i) + s\alpha)} \\ &= \prod_{i=1}^q \prod_{j=0}^{\infty} \frac{(jq + i + s)(j\nu + r(i))}{(jq + i)(j\nu + r(i) + s\alpha)} = \prod_{i=1}^q \prod_{j=0}^{\infty} \frac{(j + \frac{i+s}{q})(j + \frac{r(i)}{\nu})}{(j + \frac{i}{q})(j + \frac{r(i)}{\nu} + \frac{s}{q})} \\ &= \prod_{i=1}^q \lim_{n \rightarrow \infty} \frac{\Gamma(n + \frac{i+s}{q}) \Gamma(n + \frac{r(i)}{\nu})}{\Gamma(\frac{i+s}{q}) \Gamma(\frac{r(i)}{\nu})} \frac{\Gamma(\frac{i}{q}) \Gamma(\frac{r(i)}{\nu} + \frac{s}{q})}{\Gamma(n + \frac{i}{q}) \Gamma(n + \frac{r(i)}{\nu} + \frac{s}{q})} \\ &= \prod_{i=1}^q \frac{\Gamma(\frac{i}{q}) \Gamma(\frac{r(i)}{\nu} + \frac{s}{q})}{\Gamma(\frac{i+s}{q}) \Gamma(\frac{r(i)}{\nu})}. \end{aligned}$$

The result (6.2) follows. Furthermore, by the Gauss multiplication formula for the Gamma function [17, (5.5.6)],

$$\Gamma(s + 1) = (2\pi)^{(1-q)/2} q^{s+1/2} \prod_{k=0}^{q-1} \Gamma\left(\frac{s}{q} + \frac{1}{q} + \frac{k}{q}\right), \tag{6.6}$$

$$\Gamma(s\alpha + 2) = (2\pi)^{(1-\nu)/2} \nu^{s\alpha+3/2} \prod_{k=0}^{\nu-1} \Gamma\left(\frac{s}{q} + \frac{2}{\nu} + \frac{k}{\nu}\right), \tag{6.7}$$

which together with the special case $s = 0$ in (6.6) and (6.7) easily yield (6.3)–(6.5). \square

See Examples 8.1, 8.2 and 8.5.

Example 6.2. Taking $s = q$ in (6.3), we see that the q -th moment has the rational value

$$\mathbb{E} W^q = \frac{q^q}{\Gamma(\nu + 2)} \prod_{i=1}^q \frac{r(i)}{\nu} = \frac{q^q}{\nu^q(\nu + 1)!} \prod_{i=1}^q r(i). \tag{6.8}$$

Remark 6.3. The result extends to all complex s such that $\operatorname{Re} s > -q/\nu$ (and more generally to $\operatorname{Re} s > -r_*/\alpha$); in particular to imaginary s , which gives the characteristic function of $\log W$, see Remark 5.5 and Theorem 5.11.

Remark 6.4. Positive random variables with moments that can be expressed as a fraction of products of Gamma functions as in (6.2) are studied in e.g. [12]. In particular, [12, Theorems 5.4 and 6.1] imply that W has an infinitely differentiable density function $f_W(x)$ on $[0, \infty)$, with the asymptotic

$$f_W(x) \sim C_2 x^{c_1-1} e^{-c_2 x^{1/(1-\alpha)}}, \tag{6.9}$$

where the positive constants C_2, c_1, c_2 can be expressed explicitly in ν, q and $r(1), \dots, r(q)$. This density f_W is of a type known as H -function, see [12, Addendum].

7 Small $r(m)$

In this section we develop the results in the case of small $r(m)$, i.e., $\sum_m r(m)^{-2} = \infty$. However, we begin with some results holding in general, although their main interest is for the case of small $r(m)$.

Let, recalling (3.5),

$$\sigma_m^2 := \operatorname{Var} Y_m = \sum_{k=1}^m \frac{\delta_k}{r(k)^2} \tag{7.1}$$

and, as a simpler approximation,

$$\hat{\sigma}_m^2 := \sum_{k=1}^m \frac{1}{r(k)^2}. \tag{7.2}$$

We have, by the argument in (3.13),

$$\sigma_m^2 := \hat{\sigma}_m^2 - \sum_{\ell=2}^{r(m)} \frac{1}{\ell^2} = \hat{\sigma}_m^2 - O(1). \tag{7.3}$$

Hence, the condition that $r(m)$ is small can be written in three equivalent forms:

$$\sum_m r(m)^{-2} = \infty \iff \hat{\sigma}_m^2 \rightarrow \infty \iff \sigma_m^2 \rightarrow \infty \tag{7.4}$$

Furthermore, when this holds, then $\sigma_m \sim \hat{\sigma}_m$.

For convenience, define

$$A_m := \sum_{k=0}^m e^{Y_k} E'_k, \tag{7.5}$$

$$B_m := \sum_{k=0}^m e^{Y_k}. \tag{7.6}$$

Hence, Theorem 3.8 says that a.s. $N_m \sim A_{m-1}$ as $n \rightarrow \infty$.

Lemma 7.1. For any sequence $r(m)$,

$$B_m = r(m)e^{Y_m + O_p(1)}, \quad m \geq 1, \tag{7.7}$$

i.e.,

$$\log B_m - \log r(m) - Y_m = O_p(1). \tag{7.8}$$

Proof. Let $m \geq 1$ and let $m_1 := m - \lceil r(m)/2 \rceil \geq 0$. Then, by (7.6),

$$B_m e^{-Y_m} \geq \sum_{k=m_1}^m e^{Y_k - Y_m} \geq (m - m_1) e^{Y_{m_1} - Y_m} \geq \frac{r(m)}{2} e^{Y_{m_1} - Y_m}. \tag{7.9}$$

If $i < m - m_1 = \lceil r(m)/2 \rceil$, then $i \leq \lfloor r(m)/2 \rfloor$ and thus

$$r(m - i) \geq r(m) - i \geq r(m) - \lfloor r(m)/2 \rfloor = \lceil r(m)/2 \rceil. \tag{7.10}$$

Hence,

$$\begin{aligned} \mathbb{E} |Y_m - Y_{m_1}| &= \mathbb{E}(Y_m - Y_{m_1}) = y_m - y_{m_1} = \sum_{k=m_1+1}^m \frac{\delta_k}{r(k)} \\ &= \sum_{i=0}^{m-m_1-1} \frac{\delta_{m-i}}{r(m-i)} \leq (m - m_1) \cdot \frac{1}{\lceil r(m)/2 \rceil} = 1. \end{aligned} \tag{7.11}$$

Thus $Y_m - Y_{m_1} = O_p(1)$, which by (7.9) yields the lower bound

$$B_m e^{-Y_m} \geq r(m) e^{-(Y_m - Y_{m_1}) - \log 2} = r(m) e^{O_p(1)}. \tag{7.12}$$

In the other direction, we note that $\delta_k = 0$ for exactly $r(m) - 1$ values of $k \leq m$. Hence, if $r(m) \leq j \leq m$, then $\delta_k = 1$ for more that $j - r(m)$ values of $k \in [m - j + 1, m]$. For each such k ,

$$\mathbb{E} e^{-E_k/r(k)} = (1 + 1/r(k))^{-1} \leq (1 + 1/r(m))^{-1}. \tag{7.13}$$

Consequently, by (3.5),

$$\mathbb{E} e^{Y_{m-j} - Y_m} = \prod_{\substack{m-j < k \leq m \\ \delta_k = 1}} \mathbb{E} e^{-E_k/r(k)} \leq \left(\frac{1}{1 + 1/r(m)} \right)^{j - r(m)}. \tag{7.14}$$

In other words, if $r(m) \leq j \leq m$, then

$$\mathbb{E} e^{Y_{m-j} - Y_m} \leq (1 + 1/r(m))^{r(m) - j} \leq e(1 + 1/r(m))^{-j}, \tag{7.15}$$

and the same estimate holds trivially if $0 \leq j < r(m)$ too. Consequently,

$$\mathbb{E}(B_m e^{-Y_m}) \leq \sum_{j=0}^{\infty} e(1 + 1/r(m))^{-j} = e(r(m) + 1) \leq 2er(m) \tag{7.16}$$

and thus, noting that $\log_+ O_p(1) = O_p(1)$,

$$B_m e^{-Y_m} = r(m) O_p(1) \leq r(m) e^{O_p(1)}. \tag{7.17}$$

The result follows by combining (7.12) and (7.17). □

Lemma 7.2. For any sequence $r(m)$,

$$A_m = B_m e^{O_p(1)} = r(m) e^{Y_m + O_p(1)}. \tag{7.18}$$

Proof. First condition on the entire sequence $(Y_k)_1^\infty$. Then, by (7.5)–(7.6),

$$\begin{aligned} \mathbb{E}\left((A_m - B_m)^2 \mid (Y_k)_1^\infty\right) &= \mathbb{E}\left(\left(\sum_{k=0}^m e^{Y_k}(E'_k - 1)\right)^2 \mid (Y_k)_1^\infty\right) \\ &= \sum_{k=0}^m e^{2Y_k} \leq \sum_{k=0}^m e^{Y_k + Y_m} = B_m e^{Y_m} \end{aligned} \tag{7.19}$$

and thus

$$\mathbb{E}\left(\left(\frac{A_m}{B_m} - 1\right)^2 \mid (Y_k)_1^\infty\right) \leq B_m^{-1} e^{Y_m}. \tag{7.20}$$

Using (7.9), with $m_1 := m - \lceil r(m)/2 \rceil$ as above, and taking the expectation, we obtain

$$\mathbb{E}\left(\frac{A_m}{B_m} - 1\right)^2 \leq \mathbb{E}(B_m^{-1} e^{Y_m}) \leq \frac{2}{r(m)} \mathbb{E} e^{Y_m - Y_{m_1}}. \tag{7.21}$$

Furthermore, by (3.5) and (7.10), if $r(m) \geq 3$,

$$\begin{aligned} \mathbb{E} e^{Y_m - Y_{m_1}} &\leq \prod_{k=m_1+1}^m \mathbb{E} e^{E'_k/r(k)} \leq \left(1 - \frac{1}{r(m_1 + 1)}\right)^{-(m - m_1)} \\ &\leq \left(1 - \frac{1}{\lceil r(m)/2 \rceil}\right)^{-\lceil r(m)/2 \rceil} \leq 4. \end{aligned} \tag{7.22}$$

Consequently, if $r(m) \rightarrow \infty$ as $m \rightarrow \infty$, then (7.21) and (7.22) imply $\mathbb{E}(A_m/B_m - 1)^2 \rightarrow 0$, and thus $A_m/B_m \xrightarrow{P} 1$, which is the same as $\log(A_m/B_m) = o_p(1)$, or $A_m = B_m e^{o_p(1)}$. In particular, $A_m = B_m e^{O_p(1)}$, and (7.18) follows in this case by Lemma 7.1.

On the other hand, if $\sup_m r(m) < \infty$, we note that the right-hand side of (7.20) is ≤ 1 , and thus taking the expectation yields

$$\mathbb{E}\left(\frac{A_m}{B_m} - 1\right)^2 \leq \mathbb{E}(B_m^{-1} e^{Y_m}) \leq 1, \tag{7.23}$$

which implies that $A_m/B_m = O_p(1)$ and thus, using (7.7),

$$A_m \leq B_m e^{O_p(1)} = r(m) e^{Y_m + O_p(1)}. \tag{7.24}$$

Furthermore, assuming $r(m) = O(1)$ and thus $\log r(m) = O(1)$,

$$A_m \geq e^{Y_m} E'_m = e^{Y_m + O_p(1)} = r(m) e^{Y_m + O_p(1)}. \tag{7.25}$$

The result in this case (bounded $r(m)$) follows from (7.24) and (7.25), together with (7.7) again. \square

Theorem 7.3. For any sequence $r(m)$,

$$N_m = r(m) e^{Y_m + O_p(1)}. \tag{7.26}$$

Proof. Theorem 3.8 yields $N_m/A_{m-1} \xrightarrow{\text{a.s.}} 1$, and thus $N_m/A_{m-1} \xrightarrow{P} 1$, which is the same as $\log(N_m/A_{m-1}) = o_p(1)$. Hence, Lemma 7.2 yields

$$\log N_m = \log A_{m-1} + o_p(1) = \log r(m-1) + Y_{m-1} + O_p(1). \tag{7.27}$$

Furthermore, $\log r(m-1) = \log r(m) + O(1)$ and $Y_m - Y_{m-1} = \frac{\delta_m}{r(m)} E'_m = O_p(1)$. Hence, (7.26) follows from (7.27). \square

Theorem 7.3 holds for any sequence $r(m)$, but if $r(m)$ is large, then Theorem 4.6 gives a sharper result. In the remainder of this section we assume that $r(m)$ is small.

Lemma 7.4. *Suppose that $\sum_m r(m)^{-2} = \infty$. Then, as $m \rightarrow \infty$,*

$$\frac{Y_m - y_m}{\sigma_m} \xrightarrow{d} N(0, 1). \tag{7.28}$$

Proof. An immediate consequence of (4.3) and the central limit theorem with Lyapunov's condition [7, Theorem 7.2.2], using the estimate

$$\sum_{k=1}^m \mathbb{E} \left| \frac{\delta_k}{r(k)} (E_k - 1) \right|^3 = C \sum_{k=1}^m \frac{\delta_k}{r(k)^3} \leq C \sigma_m^2 = o(\sigma_m^3). \tag{7.29}$$

□

Theorem 7.5. *Suppose that $\sum_m r(m)^{-2} = \infty$. Then*

$$\frac{\log N_m - (y_m + \log r(m))}{\hat{\sigma}_m} \xrightarrow{d} N(0, 1). \tag{7.30}$$

Proof. Since $\hat{\sigma}_m \rightarrow \infty$ by the assumption and (7.4), Theorem 7.3 yields

$$\log N_m - (y_m + \log r(m)) = Y_m + O_p(1) - y_m = Y_m - y_m + o_p(\hat{\sigma}_m)$$

and thus

$$\frac{\log N_m - (y_m + \log r(m))}{\hat{\sigma}_m} = \frac{\sigma_m}{\hat{\sigma}_m} \cdot \frac{Y_m - y_m}{\sigma_m} + o_p(1). \tag{7.31}$$

Since (7.3) implies $\sigma_m/\hat{\sigma}_m \rightarrow 1$, the result follows from Lemma 7.4 and the standard Cramér–Slutsky theorem [7, Theorem 5.11.4]. □

The asymptotic distribution of N_m is thus log-normal, for any small sequence $r(m)$. Under weak regularity assumptions, this can be inverted to yield asymptotic normality of M_n . For convenience, we will assume that the sequence $r(m)$ is regularly varying, see e.g. [3, §1.9]. We define,

$$\hat{y}_m := \sum_{k=1}^m \frac{1}{r(k)}, \tag{7.32}$$

and see by (3.13) that

$$\hat{y}_m - y_m = \sum_{\ell=2}^{r(m)} \frac{1}{\ell} = \log r(m) + O(1). \tag{7.33}$$

Lemma 7.6. *Suppose that $\sum_m r(m)^{-2} = \infty$, and that $r(m)$ is regularly varying.*

(i) *Let $\beta(m) := r(m)\hat{\sigma}_m$ as in (1.8). Then, for all $m \geq 1$,*

$$\beta(m) \geq m^{1/2}, \tag{7.34}$$

$$\beta(m) \leq Cm^{0.51} = o(m). \tag{7.35}$$

(ii) *Suppose further that $m = m_n$ and $\tilde{m} = m_n$ are two sequences such that, as $n \rightarrow \infty$, $\tilde{m} \sim m$. Then,*

$$r(\tilde{m}) \sim r(m), \tag{7.36}$$

$$\hat{\sigma}_{\tilde{m}} \sim \hat{\sigma}_m. \tag{7.37}$$

Proof. Suppose that $r(m)$ is regularly varying of index κ . Then $r(m) = m^{\kappa+o(1)}$, as a consequence of [3, Theorem 1.9.7 or Theorem 1.5.4]. Hence, the assumption $\sum_m r(m)^{-2} = \infty$ implies $\kappa \leq 0.5$. Hence, $r(m) = O(m^{0.5+o(1)})$, and in particular,

$$r(m) = O(m^{0.51}) = o(m). \tag{7.38}$$

Since $\kappa < 0.51$, we have $r(m)/r(k) \leq C(m/k)^{0.51}$ whenever $1 \leq k \leq m$; this is another consequence of [3, Theorem 1.9.7, or Theorem 1.5.6]. Hence,

$$\beta(m)^2 = \sum_{k=1}^m \frac{r(m)^2}{r(k)^2} \leq \sum_{k=1}^m C \frac{m^{1.02}}{k^{1.02}} \leq Cm^{1.02}, \tag{7.39}$$

which yields (7.35). In the opposite direction, trivially

$$\beta(m)^2 = \sum_{k=1}^m \frac{r(m)^2}{r(k)^2} \geq \sum_{k=1}^m 1 = m. \tag{7.40}$$

(ii): Since $r(m)$ is regularly varying, if $\tilde{m} \sim m$ then $r(\tilde{m}) \sim r(m)$, as a consequence of [3, Theorem 1.9.7 or Theorem 1.5.2].

Furthermore, $r(m)^{-2}$ is regularly varying of index -2κ , and it follows, using the assumption $\sum_m r(m)^{-2} = \infty$ when $\kappa = 1/2$, that $\hat{\sigma}_m^2$ defined by (7.2) is regularly varying of index $1 - 2\kappa$, see [3, Theorem 1.9.5(ii) and Propositions 1.5.8 and 1.5.9b]. Hence $\hat{\sigma}_{\tilde{m}}^2 \sim \hat{\sigma}_m^2$. \square

Proof of Theorem 1.3. Note first that (1.7) implies that $\mu(n) \rightarrow \infty$ as $n \rightarrow \infty$. We may thus assume $\mu(n) \geq 1$. Furthermore, (7.34) yields $\gamma(n) = \beta(\lfloor \mu(n) \rfloor) \rightarrow \infty$ as $n \rightarrow \infty$. We may thus also replace $\mu(n)$ by $\lfloor \mu(n) \rfloor$, and thus in the sequel assume that $\mu(n)$ is an integer.

Fix $x \in \mathbb{R}$ and let $m := \lceil \mu(n) + x\gamma(n) \rceil$. By (1.9) and (7.35), as $n \rightarrow \infty$,

$$\gamma(n) = \beta(\mu(n)) = o(\mu(n)) \tag{7.41}$$

and thus $m \sim \mu(n)$. In particular, $m \rightarrow \infty$ as $n \rightarrow \infty$ and $m \geq 1$ for all large n ; consider only such n . Then, using (3.1),

$$\begin{aligned} \mathbb{P}(M_n \geq \mu(n) + x\gamma(n)) &= \mathbb{P}(M_n \geq m) = \mathbb{P}(N_m \leq n) \\ &= \mathbb{P}\left(\frac{\log N_m - (y_m + \log r(m))}{\hat{\sigma}_m} \leq \frac{\log n - (y_m + \log r(m))}{\hat{\sigma}_m}\right). \end{aligned} \tag{7.42}$$

In the last line, the random variable on the left of ‘ \leq ’ is asymptotically normal by (7.30); we turn to the term on the right. By (1.7) and (7.32), $\hat{y}_{\mu(n)} = \log n + O(1)$, and by (7.33) this yields

$$\log n - (y_m + \log r(m)) = \hat{y}_{\mu(n)} - \hat{y}_m + O(1). \tag{7.43}$$

Write $\tilde{m} := \mu(n)$. Since $\tilde{m} \sim m$, it follows by (7.36) that $r(k) \sim r(m)$ uniformly for all k with $\tilde{m} \leq k \leq m$ or $m \leq k \leq \tilde{m}$. Consequently, by (3.13),

$$\hat{y}_{\tilde{m}} - \hat{y}_m \sim \frac{\tilde{m} - m}{r(m)}. \tag{7.44}$$

Furthermore, by (7.36)–(7.37),

$$\gamma(n) = \beta(\tilde{m}) = r(\tilde{m})\hat{\sigma}_{\tilde{m}} \sim r(m)\hat{\sigma}_m, \tag{7.45}$$

and hence, using (7.36)–(7.37) again,

$$\begin{aligned} \frac{\tilde{m} - m}{r(m)} &= \frac{-x\gamma(n) + O(1)}{r(m)} = \frac{-xr(\tilde{m})\hat{\sigma}_{\tilde{m}}}{r(m)} + O(1) \\ &= -x\hat{\sigma}_m(1 + o(1)) + O(1). \end{aligned} \tag{7.46}$$

Since $\hat{\sigma}_m \rightarrow \infty$ by (7.4), it follows from (7.46) that

$$\frac{\tilde{m} - m}{r(m)\hat{\sigma}_m} \rightarrow -x \tag{7.47}$$

and thus (7.43) and (7.44) yield

$$\frac{\log n - (y_m + \log r(m))}{\hat{\sigma}_m} = \frac{\hat{y}_{\tilde{m}} - \hat{y}_m}{\hat{\sigma}_m} + o(1) \rightarrow -x. \tag{7.48}$$

Consequently (7.42) and Theorem 7.5 imply, with $\zeta \in N(0, 1)$,

$$\mathbb{P}(M_n \geq \mu(n) + x\gamma(n)) \rightarrow \mathbb{P}(\zeta \leq -x) = \mathbb{P}(\zeta \geq x) \tag{7.49}$$

for every $x \in \mathbb{R}$, which proves (1.10).

Finally, the limit (7.28) is evidently mixing, i.e., it holds also conditioned on any fixed set Y_1, \dots, Y_K of the variables. (See [1, Proposition 2].) Using the argument in the proof of Lemma 3.3, (7.28) holds also conditioned on the sequence of indicators $1\{X_k\}$ is accepted, $k \leq N_K$, which is equivalent to conditioning on M_1, \dots, M_{N_K} . It follows that the results above, including (7.49) and (1.10), hold also conditioned on M_1, \dots, M_{N_K} , and thus a fortiori also conditioned on M_1, \dots, M_K . Hence (1.10) is mixing. \square

Remark 7.7. It can be seen from the proof that the assumption that $r(m)$ be regularly varying may be replaced by the weaker

$$\text{if } \tilde{m} = m + O(\beta(m)), \text{ then } r(\tilde{m}) \sim r(m). \tag{7.50}$$

8 Examples

Example 8.1. Consider ‘hiring above the median’. By (1.2), this is the case $r(m) = \lfloor m/2 \rfloor + 1$. Hence, (6.1) holds with $\nu = 1$ and $q = 2$, and thus Theorem 6.1 applies and yields

$$\mathbb{E} W^s = 2^s \Gamma\left(\frac{s}{2} + 1\right), \tag{8.1}$$

which shows that $(W/2)^2 \in \text{Exp}(1)$ and that W thus has a Rayleigh distribution with density function

$$\frac{1}{2} x e^{-x^2/4}, \quad x > 0. \tag{8.2}$$

Consequently, Theorem 1.2 shows that $M_n/n^{1/2}$ converges in distribution to this Rayleigh distribution, as shown by Helmi and Panholzer [11]; moreover, Theorem 1.2 shows a.s. convergence, and convergence of all moments. ([11] treated only the mean.)

Furthermore, the definition (3.4) yields $\delta_k = \mathbf{1}\{m \text{ odd}\}$, and consequently, (4.2) yields

$$Z = \sum_{j=1}^{\infty} \frac{1}{r(2j-1)} (E_{2j-1} - 1) = \sum_{j=1}^{\infty} \frac{1}{j} (E_{2j-1} - 1) \stackrel{d}{=} \sum_{j=1}^{\infty} \frac{1}{j} (E_j - 1). \tag{8.3}$$

It is well-known that this sum yields a centered Gumbel distribution:

$$\mathbb{P}(Z \leq x) = e^{-e^{-(x-\gamma)}}. \tag{8.4}$$

| α | $\mathbb{E} W_\alpha$ | $\mathbb{E} W_\alpha^2$ |
|---------------|---|---|
| 1 | $\frac{1}{2}$ | $\frac{1}{3}$ |
| $\frac{1}{2}$ | $\frac{2\sqrt{\pi}}{3}$ | 2 |
| $\frac{1}{3}$ | $\frac{1}{4}\Gamma(\frac{1}{3})^2$ | $\frac{12}{5}\Gamma(\frac{2}{3})^2$ |
| $\frac{2}{3}$ | $\frac{3}{2^{2/3}\cdot 5}\Gamma(\frac{1}{3})$ | $\frac{9}{2^{1/3}\cdot 7}\Gamma(\frac{2}{3})$ |
| $\frac{1}{4}$ | $\frac{1}{20}\Gamma(\frac{1}{4})^3$ | $\frac{4}{3}\pi^{3/2}$ |
| $\frac{3}{4}$ | $\frac{4}{3^{3/4}\cdot 7}\Gamma(\frac{1}{4})$ | $\frac{16}{3^{3/2}\cdot 5}\pi^{1/2}$ |
| $\frac{1}{5}$ | $\frac{1}{150}\Gamma(\frac{1}{5})^4$ | $\frac{16}{35}\Gamma(\frac{2}{5})^4$ |
| $\frac{2}{5}$ | $\frac{2^{1/5}}{7}\Gamma(\frac{1}{5})\Gamma(\frac{2}{5})$ | $\frac{2^{7/5}\cdot 5}{9}\Gamma(\frac{2}{5})\Gamma(\frac{4}{5})$ |
| $\frac{3}{5}$ | $\frac{5\Gamma(\frac{13}{15})\Gamma(\frac{1}{5})}{3^{3/5}\cdot 8\Gamma(\frac{2}{3})}$ | $\frac{10\Gamma(\frac{1}{15})\Gamma(\frac{2}{3})}{3^{11/5}\cdot 11\Gamma(\frac{2}{3})}$ |
| $\frac{4}{5}$ | $\frac{5}{2^{8/5}\cdot 9}\Gamma(\frac{1}{5})$ | $\frac{25}{2^{11/5}\cdot 13}\Gamma(\frac{2}{5})$ |

Figure 1: Some values of the asymptotic first and second moments for ‘the α -percentile rule’.

(Recall that by definition, $\mathbb{E} Z = 0$.) This can be seen in several ways, for example by computing the moment generating function $\mathbb{E} e^{sZ} = e^{-\gamma s}\Gamma(1 - s)$, $\text{Re } s < 1$, by arguments similar to the proof of Theorem 6.1, or directly from (5.8) and the identification of W as a Rayleigh distribution; we omit the details.

Example 8.2. Let $0 < \alpha \leq 1$ and consider ‘the α -percentile rule’, i.e. $r(m) = \lceil \alpha m \rceil$ by (1.3). Theorem 1.2 applies and shows convergence $M_n/n^\alpha \rightarrow W_\alpha$, a.s., in distribution, and with all moments, to some positive random variable W_α with moments given by (1.6). In particular, when α is rational, the moments can be calculated (in terms of the Gamma function) by Theorem 6.1. We give a few examples in Table 1. (The expectations were given in [6], see below. Note that the results can be written in different forms, using standard Gamma functions identities; cf. the partly different but equivalent formulas for the expectations given here and in [6].) In particular, note that (6.5) yields for the special case $\nu = 1$ (where $r(i) = 1$ for $1 \leq i \leq q$)

$$\mathbb{E} W_{1/q}^s = \frac{q^{s+1}}{s+q}\Gamma\left(\frac{s}{q} + 1\right)^{q-1} = \frac{q^{s+2-q} s^{q-1}}{s+q}\Gamma\left(\frac{s}{q}\right)^{q-1} \tag{8.5}$$

and for $\nu = q - 1$ (where $r(i) = i$ for $1 \leq i \leq q - 1$ and $r(q) = \nu = q - 1$)

$$\mathbb{E} W_{\nu/q}^s = \frac{q^{s+1}}{\nu^{\nu s/q}(\nu s + q)}\Gamma\left(\frac{s}{q} + 1\right) = \frac{s q^s}{\nu^{\nu s/q}(\nu s + q)}\Gamma\left(\frac{s}{q}\right). \tag{8.6}$$

The expectations $c_\alpha := \mathbb{E} W_\alpha = \lim_{n \rightarrow \infty} \mathbb{E} M_n/n^\alpha$ have been considered before. Krieger, Pollak and Samuel-Cahn [14] found that $W_1 \sim U(0, 1)$ and thus $c_1 = 1/2$, but otherwise showed only existence of the limit c_α . Gaither and Ward [6] computed c_α (our $\mathbb{E} W_\alpha$) as

$$c_\alpha = \frac{1 + \sum_{k \geq 1} \frac{[\alpha k] - \alpha k}{[\alpha k]} \prod_{j=1}^k \frac{1}{1 + \alpha/[\alpha j]}}{(\alpha + 1)\Gamma(\alpha + 1)}. \tag{8.7}$$

In the case $\alpha = \nu/q$ rational, they showed further how this can be transformed into a form that they could evaluate symbolically; as examples they gave explicit values for all cases with $q \leq 6$. The formula (8.7) must agree with (1.6) for $s = 1$, i.e.,

$$c_\alpha = \mathbb{E} W_\alpha = \frac{1}{\Gamma(\alpha + 2)} \prod_{k=1}^{\infty} \frac{1 + 1/k}{1 + \alpha/[\alpha k]}, \tag{8.8}$$

but we do not see any direct proof of this. The explicit values for α rational are obtained more easily from (6.2)–(6.5); in particular, for the cases $\nu = 1$ and $\nu = q - 1$, we can take $s = 1$ in (8.5)–(8.6).

Gaither and Ward [6] gave a graph of the function $\alpha \mapsto c_\alpha$, and conjectured that it is continuous at all irrational α but only left-continuous at rational α . This is easily verified from our form (8.8), since the infinite product converges uniformly on each interval $[a, 1]$, and each factor in it is continuous at irrational α and left-continuous everywhere, but for each rational α there are factors that have jumps; furthermore, the jumps in the factors are always positive. Hence, c_α has a positive jump at each rational $\alpha \in (0, 1)$.

Let us consider the case $\alpha = 1/2$ in more detail. For comparisons, let W_{med} denote the limit variable for ‘hiring above the median’ in Example 8.1. Then (6.3) and (8.1) yield

$$\mathbb{E} W_{1/2}^s = \frac{2^s}{1 + s/2} \Gamma\left(\frac{s}{2} + 1\right) = \frac{1}{s/2 + 1} \mathbb{E} W_{\text{med}}^s \tag{8.9}$$

and thus

$$W_{1/2} \stackrel{d}{=} U^{1/2} W_{\text{med}}, \tag{8.10}$$

where $U \sim U(0, 1)$ is independent of W_{med} . In particular, $\mathbb{E} W_{1/2} = 2\sqrt{\pi}/3 = \frac{2}{3} \mathbb{E} W_{\text{med}}$, as shown by Helmi and Panholzer [11], who also gave a formula for the density function of $W_{1/2}$ which is equivalent to (8.10).

The relation between $W_{1/2}$ and W_{med} is studied further in Example 8.3 and Example 8.4.

Example 8.3. One way to regard the difference between ‘hiring above the median’ and ‘the $\frac{1}{2}$ -percentile rule’ is that their sequences $(r(m))_0^\infty$ are $1, 1, 2, 2, \dots$ and $1, 1, 1, 2, 2, \dots$, respectively, and thus differ only by an extra 1 in the latter case.

We can study this in general. Given a sequence $r(m)$ satisfying (1.1), define a new sequence $\tilde{r}(m)$ by inserting an extra 1 first, i.e., let $\tilde{r}(m) := r(m - 1)$, $m \geq 1$. We use $\tilde{\cdot}$ to denote variables for the new sequence. It follows from (4.2) that

$$\tilde{Z} \stackrel{d}{=} Z + E - 1 \tag{8.11}$$

with $E \in \text{Exp}(1)$ independent of Z . Suppose now that $r(m) = \alpha m + O(1)$, or more generally that (5.1) holds. Then the same is true for $\tilde{r}(m)$; furthermore it is easy to see from (5.2) that $\tilde{\rho} = \rho + 1$, and thus (5.8) yields, using (8.11),

$$\tilde{W} \stackrel{d}{=} e^{-\alpha E} W = U^\alpha W \tag{8.12}$$

where $U = e^{-E} \in U(0, 1)$ is independent of W . Equivalently,

$$\mathbb{E} \tilde{W}^s = \frac{1}{1 + \alpha s} \mathbb{E} W^s, \tag{8.13}$$

which also follows from (1.6).

In fact, this has a simple probabilistic explanation. In the modified strategy, the first candidate is, as always, accepted, and because $\tilde{r}(1) = 1$, the threshold for the next candidate is $\tilde{Y}_1 = X_1$. Since the threshold never decreases (see Lemma 3.3), this means that only candidates better than X_1 have a chance of being considered. Moreover, it is easy to see that if we consider only the subsequence of candidates with values $X_n > X_1$, then the ones hired by the modified strategy are precisely those that would have been hired by the original strategy applied to this subsequence of candidates. Conditioning on $X_1 = x_1$, the values in the subsequence will be independent with the conditional distribution $\mathcal{L}(E + x_1)$, and subtracting x_1 from all values, we obtain the original problem

for the original sequence. However, still conditioned on X_1 , if we start with a sequence of n candidates, the subsequence will contain only $\text{Bin}(n - 1, e^{-X_1})$ candidates. Note that $U := e^{-X_1} \in U(0, 1)$. It follows, using the law of large numbers, that if Theorem 1.2 applies to the original strategy, then it holds for the modified one too, with

$$\widetilde{W} \stackrel{d}{=} U^\alpha W, \tag{8.14}$$

where $U \in U(0, 1)$ is independent of W .

Example 8.4. Another way to view the difference between ‘hiring above the median’ and ‘the $\frac{1}{2}$ -percentile rule’ is that $r(m)$ has been decreased by 1 for every even $m \geq 2$. Let us consider, in general, the effect of decreasing a single value $r(m)$ by 1, assuming that this is possible (i.e., that $r(m - 1) < r(m) = r(m + 1)$). Assume also for simplicity that Theorem 1.2 applies. Then (1.6) shows that W is modified such that $\mathbb{E} W^s$ is multiplied by

$$\frac{1 + s\alpha/r(m)}{1 + s\alpha/(r(m) - 1)} = \frac{r(m) - 1}{r(m)} + \frac{1/r(m)}{1 + s\alpha/(r(m) - 1)} = \mathbb{E} V^{s\alpha/(r(m)-1)}, \tag{8.15}$$

where V has density $1/r(m)$ on $(0, 1)$ and a point mass $\mathbb{P}(V = 1) = 1 - 1/r(m)$. Hence, the modified limit $\widetilde{W} \stackrel{d}{=} V^{\alpha/(r(m)-1)} W$. This can be repeated for several changes.

In particular, looking just at the expectation, decreasing $r(2)$ from 2 to 1 in ‘hiring above the median’ multiplies $\mathbb{E} W$ by $(1 + \frac{1}{4})/(1 + \frac{1}{2}) = 5/6$. As seen above, $\mathbb{E} W$ decreases by a factor $2/3$ if we change ‘hiring above the median’ to ‘the $1/2$ -percentile rule’, and we now see that half of the decrease is due to the decrease of $r(2)$. This illustrates that, as said in Section 1, in the case of large $r(m)$, the asymptotic behaviour is heavily influenced by the effects of the first candidates.

Example 8.5. Another variation of ‘hiring above the median’ is to take the sequence $1, 2, 2, 3, 3, \dots$, i.e.,

$$r(m) = \lceil m/2 \rceil + 1, \quad m \geq 0. \tag{8.16}$$

Theorem 6.1 applies and yields, e.g. by (6.3),

$$\mathbb{E} W^s = 2^s \Gamma\left(\frac{s}{2} + 2\right). \tag{8.17}$$

This follows also from Example 8.3, since if we insert an extra 1 first in this sequence $r(m)$, we obtain ‘hiring above the median’ as in Example 8.1, and thus by (8.13) and (8.1),

$$2^s \Gamma\left(\frac{s}{2} + 1\right) = \frac{1}{s/2 + 1} \mathbb{E} W^s. \tag{8.18}$$

It follows from (8.17) that $W^2/4 \in \Gamma(2)$, and thus W has the density

$$f_W(x) = \frac{x^3}{8} e^{-x^2/4}, \quad x > 0. \tag{8.19}$$

Equivalently, $W/\sqrt{2} \sim \chi(4)$, a chi distribution.

We return to the significance of this example in Section 9.

Example 8.6. The extreme case of small $r(m)$ is $r(m) = 1$, $m \geq 0$. This means the we only accept candidates that are better than all previous candidates, i.e., the record values in the sequence (X_n) .

Theorem 1.3 applies with $\mu(n) = \log n$, $\beta(m)^2 = m$ and $\gamma(n)^2 = \lfloor \log n \rfloor \sim \log n$, which yields

$$\frac{M_n - \log n}{\log^{1/2} n} \xrightarrow{d} N(0, 1). \tag{8.20}$$

This is a well-known result for the number of records, see e.g. [7, Theorem 7.4.2], and is easily proved directly by the central limit theorem, observing that the indicators $I_k := \mathbf{1}\{X_k \text{ is a record}\}$ are independent with $I_k \sim \text{Be}(1/k)$. See further the next example. (This connection between records and the hiring problem was noted by [4].)

Example 8.7. More generally, consider ‘hiring above the r -th best’ for a fixed $r \geq 1$, with $r(m)$ given by (1.4). Thus Example 8.6 is the case $r = 1$. This strategy was studied by Archibald and Martínez [2] and, in great detail, by Helmi, Martínez and Panholzer [9]. A value X_k is accepted if it is an r -record, in the sense that it is one of the r best values seen so far. (In particular, the first r values X_k are always accepted.)

Theorem 1.3 applies with $\mu(n) = r \log n$, $\beta(m)^2 \sim m$ and $\gamma(n)^2 \sim r \log n$, which yields

$$\frac{M_n - r \log n}{(r \log n)^{1/2}} \xrightarrow{d} N(0, 1), \tag{8.21}$$

as shown by [9] (who also gave many other results, including for fixed n , and for the case when both $n, r \rightarrow \infty$). Again, this is easily shown directly by the central limit theorem, using the fact that the indicators $I_k := \mathbf{1}\{X_k \text{ is an } r\text{-record}\}$ are independent with $I_k \sim \text{Be}(r/k)$ for $k \geq r$, which is noted in [9], see also the further references given there.

Example 8.8. Let $r(m) := \lfloor \sqrt{m} \rfloor$, $m \geq 1$. This is an example of small $r(m)$. (But rather large among the small ones.) We have

$$\sigma_m^2 = \hat{\sigma}_m^2 + O(1) = \sum_{k=1}^m \frac{1}{\lfloor \sqrt{k} \rfloor^2} + O(1) = \sum_{k=1}^m \frac{1}{k} + O(1) = \log m + O(1)$$

and

$$\hat{y}_m = \sum_{k=1}^m \frac{1}{\lfloor \sqrt{k} \rfloor} = \sum_{j=1}^{\lfloor \sqrt{m} \rfloor} \frac{2j+1}{j} + O(1) = 2\sqrt{m} + \frac{1}{2} \log m + O(1). \tag{8.22}$$

It follows that Theorem 1.3 applies with

$$\mu(n) = \frac{1}{4} \log^2 n - \frac{1}{2} \log n \log \log n, \tag{8.23}$$

$\beta(m)^2 = r(m)^2 \hat{\sigma}_m^2 \sim m \log m$ and

$$\gamma(n)^2 \sim \mu(n) \log \mu(n) \sim \frac{1}{2} \log^2 n \log \log n. \tag{8.24}$$

Hence, Theorem 1.3 yields

$$\frac{M_n - (\frac{1}{4} \log^2 n - \frac{1}{2} \log n \log \log n)}{\log n (\log \log n)^{1/2}} \xrightarrow{d} N\left(0, \frac{1}{2}\right). \tag{8.25}$$

Example 8.9. We give an example showing how Theorem 1.3 can fail when the sequence $r(m)$ is too irregular.

Let $r(m)$ be such that $r(8^i) = 2^i$, $r(m)$ increases linearly on each interval $[8^i, 8^i + 2^i]$, and is constant between these intervals; i.e.,

$$r(m) := \begin{cases} 2^i + m - 8^i, & 8^i \leq m \leq 8^i + 2^i, \quad i \geq 0 \\ 2^{i+1}, & 8^i + 2^i \leq m \leq 8^{i+1}, \quad i \geq 0. \end{cases} \tag{8.26}$$

Thus $r(m) \asymp m^{1/3}$ (meaning that $cm^{1/3} \leq r(m) \leq Cm^{1/3}$ for some constants c and C), and thus (7.2)–(7.3) yield $\sigma_m^2 \sim \hat{\sigma}_m^2 \asymp m^{1/3}$. Hence $\beta(m)^2 = r(m)^2 \hat{\sigma}_m^2 \asymp m$ so $\beta(m) \asymp m^{1/2}$. We see also that $\tilde{m} \sim m$ implies $\hat{\sigma}_{\tilde{m}} \sim \hat{\sigma}_m$.

Define $m_i = 8^i$ and $n_i := \lfloor \exp(\hat{y}_{8^i}) \rfloor$, $i \geq 1$. Consider only the subsequence $(n_i)_i$. We have chosen n_i such that (1.7) holds with $n = n_i$ and $\mu(n) = 8^i$, so we may choose $\mu(n)$ such that $\mu(n_i) = 8^i = m_i$ for $i \geq 1$. Then $\gamma(n_i) = \beta(8^i) \asymp 2^{3i/2}$.

Now argue as in the proof of Theorem 1.3 for $n = n_i$. Thus $\tilde{m} = \mu(n_i) = m_i = 8^i$. Suppose first that $x < 0$. Then $m \leq \tilde{m}$, and $r(k) = 2^i = r(m_i)$ for every $k \in [m, \tilde{m}]$ (at least if i is large); hence, as $i \rightarrow \infty$,

$$\hat{y}_{\tilde{m}} - \hat{y}_m = \frac{\tilde{m} - m}{r(m_i)} \sim \frac{-x\gamma(n_i)}{r(m_i)} = \frac{-x\beta(m_i)}{r(m_i)} = -x\hat{\sigma}_{m_i}, \tag{8.27}$$

and since $\hat{\sigma}_m \sim \hat{\sigma}_{m_i}$,

$$\frac{\hat{y}_{\tilde{m}} - \hat{y}_m}{\hat{\sigma}_m} \rightarrow -x. \tag{8.28}$$

On the other hand, if $x > 0$, then $m \geq \tilde{m}$, and for most $k \in [\tilde{m}, m]$, $r(k) = 2^{i+1} = 2r(m_i)$. Hence, as $i \rightarrow \infty$, similarly,

$$\hat{y}_{\tilde{m}} - \hat{y}_m \sim -\frac{m - \tilde{m}}{2r(m_i)} \sim -\frac{x\gamma(n_i)}{2r(m_i)} = -\frac{x}{2}\hat{\sigma}_{m_i}, \tag{8.29}$$

and

$$\frac{\hat{y}_{\tilde{m}} - \hat{y}_m}{\hat{\sigma}_m} \rightarrow -\frac{x}{2}. \tag{8.30}$$

Instead of (7.49), we thus obtain

$$\mathbb{P}(M_{n_i} \geq \mu(n_i) + x\gamma(n_i)) \rightarrow \begin{cases} \mathbb{P}(\zeta \leq -x) = \mathbb{P}(\zeta \geq x), & x < 0, \\ \mathbb{P}(\zeta \leq -x/2) = \mathbb{P}(2\zeta \geq x), & x > 0. \end{cases} \tag{8.31}$$

Consequently, along the sequence (n_i) , M_n after normalization as in (1.10) converges to the non-normal random variable

$$\eta := \begin{cases} \zeta, & \zeta \leq 0, \\ 2\zeta, & \zeta > 0, \end{cases} \tag{8.32}$$

where $\zeta \in N(0, 1)$.

On the other hand, for many other subsequences there is asymptotic normality (by the same proof), for example for $\lfloor \exp(\hat{y}_{2 \cdot 8^i}) \rfloor$.

9 Conditioning on the first value

We have seen above that in the case of large $r(m)$, the asymptotics depend heavily on the first values X_k , and thus in particular on the first value X_1 . Furthermore, as have been remarked by [4], assuming $r(1) = 1$, so that the second accepted candidate is the first one with $X_n > X_1$, the waiting time $N_2 - N_1$ until the second candidate is accepted has, conditioned on X_1 , the distribution $\text{Ge}(e^{-X_1})$ with expectation

$$\mathbb{E}(N_2 - N_1 \mid X_1) = e^{X_1} = 1/U, \tag{9.1}$$

where $U := e^{-X_1} \in U(0, 1)$. Consequently, $\mathbb{E}(N_2 - N_1) = \mathbb{E}U^{-1} = \infty$, and thus $\mathbb{E}N_m = \infty$ for every $m \geq 2$.

These effects led [4] to consider ‘hiring above the median’ conditioned on X_1 . We can do this in general. We assume that $r(m)$ is large, since for small $r(m)$, conditioning on X_1 has no effect on the asymptotics, see e.g. the mixing property in Theorem 1.3.

Theorem 9.1. Suppose that $r(m)$ is large and that $r(1) = 1$, and define $\check{r}(m) := r(m + 1)$, $m \geq 0$. Conditioned on $X_i = x_1$, the results in Theorem 1.2 and Sections 4–6 hold, mutatis mutandis, with $r(m)$ replaced by $\check{r}(m)$ and n replaced by pn , where $p := \mathbb{P}(X_2 > x_1) = e^{-x_1}$.

In particular, when Theorem 1.2 applies, (1.5) extends to

$$M_n/n^\alpha \xrightarrow{\text{a.s.}} W = p^\alpha \check{W} = e^{-\alpha X_1} \check{W}, \tag{9.2}$$

where \check{W} is independent of X_1 and has moments as in (1.6) for the sequence $\check{r}(m)$.

For a different distribution of the values X_n , e.g. uniform, p is of course given by the corresponding tail probability.

Proof. This was explained already in Example 8.3, although we here modify in the opposite direction, so the original sequence here is the modified one there. As explained in Example 8.3, of the first n candidates, the ones accepted are the first one and then the candidates accepted using the strategy given by $\check{r}(m)$ on the candidates that pass the test $X_k > x_1$. For asymptotics, we can ignore the first accepted candidate, and thus the results are the same as for $\check{r}(m)$ with n replaced by \check{N}_n , the (random) number of values X_k , $2 \leq k \leq n$, such that $X_k > x_1$. By the law of large numbers, a.s. $\check{N}_n \sim pn$, and the result follows. We omit the details. \square

Example 9.2. Consider again ‘hiring above the median’ as in Example 8.1, but condition on X_1 . The sequence $\check{r}(m) := r(m + 1)$ then is the one studied in Example 8.5; thus we find, for example, see (9.2), that conditioned on $X_1 = x_1$,

$$M_n/n^{1/2} \xrightarrow{\text{a.s.}} p^{1/2} \check{W} \tag{9.3}$$

where $p = e^{-x_1}$ and W has the distribution with density (8.19). This (and more) has been shown by Helmi and Panholzer [11, Theorem 5]. Moment convergence holds too, and thus, using (8.17), for $-2 < s < \infty$,

$$\mathbb{E}(M_n^s \mid X_1 = x_1) \sim e^{-sx_1/2} \mathbb{E} \check{W}^s n^{s/2} = e^{-sx_1/2} 2^s \Gamma\left(\frac{s}{2} + 2\right) n^{s/2}. \tag{9.4}$$

Problem 9.3. What happens if we condition on X_1 in a case with $r(1) = 2$?

10 Probability of accepting and length of gaps

Let $I_n := 1\{X_n \text{ is accepted}\}$ and $p_n := \mathbb{E} I_n$ be the indicator and the probability that candidate n is accepted; thus $M_n = \sum_{k=1}^n I_k$ and $\mathbb{E} M_n = \sum_{k=1}^n p_k$. Let Y_n^* be the current threshold when candidate n is examined. We then have accepted M_{n-1} candidates, and thus

$$Y_n^* = Y_{M_{n-1}}. \tag{10.1}$$

Furthermore, if P_n is the conditional probability that X_n is accepted given the past,

$$P_n := \mathbb{E}(I_n \mid X_1, \dots, X_{n-1}) = e^{-Y_n^*} = e^{-Y_{M_{n-1}}} \tag{10.2}$$

and thus

$$p_n = \mathbb{E} P_n = \mathbb{E} e^{-Y_{M_{n-1}}}. \tag{10.3}$$

We return to a more explicit asymptotic result in the case (5.1) in Theorem 10.5 below.

Conditioned on Y_n^* , or equivalently on P_n , the waiting time until the next candidate is accepted is $\text{Ge}(P_n)$. We will see that asymptotically, the same holds if we go back in time from n to the last acceptance. The next lemma excludes some extreme cases.

Lemma 10.1. *If*

$$\sum_{k=1}^{\infty} \frac{\delta_k}{r(k)} = \infty, \tag{10.4}$$

then $Y_m \xrightarrow{\text{a.s.}} \infty$ as $m \rightarrow \infty$, and thus $P_n \xrightarrow{\text{a.s.}} 0$ and $p_n \rightarrow 0$ as $n \rightarrow \infty$.

Conversely, if the sum in (10.4) converges, then $Y_m \xrightarrow{\text{a.s.}} Y_\infty < \infty$.

Proof. It follows from (3.5) that

$$Y_m \rightarrow Y_\infty := \sum_{m=1}^{\infty} \frac{\delta_k}{r(k)} E_k \leq \infty. \tag{10.5}$$

Note that the sum in (10.4) is $\mathbb{E}Y_\infty$. Hence, if the sum is finite, then $Y_\infty < \infty$ a.s.

Conversely, assume that (10.4) holds. Then the a.s. divergence of the sum in (10.5) follows by the Kolmogorov three series theorem [7, Theorem 6.5.5], or by Lemma 4.1 in the case of large $r(m)$ and otherwise by Lemma 7.4, which implies first $y_m/\sigma_m \rightarrow \infty$ and then $Y_m/y_m \xrightarrow{P} 1$. Furthermore, $M_{n-1} \xrightarrow{\text{a.s.}} \infty$, and thus $Y_{M_{n-1}} \xrightarrow{\text{a.s.}} \infty$ as $n \rightarrow \infty$; hence $P_n \xrightarrow{\text{a.s.}} 0$ by (10.2) and $p_n \rightarrow 0$ by (10.3) and dominated convergence. \square

Theorem 10.2. *Suppose that (10.4) holds and that $r(m) \rightarrow \infty$ as $m \rightarrow \infty$. Then there exists a sequence $a_n \rightarrow \infty$, such that on the interval $\mathcal{J}_n := [n - a_n P_n^{-1}, n]$, the stochastic process $(I_k)_{k \in \mathcal{J}_n}$ w.h.p. agrees with a sequence $(I'_k)_{k \in \mathcal{J}_n}$ of indicator variables that conditioned on P_n are i.i.d. with $I'_k \in \text{Be}(P_n)$.*

Proof. Fix an integer $K > 0$, and define the stopping time $\nu_n := \min\{k : M_{k-1} \geq M_{n-1} - K\}$. Thus $\nu_n \leq n$ and, assuming n is so large that $M_{n-1} \geq K$,

$$Y_{\nu_n}^* = Y_{M_{n-1}-K} \geq Y_{M_{n-1}} - \frac{K}{r(M_{n-1}-K)} = Y_{M_{n-1}} + o(1) = Y_n^* + o(1) \tag{10.6}$$

a.s. as $n \rightarrow \infty$, since $M_n \rightarrow \infty$ and thus $r(M_{n-1} - K) \rightarrow \infty$ a.s. Consequently, a.s. as $n \rightarrow \infty$,

$$P_{\nu_n} \sim P_n. \tag{10.7}$$

By definition, K candidates are accepted in the interval $\mathcal{J}^* := [\nu_n, n]$ (provided $\nu_n > 1$), and, conditioned on P_{ν_n} , each candidate in \mathcal{J}^* is accepted with probability at most P_{ν_n} .

Let a be a fixed large number and define $n_1 := \lceil n - aP_{\nu_n}^{-1} \rceil$. If $\nu_n \geq n_1$, then $|\mathcal{J}^*| \leq n - n_1 \leq aP_{\nu_n}^{-1}$, and thus at least K candidates are accepted in the interval $[\nu_n, \nu_n + \lceil aP_{\nu_n}^{-1} \rceil]$. Hence, using Markov's inequality,

$$\mathbb{P}(\nu_n \geq n_1 \mid P_{\nu_n} = p) \leq \mathbb{P}(\text{Bin}(\lceil ap^{-1} \rceil, p) \geq K) \leq a/K. \tag{10.8}$$

Consequently, given any $\varepsilon > 0$, we may by choosing $K > a/\varepsilon$ make this probability $< \varepsilon$, uniformly in $p > 0$. Hence, we may in the rest of the proof assume that $\nu_n < n_1$. This means that for every k in the interval $[n_1, n]$, $\nu_n < k \leq n$, and thus $P_{\nu_n} \geq P_k \geq P_n$. It follows that, conditioned on P_{ν_n} , we may couple the Markov process $(I_k)_{k \in [n_1, n]}$ with a sequence of (conditionally) i.i.d. variables $(I''_k)_{k \in [n_1, n]}$ with $\mathbb{P}(I''_k) = P_{\nu_n}$, with an error probability at most, using (10.7),

$$(n - n_1 + 1)(P_{\nu_n} - P_n) \sim aP_{\nu_n}^{-1}(P_{\nu_n} - P_n) \xrightarrow{\text{a.s.}} 0. \tag{10.9}$$

We now uncondition, and see (using (10.9) and dominated convergence) that we may couple $(I_k)_{k \in [n_1, n]}$ and $(I''_k)_{k \in [n_1, n]}$ with error probability $o(1)$. We may then instead

couple with $(I'_k)_{k \in [n_1, n]}$ where I'_k are conditionally i.i.d. with $\mathbb{P}(I'_k) = P_n$, introducing an additional error $o(1)$ by an estimate similar to (10.9).

Furthermore, a.s. $P_{\nu_n} \sim P_n$ by (10.7), and thus $aP_{\nu_n}^{-1} > (a-1)P_n^{-1}$ for large n . Thus, we may as well couple (I_k) and (I'_k) on $[n - (a-1)P_n^{-1}, n]$, for any fixed a .

We may here replace a by $a+1$. Moreover, by a simple general argument, since this coupling with error probability $o(1)$ is possible for every fixed $a > 0$, it is also possible for some sequence $a_n \rightarrow \infty$; this follows by the following elementary lemma, taking $x(a, n)$ to be the total variation distance between the two sequences, which completes the proof. \square

Lemma 10.3. *Suppose that $x(a, n)$, $a, n \in \mathbb{N}$, are real numbers such that for every fixed a , $x(a, n) \rightarrow 0$ as $n \rightarrow \infty$. Then there exists a sequence $a_n \rightarrow \infty$ such that $x(a_n, n) \rightarrow 0$.*

Proof. Let $n_0 = 1$. For every $k \geq 1$, choose $n_k > n_{k-1}$ such that $|x(k, n)| < 1/k$ when $n \geq n_k$. Define $a_n = k$ when $n_k \leq n < n_{k+1}$. \square

Let $L_n := n - N_{M_n}$ be the number of candidates examined after the last accepted one. Let $d_{TV}(X, Y)$ denote the total variation distance between two distributions or random variables.

Corollary 10.4. *Suppose that (10.4) holds and that $r(m) \rightarrow \infty$ as $m \rightarrow \infty$. Then, conditioned on P_n , $d_{TV}(L_n, \text{Ge}(P_n)) \rightarrow 0$ as $n \rightarrow \infty$. Consequently, still conditioned on P_n ,*

$$P_n L_n \xrightarrow{d} \text{Exp}(1). \tag{10.10}$$

Proof. Immediate from Theorem 10.2 and the fact that $P_n \xrightarrow{\text{a.s.}} 0$ by Lemma 10.1. \square

We can also find the unconditional distribution of L_n . For convenience, and in order to obtain more explicit results, we consider only the case in Section 5, and we assume $\alpha < 1$, which implies (10.4). We first study P_n .

Theorem 10.5. *Suppose that (5.1) holds for some $\alpha \in (0, 1)$. Then, a.s.,*

$$P_n \sim \alpha M_n / n \sim \alpha W n^{\alpha-1}, \tag{10.11}$$

where W is as in Theorem 1.2.

Proof. Let, for convenience $\xi := (\alpha^{-1} - 1)\gamma + \rho + 1$, so $W = e^{-\alpha\xi - \alpha Z}$ by (5.8). Then, by Lemma 4.1 and (5.4), a.s.,

$$Y_m = y_m + Z + o(1) = (\alpha^{-1} - 1) \log m - \log \alpha + \xi + Z + o(1) \tag{10.12}$$

and thus, by (5.8),

$$\begin{aligned} Y_n^* &= (\alpha^{-1} - 1) \log M_n - \log \alpha + \xi + Z + o(1) \\ &= (1 - \alpha) \log n + (\alpha^{-1} - 1) \log W - \log \alpha + \xi + Z + o(1) \\ &= (1 - \alpha) \log n - \log \alpha + \alpha\xi + \alpha Z + o(1). \end{aligned} \tag{10.13}$$

Hence, by (10.2),

$$P_n \sim \alpha e^{-\alpha\xi - \alpha Z} n^{\alpha-1} = \alpha W n^{\alpha-1}. \tag{10.14}$$

\square

Theorem 10.6. *Suppose that (5.1) holds for some $\alpha \in (0, 1)$. Then*

$$L_n / n^{1-\alpha} \xrightarrow{d} \widehat{L} := \alpha^{-1} W^{-1} E, \tag{10.15}$$

where W is as in Theorem 1.2 and $E \in \text{Exp}(1)$ is independent of W .

Proof. A consequence of Corollary 10.4 and Theorem 10.5. □

Remark 10.7. The moments

$$\mathbb{E} \widehat{L}^s = \alpha^{-s} \mathbb{E} [W^{-s}] \mathbb{E} [E^s] = \alpha^{-s} \Gamma(s + 1) \mathbb{E} W^{-s}, \quad s > -1, \quad (10.16)$$

follow from (1.6). Note that for real s , $\mathbb{E} \widehat{L}^s < \infty$ if $-1 < s < r_*/\alpha$, but not outside this interval, see Remark 5.5.

Example 10.8. For ‘hiring above the median’, W has the density function (8.2), and thus $\widehat{L} = 2W^{-1}E$ has the density function

$$\begin{aligned} f_{\widehat{L}}(x) &= \int_0^\infty f_W(y) f_{2y^{-1}E}(x) dy = \int_0^\infty \frac{y}{2} e^{-y^2/4} \frac{y}{2} e^{-yx/2} dy \\ &= 2 \int_0^\infty t^2 e^{-t^2-tx} dt, \end{aligned} \quad (10.17)$$

as found by [11, Theorem 3]. Furthermore, (10.16) and (8.1) yield

$$\mathbb{E} \widehat{L}^s = \Gamma(s + 1) \Gamma\left(1 - \frac{s}{2}\right), \quad -1 < s \leq 2. \quad (10.18)$$

In particular, $\mathbb{E} \widehat{L} = \sqrt{\pi}$. [11] proved convergence of $\mathbb{E} L_n/n^{1/2}$ to this limit.

Remark 10.9. Theorem 10.2 implies also the same limit results for, e.g., the distance between the last two accepted. See [11, Theorem 4] for ‘hiring above the median’.

11 The distribution of accepted values

Finally, we study the distribution of the accepted values. For simplicity we consider again only the situation in Section 5. We also assume for simplicity that $\alpha < 1$, leaving the case $\alpha = 1$ to the reader.

Let, for a real number x , $M_n^{\leq x}$ be the number of values X_k with $k \leq n$ that are accepted and furthermore satisfy $X_k \leq x$. Define $M_n^{> x} = M_n - M_n^{\leq x}$ similarly.

Theorem 11.1. *Suppose that (5.1) holds for some $\alpha \in (0, 1)$. Then, a.s., for every $u \in \mathbb{R}$,*

$$\frac{M_n^{\leq Y_n^* + u}}{M_n} \rightarrow F(u) := \begin{cases} (1 - \alpha)e^{\alpha u/(1-\alpha)}, & u \leq 0, \\ 1 - \alpha e^{-u}, & u \geq 0. \end{cases} \quad (11.1)$$

In other words, the empirical distribution function of the differences $X_k - Y_n^$ for the M_n accepted candidates converges a.s. to the distribution with distribution function $F(u)$. Hence, if \bar{X}_n is the value of one of the M_n accepted candidates, chosen uniformly at random, then,*

$$\bar{X}_n - Y_n^* \xrightarrow{d} V, \quad (11.2)$$

where V has the distribution $F(u)$.

The proof is given later. Note that V has density

$$f(u) := F'(u) = \begin{cases} \alpha e^{\alpha u/(1-\alpha)}, & u < 0, \\ \alpha e^{-u}, & u > 0. \end{cases} \quad (11.3)$$

Thus, V has an asymmetric double exponential distribution (Laplace distribution); if $\alpha = 1/2$, V has the usual Laplace distribution.

In order to prove Theorem 11.1, we introduce a simpler strategy. Fix a real number z and define

$$x_n = x_n(z) := (1 - \alpha) \log n + \alpha z - \log \alpha \quad (11.4)$$

and, for later convenience,

$$w = w(z) := e^{-\alpha z} \in [0, \infty). \tag{11.5}$$

Define the strategy $\mathcal{H}(z)$ as ‘accept if $X_n > x_n$ ’. (Thus $\mathcal{H}(z)$ is not a rank-based strategy.)

Lemma 11.2. *For the strategy $\mathcal{H}(z)$, a.s.,*

$$M_n \sim wn^\alpha \tag{11.6}$$

and, for every real u ,

$$\frac{M_n^{\leq x_n+u}}{M_n} \rightarrow F(u) := \begin{cases} (1-\alpha)e^{\alpha u/(1-\alpha)}, & u \leq 0, \\ 1 - \alpha e^{-u}, & u \geq 0. \end{cases} \tag{11.7}$$

Proof. Consider first a fixed $u \geq 0$. Then every value $X_k > x_n + u$ with $k \leq n$ will have $X_k > x_n \geq x_k$, and thus be accepted. Hence, $M_n^{>x_n+u}$ is the number of all such values, and since the indicators $\mathbf{1}\{X_k > x_n + u\}$ are i.i.d. with $\mathbb{P}(X_k > x_n + u) = e^{-x_n-u}$,

$$M_n^{>x_n+u} \in \text{Bi}(n, e^{-x_n-u}). \tag{11.8}$$

This binomial random variable has mean $ne^{-x_n-u} = \alpha w e^{-u} n^\alpha$, recalling (11.4) and (11.5). A standard Chernoff bound, see e.g. [13, Corollary 2.3], shows that for every $\varepsilon > 0$,

$$\mathbb{P}(|M_n^{>x_n+u} / \mathbb{E} M_n^{>x_n+u} - 1| > \varepsilon) \leq 2 \exp(-c(\varepsilon, u)n^\alpha). \tag{11.9}$$

It follows, by the Borel–Cantelli lemma, that a.s. $M_n^{>x_n+u} / \mathbb{E} M_n^{>x_n+u} \rightarrow 1$, i.e.,

$$M_n^{>x_n+u} \sim \mathbb{E} M_n^{>x_n+u} = \alpha w n^\alpha e^{-u}. \tag{11.10}$$

Consider next M_n . This too is a sum of independent indicators $I_k := \mathbf{1}\{X_k > x_k\}$. Furthermore, for k large enough so that $x_k \geq 0$,

$$p_k := \mathbb{E} I_k = e^{-x_k} = \alpha e^{-\alpha z} k^{\alpha-1} = \alpha w k^{\alpha-1}. \tag{11.11}$$

Since $M_n = \sum_{k=1}^n I_k$, we have

$$\mathbb{E} M_n = \sum_{k=1}^n p_k = \sum_{k=1}^n \alpha w k^{\alpha-1} + O(1) = wn^\alpha + O(1). \tag{11.12}$$

The random variables I_k are not identically distributed, but the Chernoff bound holds for sums of arbitrary independent indicator variables [13, Theorem 2.8], and thus (11.9) holds for M_n too, and we obtain as above

$$M_n / \mathbb{E} M_n \xrightarrow{\text{a.s.}} 1, \tag{11.13}$$

which together with (11.12) yields (11.6).

Furthermore, (11.6) and (11.10) show that for every fixed $u \geq 0$,

$$M_n^{>x_n+u} / M_n \xrightarrow{\text{a.s.}} \alpha e^{-u}, \tag{11.14}$$

and thus (11.7) holds for $u \geq 0$.

Finally, consider a fixed $u \leq 0$. Similarly as above, we write $M_n^{\leq x_n+u}$ as a sum of independent indicators $I'_k := \mathbf{1}\{x_k < X_k \leq x_n + u\}$. Note that $I'_k = 0$ unless $x_k < x_n + u$, which by (11.4) is equivalent to

$$(1-\alpha) \log k < (1-\alpha) \log n + u \tag{11.15}$$

or

$$k < e^{u/(1-\alpha)}n. \tag{11.16}$$

For such k , except possibly for some small k with $x_k < 0$,

$$\mathbb{E} I'_k := e^{-x_k} - e^{-(x_n+u)} = \alpha w k^{\alpha-1} - \alpha w e^{-u} n^{\alpha-1}. \tag{11.17}$$

Hence,

$$\begin{aligned} \mathbb{E} M_n^{\leq x_n+u} &= \sum_{k=1}^{e^{u/(1-\alpha)}n} (\alpha w k^{\alpha-1} - \alpha w e^{-u} n^{\alpha-1}) + O(1) \\ &= w (e^{u/(1-\alpha)}n)^\alpha - \alpha w e^{-u} e^{u/(1-\alpha)}n^\alpha + O(1) \\ &= w n^\alpha e^{\alpha u/(1-\alpha)}(1 - \alpha) + O(1). \end{aligned} \tag{11.18}$$

The Chernoff argument applies again, and yields

$$M_n^{\leq x_n+u} / \mathbb{E} M_n^{\leq x_n+u} \xrightarrow{\text{a.s.}} 1 \tag{11.19}$$

and thus, using (11.18) and (11.6), a.s.,

$$M_n^{\leq x_n+u} \sim (1 - \alpha) w n^\alpha e^{\alpha u/(1-\alpha)} \tag{11.20}$$

and

$$M_n^{\leq x_n+u} / M_n \rightarrow (1 - \alpha) e^{\alpha u/(1-\alpha)} = F(u). \tag{11.21}$$

We have proved that (11.7) holds a.s. for every fixed $u \in \mathbb{R}$. Hence, it holds a.s. for every rational u , but this implies that it holds for all u simultaneously, since the left-hand side is monotone in u and the right-hand side is continuous; we omit the details. \square

Proof of Theorem 11.1. Consider the rank-based strategy (as in the rest of the paper), which we denote by \mathcal{R} , together with the strategies $\mathcal{H}(z)$ for all rational $z \in \mathbb{R}$, acting on the same sequence X_n . We indicate quantities for the strategy $\mathcal{H}(z)$ with an extra argument z ; for example, $M_n(z)$ for the number of accepted values of the n first ones.

Recall that the strategy \mathcal{R} is to accept X_n if $X_n > Y_n^*$. By (10.13) and (11.4), if $z < Z + \xi$, then a.s. $Y_n^* > x_n(z)$ for all large n , and thus (for large n) every value accepted by \mathcal{R} is also accepted by $\mathcal{H}(z)$. Consequently, a.s.

$$M_n \leq M_n(z) + O(1). \tag{11.22}$$

Similarly, for every fixed real u , a.s.,

$$M_n^{> Y_n^*+u} \leq M_n^{> x_n(z)+u}(z) + O(1). \tag{11.23}$$

For every rational $z > Z + \xi$, (11.22) and (11.23) hold a.s. with the inequalities in the opposite direction.

Consequently, with $G(u) := 1 - F(u)$, a.s., for every rational $z < Z + \xi$ and $z' > Z + \xi$, using (11.6), (11.5), and (11.7),

$$\begin{aligned} \frac{M_n^{> Y_n^*+u}}{M_n} &\leq \frac{M_n^{> x_n(z)+u}(z) + O(1)}{M_n(z') + O(1)} = \frac{M_n^{> x_n(z)+u}(z) + O(1)}{M_n(z)} \frac{M_n(z)}{M_n(z') + O(1)} \\ &\rightarrow G(u) \frac{w(z)}{w(z')} = G(u) e^{\alpha(z'-z)}. \end{aligned} \tag{11.24}$$

Hence, a.s.,

$$\limsup_{n \rightarrow \infty} \frac{M_n^{>Y_n^*+u}}{M_n} \leq G(u)e^{\alpha(z'-z)} \tag{11.25}$$

for every rational z and z' with $z < Z + \xi < z'$. Consequently, a.s.

$$\limsup_{n \rightarrow \infty} \frac{M_n^{>Y_n^*+u}}{M_n} \leq G(u) = 1 - F(u). \tag{11.26}$$

A lower bound follows in the same way, now comparing in the opposite directions with $z > Z + \xi > z'$. This proves (11.1) a.s. for a fixed u , and thus for all rational u simultaneously, which again implies the result for all real u simultaneously by monotonicity and continuity. \square

Corollary 11.3. *Suppose that (5.1) holds for some $\alpha \in (0, 1)$. Then, the fraction of the accepted values that are larger than the current threshold, and thus would have been accepted now, converges a.s. to α .*

Proof. This fraction is $M_n^{>Y_n^*}/M_n$, so the result is the case $u = 0$ of Theorem 11.1. \square

A Proof of (5.3)

Lemma A.1. *Suppose that (5.1) holds for some $\alpha \in (0, 1]$. Then (5.3) holds.*

Proof. Fix $\delta > 0$ and suppose that m is such that $r(m) \geq (1 + \delta)^2 \alpha m$. Then, for every k with $m \leq k \leq (1 + \delta)m$, we have $r(k) \geq r(m)$ and thus

$$(\alpha k)^{-1} - r(k)^{-1} \geq (1 + \delta)^{-1} (\alpha m)^{-1} - (1 + \delta)^{-2} (\alpha m)^{-1} = c(\delta) m^{-1}. \tag{A.1}$$

Hence,

$$\sum_{k=m}^{\infty} |r(k)^{-1} - (\alpha k)^{-1}| \geq \sum_{k=m}^{\lfloor (1+\delta)m \rfloor} c(\delta) m^{-1} \geq c(\delta). \tag{A.2}$$

Since the sum in (5.1) converges, (A.2) cannot hold for arbitrarily large m , and thus $r(m) < (1 + \delta)^2 \alpha m$ for large m . Similarly, if $0 < \delta < 1$, then $r(m) > (1 - \delta)^2 \alpha m$ for large m . This proves (5.3). \square

References

- [1] D. J. Aldous & G. K. Eagleson: On mixing and stability of limit theorems. *Ann. Probability* **6** (1978), no. 2, 325–331. MR-0517416
- [2] Margaret Archibald & Conrado Martínez: The hiring problem and permutations. *Proc. 21st International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2009)*, 63–76, Discrete Math. Theor. Comput. Sci. Proc. **AK**, Nancy, 2009. MR-2721502
- [3] N. H. Bingham, C. M. Goldie & J. L. Teugels: *Regular Variation*. Cambridge Univ. Press, Cambridge, 1987. MR-0898871
- [4] Andrei Z. Broder, Adam Kirsch, Ravi Kumar, Michael Mitzenmacher, Eli Upfal & Sergei Vassilvitskii: The hiring problem and Lake Wobegon strategies. *SIAM J. Comput.* **39** (2009), no. 4, 1233–1255. MR-2551799
Preliminary version in *Proc. Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2008)*, 1184–1193, ACM, New York, 2008. MR-2490490
- [5] William Feller: *An Introduction to Probability Theory and its Applications, Volume II*. 2nd ed., Wiley, New York, 1971. MR-0270403
- [6] J. Gaither & M.D. Ward: Analytic methods for select sets. *Probab. Eng. Inf. Sci.* **26** (2012), 561–568. MR-2983206

- [7] Allan Gut: *Probability: A Graduate Course*, 2nd ed., Springer, New York, 2013. MR-2977961
- [8] Ahmed Helmi, Conrado Martínez & Alois Panholzer: Hiring above the m -th best candidate: a generalization of records in permutations. *Proc. 10th Latin American Theoretical Informatics Conference (LATIN 2012)*, 470–481, Lecture Notes in Computer Science **7256**, Springer, Berlin, 2012. MR-2979069
- [9] Ahmed Helmi, Conrado Martínez & Alois Panholzer: Analysis of the strategy “hiring above the m -th best candidate”. *Algorithmica* **70** (2014), no. 2, 267–300. MR-3239792
- [10] Ahmed Helmi & Alois Panholzer: Analysis of “hiring above the median”: a “Lake Wobegon” strategy for the hiring problem. *Proc. Meeting on Analytic Algorithmics and Combinatorics (ANALCO 2012)*, 75–83, SIAM, Philadelphia, PA, 2012. MR-3240913
- [11] Ahmed Helmi & Alois Panholzer: Analysis of the “hiring above the median” selection strategy for the hiring problem. *Algorithmica* **66** (2013), no. 4, 762–803. MR-3071846
- [12] Svante Janson: Moments of Gamma type and the Brownian supremum process area. *Probability Surveys* **7** (2010), 1–52. MR-2645216
Addendum: *Probability Surveys* **7** (2010), 207–208. MR-2771251
- [13] Svante Janson, Tomasz Łuczak & Andrzej Ruciński: *Random Graphs*. Wiley, New York, 2000. MR-1782847
- [14] Abba M. Krieger, Moshe Pollak & Ester Samuel-Cahn: Select sets: rank and file. *Ann. Appl. Probab.* **17** (2007), no. 1, 360–385. MR-2292590
- [15] Abba M. Krieger, Moshe Pollak & Ester Samuel-Cahn: Beat the mean: sequential selection by better than average rules. *J. Appl. Probab.* **45** (2008), no. 1, 244–259. MR-2409324
- [16] Abba M. Krieger, Moshe Pollak & Ester Samuel-Cahn: Extreme(ly) mean(ingful): sequential formation of a quality group. *Ann. Appl. Probab.* **20** (2010), no. 6, 2261–2294. MR-2759734
- [17] *NIST Handbook of Mathematical Functions*. Edited by Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert & Charles W. Clark. Cambridge Univ. Press, 2010. MR-2723248
Also available as *NIST Digital Library of Mathematical Functions*, <http://dlmf.nist.gov/>.
- [18] J. Preater: Sequential selection with a better-than-average rule. *Statist. Probab. Lett.* **50** (2000), no. 2, 187–191. MR-1791513

Acknowledgments. This research was begun during a lecture by Conrado Martínez at the conference celebrating the 60th birthday of Philippe Flajolet in December 2008; Conrado talked about the hiring problem, and I got the basic idea of the method used here, and made some notes. It then took me almost 10 years to return to the problem and develop the ideas in the notes, with further inspiration from papers by Conrado and others that were written in the meantime. I thank Conrado Martínez for an inspiring lecture, and I find it fitting to dedicate this paper to the memory of Philippe Flajolet.

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS¹)
- Easy interface (EJMS²)

Economical model of EJP-ECP

- Non profit, sponsored by IMS³, BS⁴, ProjectEuclid⁵
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

²EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

³IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

⁴BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁵Project Euclid: <https://projecteuclid.org/>

⁶IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>